

Visualizing Jane Addams: Interactive Archives Through AI

Anupama Bhatta (axbhatta@umich.edu)

Christy Choi (choicr@umich.edu)

Haashim Mirza (haashim@umich.edu)

April 21, 2025

EXECUTIVE SUMMARY

Our capstone project aimed to transform how researchers engage with the Jane Addams Digital Edition through innovative visualization tools and intuitive interfaces. Working with Ramapo College of New Jersey, our team created three complementary tools that bring to life the complex narratives contained within this extensive historical archive.

The project addressed key challenges identified by the Jane Addams Papers Project, including the need for intuitive tools to explore relationships among the collection's over 55,000 documents and their associated metadata. Our solution integrates data visualization with natural language processing, resulting in three interconnected components: an interactive network visualization mapping Jane Addams' social and professional relationships, a topics-over-time visualization tracking thematic evolution in her work, and an LLM-powered query interface enabling conversational interaction with the archive.

These tools collectively transform static historical documents into dynamic, interactive data that can be explored through multiple dimensions. By visualizing connections between people and organizations, tracking the evolution of ideas chronologically, and enabling conversation-based exploration, our system makes the rich history of Progressive Era reform movements more accessible to both scholars and the general public.

This report details our methodology, technical implementation, key insights generated by these visualizations, and recommendations for future enhancements. Our work demonstrates how modern computational approaches can democratize access to complex historical collections while generating new insights about Jane Addams' pivotal role in early 20th century social reform.

INTRODUCTION & PROJECT BACKGROUND

The Jane Addams Papers Project, hosted at Ramapo College of New Jersey, stands as one of the most significant digital humanities initiatives documenting the Progressive Era in America. Jane Addams

(1860-1935), a pioneering social worker, settlement house founder, peace activist, and the first American woman to win the Nobel Peace Prize, left behind a rich legacy of correspondence, publications, and organizational records that illuminate critical developments in social reform, women's suffrage, labor rights, and the peace movement during the early 20th century.

The Jane Addams Digital Edition currently contains over 55,000 transcribed documents, along with extensive metadata captured in Dublin Core format and stored in a MySQL database. These records link to thousands of biographical entries for people, organizations, and events, providing crucial context for understanding the documents. This vast collection spans the years 1900-1930 and chronicles the formation of settlement houses, the professionalization of social work, Progressive reform movements, woman suffrage activism, child labor advocacy, and transnational peace efforts.

Despite the richness of this digital archive, the project team identified several challenges that limited its accessibility and utility for researchers:

1. The sheer volume of textual data, which made it difficult for users to identify patterns and relationships across documents
2. Limited tools for visualizing connections between people, organizations, and concepts
3. Inconsistent metadata that complicated comparative analysis
4. Barriers to entry for researchers without extensive technical knowledge

Our capstone project addresses these challenges by leveraging the power of interactive visualization and natural language processing. Working collaboratively with the Jane Addams Papers Project team, we developed three complementary tools:

1. A network visualization interface that maps relationships between individuals and organizations, revealing the evolving social and professional networks that shaped Addams' work
2. A topical timeline visualization showing how Addams' interests and advocacy evolved over time, highlighting key themes and how they intersected with historical events
3. An LLM-powered query interface that enables users to explore the collection through natural language questions, making the archive more accessible to researchers without specialized technical skills or pre-requisite knowledge.

This integrated suite of tools transforms how users interact with historical archives, moving beyond traditional keyword searching to enable intuitive exploration through visualized relationships and conversation-based discovery. Our approach demonstrates how digital humanities techniques can both democratize access to historical collections and generate new insights from well-studied materials.

PROJECT GOALS & OBJECTIVES

Our project was guided by several core objectives aimed at enhancing accessibility and analytical capabilities for users of the Jane Addams Digital Edition:

1. **Relationship Visualization:** Develop an interactive network visualization that reveals the complex social and professional relationships that shaped Jane Addams' work, allowing users to explore how these networks evolved over time and across different spheres of influence.
2. **Thematic Evolution Tracking:** Create a visualization tool that tracks how the themes and topics in Addams' writings and correspondence evolved chronologically, highlighting shifts in her focus areas and connecting these to key historical events.
3. **Natural Language Archive Exploration:** Implement an LLM-powered interface that allows users to ask questions about the archive in conversational language, making the collection accessible to researchers without technical expertise in database querying.
4. **Metadata Standardization:** Address inconsistencies in the existing metadata to enable more reliable cross-document analysis and visualization.
5. **User-Centered Design:** Ensure all tools are intuitive and responsive for users with varying levels of technical expertise, from undergraduate students to professional researchers.
6. **Scalable Implementation:** Design solutions that can accommodate the ongoing growth of the digital edition as more documents continue to be added.
7. **Knowledge Discovery:** Facilitate new insights about Addams' work and influence by revealing patterns and connections that might not be apparent through traditional document-by-document research.

