Relation Extraction Across Entire Books to Reconstruct Community Networks: The AFFILKG Datasets

Erica Cai, Sean McQuade, Kevin Young, Brendan O'Connor
University of Massachusetts Amherst
140 Governors Drive Amherst, MA 01003
{ecai,spmcquade,keviny,brenocon}@umass.edu

Abstract

When knowledge graphs (KGs) are automatically extracted from text, are they accurate enough for downstream analysis? Unfortunately, current annotated datasets cannot be used to evaluate this question, since the knowledge graphs that they correspond to, constructed by mapping entities in the text to nodes and relations to edges, are typically highly disconnected, too small, or overly complex. To address this gap, we introduce AFFILKG, which is a collection of six datasets that are the first to pair complete book scans with large, labeled knowledge graphs. Each dataset features affiliation graphs, which are simple KGs that capture MEMBER relationships between PERSON and ORGANIZATION entities—useful in studies of migration, community interactions, and other social phenomena. In addition, three datasets include expanded KGs with a wider variety of relation types. Our preliminary experiments demonstrate significant variability in model performance across datasets, underscoring AFFILKG's ability to enable two critical advances: (1) benchmarking how extraction errors propagate to graph-level analyses (e.g., community structure), and (2) validating KG extraction methods for real-world social science research.

Keywords: knowledge graph extraction, OCR, relation extraction, social networks

1. Introduction

Extracting social networks from text is vital for social scientists aiming to understand how elite communities interact and how network structures evolve over time (O'Brien, 2025; Broad, 1996; Wellman and Wetherell, 1996; Kossinets and Watts, 2006). Traditionally, networks have been constructed through painstaking manual curation with high precision, for example in mapping financial acquisitions from news articles (Gugler et al., 2003; Clougherty et al., 2014) or interpersonal relationships in literary sagas (Mac Carron and Kenna, 2013). However, manual annotation remains labor-intensive and difficult to scale. Automated approaches are increasingly popular for their efficiency, but their evaluation is often insufficiently rigorous (Drury et al., 2022), raising concerns about their reliability for downstream graph analyses.

In parallel, hundreds of papers on relation extraction (RE), a.k.a. knowledge graph (KG) extraction, aim to automate the construction of these networks from text. In this approach, information is represented using triplets (e.g., ⟨BOB; MEMBER; IEEE⟩), where each triplet encodes a relation (MEMBER) between two entities (BOB, IEEE) that is expressed in the text. In the resulting KG, entities correspond to nodes and relations correspond to edges (Fig.1)(Ding et al., 2021; Nadgeri et al., 2021; Xu and Barbosa, 2019; Trisedya et al., 2019; Han et al., 2021). While early work found that large language models (LLMs) performed poorly on RE tasks (Jimenez Gutierrez et al., 2022), recent research demonstrates increasing success, even with limited supervision (Wadhwa et al., 2023; Wan et al., 2023; Zhou et al., 2024; Zhang and Soh, 2024; Tao et al., 2024; Xue et al., 2024; Zhang et al., 2024; Rajpoot and Parikh, 2023). However, a fundamental question remains: Do widely used evaluation strategies for RE and KG extraction (Zhao et al., 2024; Detroja et al., 2023) reliably predict model performance on downstream, graph-level analyses that actually matter in practice—for example, identifying key individuals or communities in a social network (Cai et al., 2024)?

Most existing NLP datasets for RE make it difficult to evaluate how KG extraction errors propagate to downstream graph analyses. Many such datasets label triplets drawn from collections of unrelated sentences or paragraphs, resulting in knowledge graphs that are fragmented and disconnected, since their entities and relations rarely form a coherent network structure (Zhang et al., 2017; Han et al., 2018; Riedel et al., 2010; Chen et al., 2021; Gardent et al., 2017; Huguet Cabot et al., 2023). Other datasets contain highly complex KGs with hundreds of relation types, or focus only on brief documents, leading to small or fragmented graphs even when all relations are extracted (Yao et al., 2019). Furthermore, in datasets that do yield connected KGs, extraction models often achieve low performance (typically below 50 F1 points), making it difficult to systematically assess the impact of extraction errors on real-world graph-based analyses or practical applications (Luan et al., 2018; Jain et al., 2020).

To address these challenges, we introduce AF-

Proceedings of the Fifteenth Language Resources and Evaluation Conference (LREC 2026), pages 7744–7754
11–16 May 2026.