These objectives align with the Jane Addams Papers Project's mission to serve as a laboratory for undergraduate research while making this important historical collection more accessible to the public. By transforming how users interact with the archive, our project supports both educational goals and scholarly research, enabling new ways of understanding Jane Addams' pivotal role in Progressive Era reform movements.

METHODOLOGY

Our project employed a multifaceted methodology that combined data analysis, visualization development, and natural language processing. Each component required specific approaches tailored to the unique challenges of working with historical archives.

Data Assessment & Preparation

The first phase of our project involved a comprehensive assessment of the Jane Addams Digital Edition database. We analyzed the structure of the MySQL database, identified metadata inconsistencies, and developed cleaning protocols to standardize the data.

For the network visualization component, we faced several specific challenges:

- Extracting meaningful relationship data from document titles and metadata
- Standardizing name variations across the collection (e.g., "Miss Addams" vs. "Jane Addams")
- Identifying non-person entities that should be excluded from the network
- Categorizing relationships by type (peace work, social reform, etc.)

We developed a systematic approach to extract sender-receiver pairs from document titles, using pattern matching and named entity recognition techniques. This extraction process included:

- Identifying common patterns in correspondence titles (e.g., "Person1 to Person2")
- Creating rules to standardize name variations
- Developing exclusion lists for non-person entities
- Implementing verification steps to ensure data quality

For text data used in the topics-over-time visualization, we focused on extracting thematic metadata from the Jane Addams Digital Edition. Our approach involved several steps to ensure the data was appropriately cleaned and structured for analysis:

- **Data Loading and Initial Exploration:** We began by downloading and loading the existing dataset from a CSV file. An initial exploration was conducted to understand the structure of the data and identify any immediate issues, such as missing values or malformed data entries.
- **Tag Filtering and Standardization:** While the dataset contained various thematic tags, we focused on filtering out specific tags that were not relevant to our analysis, such as administrative tags ('DHLab', 'help', etc.). We standardized these tags to ensure uniformity across the dataset.
- **Date Handling and Year Extraction:** Our next step was to handle the date information associated with each document. We extracted year data from the date fields for accurate temporal alignment, manually correcting entries where necessary to maintain data integrity. This was crucial for analyzing trends over time.
- **Aggregation and Longitudinal Analysis:** We grouped the filtered data by year and tag name to calculate the frequency of each theme appearing in the documents year over year. This aggregation enabled us to conduct a longitudinal analysis of thematic trends.
- **Manual Corrections and Quality Assurance:** Certain document entries required manual intervention to ensure the year and date align correctly with known historical data. We implemented corrections to specific entries based on item identifiers to maintain consistency.
- **Visualization Preparation:** We prepared the dataset for visualization using Plotly and Dash, allowing for interactive exploration of the themes over time. The setup involved creating a responsive and user-friendly interface to examine how specific themes evolved throughout Jane Addams' career and aligned with key historical events.

This structured approach allowed us to systematically extract and process thematic data from a large volume of text, facilitating the creation of a dynamic topics-over-time visualization.

The LLM query interface required preparing the document corpus for semantic search by creating vector embeddings over 7000 documents in the archive. We used OpenAI's "text-embedding-3-large" to convert all the text into numerical representations which can then be used to perform vector search. We ended up leveraging "faiss"¹ an open source library developed by META for performing efficient similarity search on dense vector representations. Given that the text also contained HTML, we leveraged BeautifulSoup to clean up and remove all html anchor tags and text within them to ensure that the vector embeddings were

¹ <https://github.com/facebookresearch/faiss>

as clean as possible for future similarity search between the user's question and the final AI response. These were finally persisted as a faiss index file for future use without having to re-index all the data.

Network Visualization Development

The network visualization component was developed using Plotly, a powerful visualization library that enables interactive web-based visualizations. Our approach involved:

1. **Data Processing Pipeline:** We created a pipeline that processed the extracted relationship data, calculated connection frequencies, and prepared the network structure for visualization.
2. **Bidirectional Network Modeling:** Rather than treating connections as one-directional (sender to receiver), we implemented a bidirectional network that represents the mutual nature of correspondence relationships. This approach:
 - a. Treats "Jane Addams to Person X" and "Person X to Jane Addams" as part of the same relationship
 - b. Combines interaction counts in both directions to represent total correspondence volume
 - c. Better represents the reciprocal nature of social and professional relationships
3. **Dynamic Time-Based Filtering:** We implemented year-based filtering to allow users to see how Jane Addams' network evolved over time from 1901 to 1935.
4. **Categorical Classification:** We developed a classification system that assigned Jane Addams' correspondents to five categories based on the nature of their relationship:
 - a. Peace Work
 - b. Social Reform
 - c. Political Activism
 - d. Personal Relations
 - e. Academic Work
5. **Historical Context Integration:** To ensure accurate categorization beyond what automated analysis could provide, we incorporated domain expertise from the Ramapo team through a weighting system. This allowed us to correctly categorize relationships where document metadata alone might be misleading - for instance, identifying Mary Rozet Smith as Addams' lifelong partner (Personal Relations) and properly classifying key peace activists like Emily Greene Balch despite varied document topics.
6. **Intuitive Visual Encoding:** The visualization uses several visual elements to convey information:
 - a. Node size represents the frequency of interactions with Jane Addams
 - b. Node color indicates the relationship category
 - c. Node positioning uses a force-directed graph layout optimized for readability
 - d. Hover functionality provides detailed information about each correspondent
7. **Rich Interactive Information Display:** When users hover over a node representing a correspondent, they receive a comprehensive information card displaying:
 - a. Full name of the correspondent
 - b. Total number of interactions with Jane Addams in the selected year
 - c. Relationship category with color-coding
 - d. Active correspondence period (years of first and last documented interaction)
 - e. Total lifetime interactions across all years

- f. Top topics discussed in their correspondence, derived from document tags
 - g. This contextual information helps researchers quickly understand the nature and significance of each relationship without requiring further database queries.
8. **Interactive Navigation:** We implemented play/pause functionality and a timeline slider to allow users to explore the network's evolution chronologically.

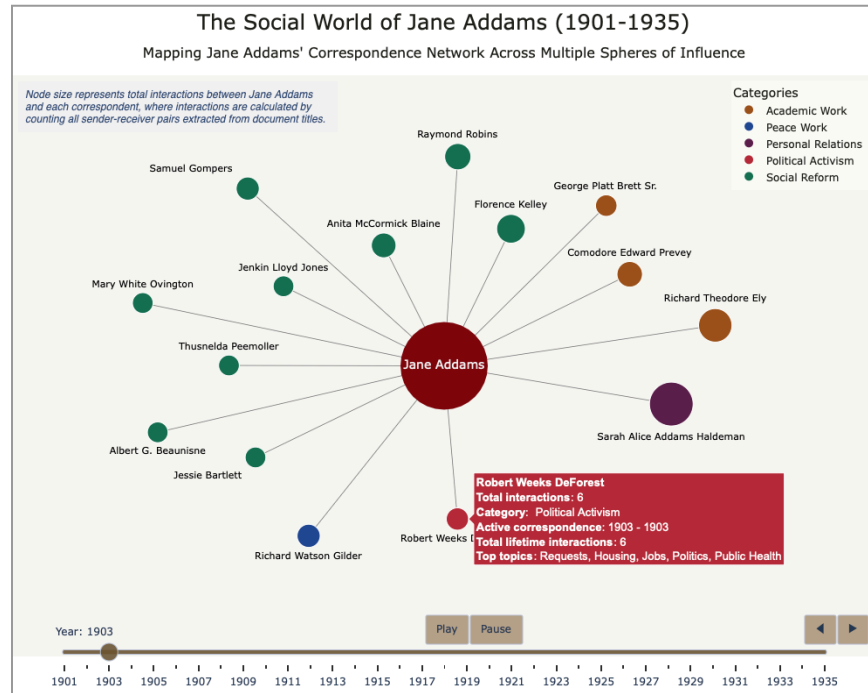


Figure 1: Interactive visualization of Jane Addams' correspondence over time.

Topics-Over-Time Visualization

The topics-over-time visualization component was developed to track thematic shifts in Jane Addams' work over time, using Plotly and Dash for interactive visualization. Our approach involved:

- **Data Processing Pipeline:** We established a comprehensive pipeline for processing thematic data, beginning with the extraction and cleaning of metadata tags associated with each document. This involved standardizing tag terminology to ensure consistency throughout the dataset.
- **Temporal Frequency Analysis:** By aggregating the frequency of each tag by year, we were able to track how often themes appeared over the years, providing insights into Jane Addams' evolving focus areas.
- **Interactive Visualization Setup:** Utilizing Plotly and Dash, we created a web-based, interactive visualization interface. This interface allows users to explore thematic trends, select and compare multiple themes, and observe how these themes intersect with significant historical events over time.
- **Intuitive Visual Encoding:** The time series graph uses color and line styles to differentiate between themes, with hover-over functionality to reveal additional information such as document titles for each year.

- **Rich Interactive Features:**
 - **Tag Dropdown Menu:** Users can select specific tags for display, with options to choose multiple tags for comparative analysis.
 - **Select All Button:** A single-click option to visualize all available tags simultaneously, providing a comprehensive view of document trends.
 - **Geographical Filter:** An option to filter and display only geographical tags, supporting focused analysis on location-based topics.
 - **Hover Data Display:** As users hover over points in the graph, document titles are displayed below, with clickable links to view the full documents.
- **Interactive Navigation:** Features such as a zoom function, which allows users to focus on specific time periods, and a reset view option for a full dataset overview, ensure an intuitive exploration process for users.

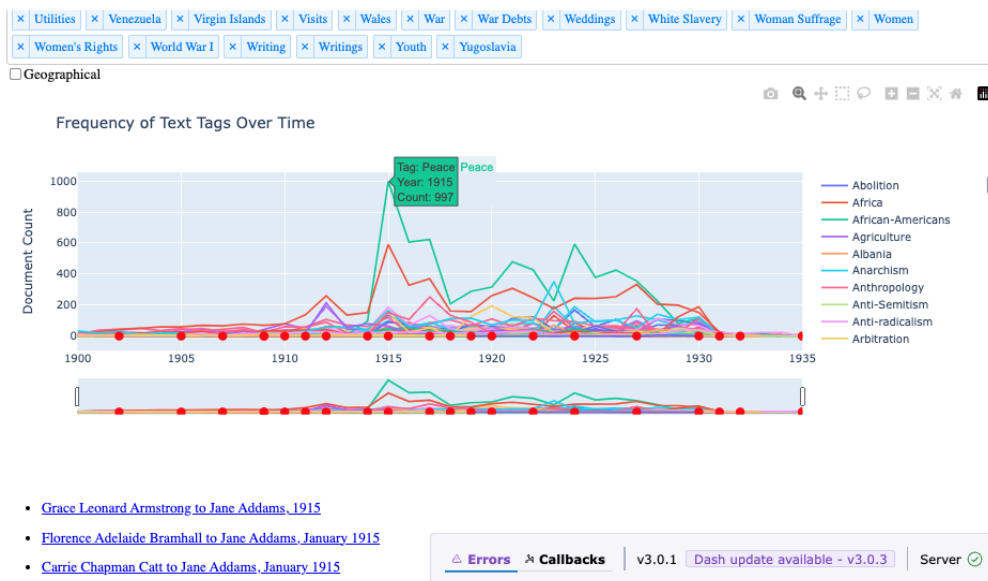


Figure 2: Interactive visualization showing the frequency of text tags over time in the Jane Addams Digital Edition.

This dynamic and interactive visualization empowers researchers and the general public to gain a deeper understanding of the thematic evolution within the Jane Addams archive, making connections between her documented work and broader historical contexts readily accessible and engaging.

LLM Query Interface

LLMs were a significant point of exploration. Given that we had the vector embeddings, we needed to explore a way to allow users to interact with the indexed data. We ended up leveraging LangChain, an open source library designed to aid in GenAI orchestration. We leveraged its “langchain_openai” and “langchain_community” submodules to allow us to orchestrate the whole conversational chatbot. This works by performing cosine similarity on the input question along with all the vector embeddings in our vector store, retrieving the top k documents and embedding them into the prompt which will be passed onto the AI with the past user and AI responses. Based on this updated context, the AI would respond and

give tailored responses based on the knowledge provided to it in the documents and in the system prompt, making it aware of who Jane Addams is.

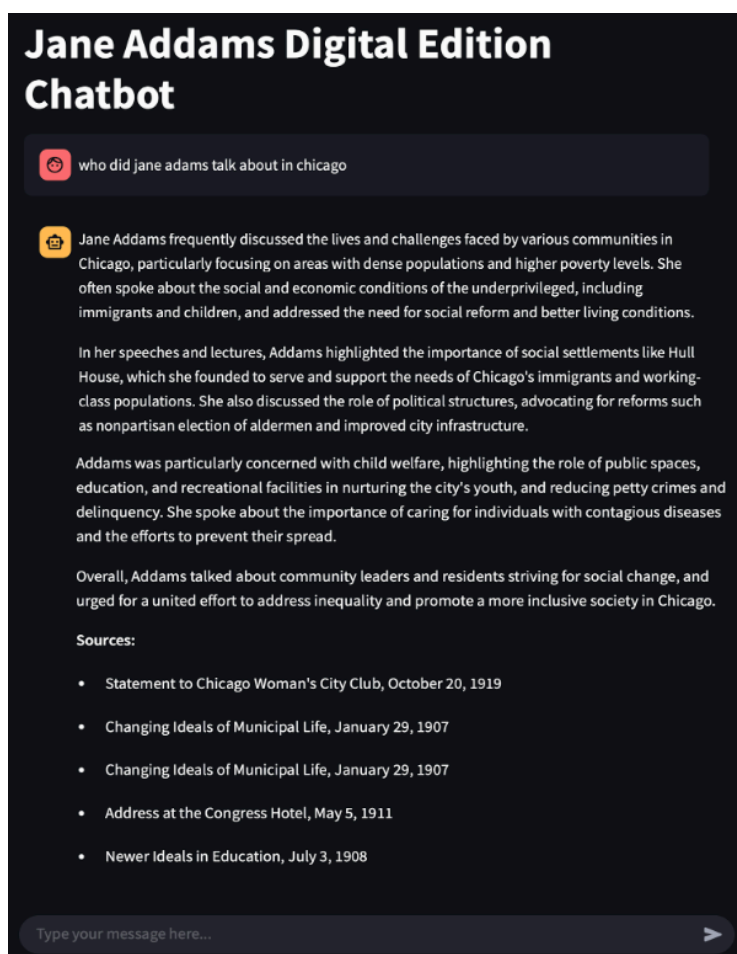


Figure 3: Example input and output of the Jane Addams Digital Edition Chatbot.

Iterative Development Process

1. Our development process was highly iterative, guided by regular feedback from the Jane Addams Papers Project stakeholders at Ramapo College. This collaboration included:
2. **Regular Check-ins:** We held periodic meetings with the project team to demonstrate progress and gather feedback.
3. **Functionality Reviews:** The Ramapo team provided valuable insights on the accuracy of data representation and historical context.
4. **Usability Feedback:** Project stakeholders tested early versions of the tools and provided suggestions for improving the interface and functionality.
5. **Feature Prioritization:** Based on stakeholder feedback, we adjusted our development priorities to focus on features with the greatest scholarly and educational value.

This iterative approach ensured that our technical solutions remained aligned with the scholarly needs of the Jane Addams Papers Project. After each feedback session, we made targeted adjustments to improve

accuracy, usability, and analytical capabilities. This collaborative process was instrumental in developing tools that not only showcased technical innovation but also provided genuine value for historical research and education.

RESULTS & DELIVERABLES

Our project has yielded three complementary tools that enhance the Jane Addams Digital Edition by revealing patterns and connections across the archive's extensive collection of documents. Each component addresses different analytical needs while contributing to a more comprehensive understanding of Jane Addams' life and work.

Network Visualization: Social Landscapes Revealed

The network visualization illuminates Jane Addams' world in ways traditional research methods cannot. By modeling relationships as bidirectional connections and employing an intuitive visual language of size, color, and position, we've made visible the otherwise hidden architecture of Addams' social influence.

What emerges is a striking portrayal of how Addams' network evolved over time. Most notably, we observed a dramatic shift toward peace activism connections following 1914-1915, with figures like Emily Greene Balch and Rosika Schwimmer becoming increasingly prominent. The visualization reveals "bridge figures" like Florence Kelley who connected Addams to multiple communities, showing how social reform movements were interconnected through key individuals rather than operating in isolation.

The year-by-year exploration feature allows researchers to track how major historical events reshaped Addams' correspondence patterns. During periods of legislative activism, her political connections intensified, while international crises triggered increased engagement with her peace network. These patterns would be difficult to discern through traditional document-by-document research but become immediately apparent through our visualization's temporal navigation system.

Topics-Over-Time Visualization

The topics-over-time visualization provides a dynamic exploration of Jane Addams' evolving thematic focus, offering insights into the shifting priorities and historical contexts of her extensive work. This tool transforms static archival data into a living timeline of ideas, allowing researchers to visualize the progression of key themes over several decades.

Through this visualization, we can observe how certain themes, such as settlement work and labor advocacy, gained prominence in the early 1900s, corresponding to Jane Addams' foundational years at Hull House. As the timeline progresses, a marked transition toward themes of peace activism and international diplomacy is apparent, particularly during and after the aftermath of World War I.

The interactive nature of the visualization allows users to filter tags, compare thematic trajectories, and focus on geographical tags to understand the regional aspects of Addams' influence. Hover functionality and detailed document title displays provide deeper insights into specific years, making it easier to connect thematic trends with significant historical events and Addams' key publications.

This component of our project not only reveals broad patterns in Addams' thought and action but also facilitates a nuanced understanding of how her interests were shaped by, and contributed to, the social and political currents of her time. By bridging micro-level document content with macro-level historical context, the topics-over-time visualization enriches both scholarly research and public education regarding the legacy of Jane Addams.

These findings would be challenging to compile through traditional archival approaches, yet through this visualization's interactive features, we uncover a rich tapestry of ideas that align with the complex history of the early 20th century.

LLM Query Interface

Our final AI deliverable has all the indexing pipeline present in case Ramapo wants to update the vector store with additional documents they will transcribe in the future. Alongside this, we have also given them a fully functional application with a working frontend and backend which also logs any errors and scales according to the needs of their website. They can ultimately embed the streamlit app as an iframe within their own website, making it an interactive component that helps people get to know more about the data present in the repository.

Implementation Resources

Our deliverables include a comprehensive GitHub repository with all visualization components, data preprocessing scripts, and integration documentation. We will support the Ramapo team throughout the implementation process, ensuring these tools become a permanent part of the Jane Addams Digital Edition ecosystem.

CHALLENGES & SOLUTIONS

Throughout the development process, we encountered several meaningful challenges that shaped our approach to visualizing the Jane Addams Digital Edition.

Technical Challenges in Network Visualization

One of our primary challenges involved extracting accurate relationship data from the document database. While the Jane Addams Papers Project maintained detailed metadata about each document, the sender-receiver relationships were embedded within document titles rather than existing as separate database fields. We developed pattern matching techniques to identify these relationships from title

formats like "Jane Addams to Florence Kelley," along with name standardization to account for variations in how individuals were referenced across documents.

The visual representation of these networks presented another significant challenge. Creating a readable visualization that could accommodate dozens of connections while maintaining clarity required careful design decisions. We experimented with various node positioning approaches before developing a category-based layout that balanced visual clarity with accurate relationship representation. Synchronizing the interactive elements – including year filtering, animation controls, and information display – required additional refinement to ensure a seamless user experience.

These challenges ultimately led to a more robust visualization system that effectively communicates the complex social landscape surrounding Jane Addams while remaining accessible to users with varying levels of technical expertise.

Topics-Over-Time Visualization Challenges

Developing the topics-over-time visualization presented several key challenges that required strategic solutions to ensure the accuracy and usability of the tool.

One prominent challenge was the inconsistency in metadata, particularly concerning date recordings across different documents. Variations in date formats and occasional errors hindered our ability to perform accurate temporal alignments necessary for reliable longitudinal analysis. To address this issue, we undertook a data cleaning process involving manual corrections to date entries, which ensured temporal accuracy across the dataset. This careful attention to detail allowed us to provide a faithful representation of thematic shifts over time, enhancing the visualization's ability to reveal meaningful historical insights.

Additionally, the extensive variety of thematic tags posed another significant challenge, resulting in a cluttered visualization that was difficult to navigate and interpret. The overwhelming number of tags obscured thematic clarity and insight extraction. To tackle this, we implemented an interactive filtering option that allows users to select specific tags they wish to explore, thereby making the visualization more user-friendly and focused. Furthermore, we created a geographical tags group to help users easily identify and analyze location-based themes. By grouping all geographical tags together, the visualization offers a focused analysis of regional influences and Addams' geographical reach, adding another dimension of interpretive power to the tool.

These solutions not only addressed the immediate challenges faced but also significantly enhanced the overall functionality and value of the topics-over-time visualization. By ensuring both metadata consistency and user-customizable views, we improved the tool's accessibility and analytical power for diverse audiences, including scholars and educators.

LLM Project Challenges

The LLM side of things also had some significant challenges associated with it. Firstly, a component that was being proposed was to use LLMs to automate the transcription process. However, by even checking

multiple LLMs and vendors, we quickly realized that LLMs currently don't have large enough output context windows to reliably transcribe all the text present in some of their corpuses which is why we had to abandon that idea.

Additionally, LLMs have really high associated costs attached to it. Given how Ramapo had limited financial resources to support this project, we had to rely on our own resources to create an application that could demonstrate how this would work for the project. For example, "text-embedding-3-large" costs around \$0.13 per 1 Million tokens (which is about ~600,000 words but a lot of tokens are also used by entities such as spaces and punctuation marks). However, we had around 50,000 documents and \$20 worth of budget for the whole project and we ran out of the whole amount at around 7000 documents indexed.

Initially we tried to use open source alternatives. Huggingface has a host of vector embedding and inference models which we used, however, they were taking much longer than anticipated and would crash our local PCs due to how memory intensive the process is. While ARC was also an option, there were issues in making the environment work in a manner which could be replicated by Ramapo which is why we went the more traditional API route. We were able to get some additional funds for demoing the chatbot but further AI API calls are also expensive (even when using smaller models like gpt-4o-mini). The scale of the project could have been increased with more financial resources to make sure it was fully populated with all the data Ramapo had.

Throughout the development process, we maintained regular communication with the Ramapo team, using their feedback to guide our approach and ensure the resulting tools would meet the needs of both researchers and students engaging with the Jane Addams Papers Project. This collaborative process was essential for navigating the technical and conceptual challenges of transforming historical documents into interactive visualizations.

FUTURE RECOMMENDATIONS

Based on our experience developing these visualization tools, we have identified several promising directions for future enhancement and expansion of the Jane Addams Digital Edition.

Network Visualization Enhancements

While the current network visualization provides valuable insights into Jane Addams' social and professional relationships, several enhancements could further increase its analytical power:

1. **Multi-level Network Analysis:** Expanding the visualization to show connections between Jane Addams' correspondents (not just between Addams and each individual) would reveal additional dimensions of these social networks.
2. **Geographic Overlay:** Integrating location data to show the spatial distribution of Addams' network would illuminate how her influence extended across different regions and countries.

3. **Thematic Filtering:** Developing more granular filtering options based on document tags would allow researchers to focus on specific aspects of Addams' work, such as child labor reform or peace activism.

Tags-Over-Time Visualization Enhancements

While the current topics-over-time visualization offers valuable insights into the thematic evolution of Jane Addams' work, several enhancements could further enrich user analysis and interpretation:

- **Dynamic Theme Annotation:** Incorporating automatic annotations that provide context for significant spikes or shifts in thematic trends, potentially tied to specific historical events or influential publications, could deepen users' understanding of these changes.
- **Advanced Aggregation Techniques:** Enhancing data aggregation methods by incorporating advanced statistical techniques, such as sentiment analysis or thematic clustering, would allow users to explore not just the frequency but the intensity and nature of themes over time.
- **Interactive Historical Context:** Developing an interactive layer of historical commentary that users can toggle on and off, allowing them to see explanations of how external events may have influenced thematic developments within the Jane Addams collection.
- **Correlated Theme Exploration:** Implementing functionality to highlight correlations between themes that often appear together, giving users deeper insight into the interconnectedness of ideas and their collective impact on Addams' advocacy and activism.

Cross-tool Integration

The current suite of tools offers complementary perspectives on the Jane Addams archive. Future development could strengthen the connections between these components:

1. **Unified Dashboard:** Creating an integrated interface where selections in one visualization automatically update the others would enable more seamless exploration.
2. **Enhanced Search Integration:** Deeper integration with the archive's existing search functionality would create a more cohesive user experience.
3. **Search-enhanced LLM:** In addition to adding context from Jane Addams' documents provided by Ramapo, adding a web search tool integrated within the LLM would have allowed for more real time knowledge sharing and more accurate results.

Expanding the Approach

The methodologies developed for this project could be extended to other aspects of the Jane Addams Digital Edition or similar historical archives:

1. **Organization Networks:** Applying similar visualization techniques to organizational relationships would reveal institutional connections that shaped Progressive Era reform movements.

2. **Comparative Networks:** Developing tools to compare Jane Addams' network with those of her contemporaries would provide valuable context for understanding her unique position in Progressive Era social movements.
3. **Pedagogical Applications:** Creating simplified versions of these tools specifically designed for classroom use would support the Jane Addams Papers Project's educational mission.
4. **Temporal Sentiment Analysis:** Incorporate sentiment analysis over time on Jane Addams' correspondence and writings to measure changes in tone and sentiment, providing insights into her emotional and rhetorical evolution in response to historical events.
5. **OCR based transcriptions:** Ramapo College is still mostly using human resources to transcribe the documents they have. Using state of the art tools like [Amazon Textract](#) would make their transcription process much more efficient.

These recommendations represent opportunities to build upon the foundation established by our current work, further enhancing the value of the Jane Addams Digital Edition as a resource for both researchers and the public.

CONCLUSION

The Jane Addams Papers Project represents an invaluable resource for understanding a pivotal era in American social history. Our visualization tools transform how researchers and students can engage with this rich archive, revealing patterns and connections that might otherwise remain hidden within thousands of individual documents.

Through the network visualization, we can now see how Jane Addams built and maintained relationships across different spheres of influence – from social reform to peace activism, from political engagement to academic work. The visualization reveals not just individual connections but the evolving structure of Progressive Era social movements, showing how key figures served as bridges between different communities and how Addams' network responded to historical events.

The topics-over-time visualization offers a dynamic lens through which the thematic evolution of Jane Addams' work can be explored, helping to contextualize her contributions within the broader tapestry of the Progressive Era. By tracing the frequency and prominence of various themes across decades, this tool reveals shifting priorities in Addams' advocacy, from her foundational work in settlement houses to her later emphasis on peace and international diplomacy. The visualization enables users to align thematic trajectories with key historical milestones, offering rich insights into how Addams' focus adapted to and influenced the social currents of her time. This interactive approach enhances our understanding of the interconnectedness of ideas and events, offering a nuanced perspective that enriches both academic research and public knowledge.

Finally, the Chatbot is a unique insight into understanding the documents present in the archive where people can add additional context to the texts in a manner where they don't have to go through all the documents to find answers, increasing accessibility and providing a more human interface to interact with the data.

Beyond the specific insights these tools provide about Jane Addams and her world, our project demonstrates the broader value of applying visualization and natural language processing techniques to historical archives. By transforming static documents into interactive knowledge systems, we enable new modes of exploration and discovery that complement traditional research approaches.

We are grateful for the opportunity to contribute to the ongoing development of the Jane Addams Digital Edition and to collaborate with the dedicated team at Ramapo College of New Jersey. We hope these tools will serve researchers, students, and the public for years to come, helping to ensure that Jane Addams' legacy continues to inform our understanding of social activism, reform movements, and civic engagement.

ACKNOWLEDGEMENTS

We would like to express our sincere gratitude to Dr. Cathy Moran Hajo, Director of the Jane Addams Papers Project, and Caitlin Biebrich, Assistant Editor, for their invaluable guidance, historical expertise, and consistent feedback that helped shape our visualizations and ensure their accuracy.

We are also deeply thankful to Anneliese Dehner, web developer for the Jane Addams Papers Project, for providing the SQL data export that formed the foundation of our project, as well as her technical insights throughout the development process.

We extend our appreciation to Dr. Elle O'Brien and Dr. Ceren Budak at the University of Michigan School of Information for connecting us with the Ramapo team and for their support and guidance throughout the capstone process.

Finally, we acknowledge the technological communities behind the open-source tools that made this project possible, particularly the developers of Plotly and Python.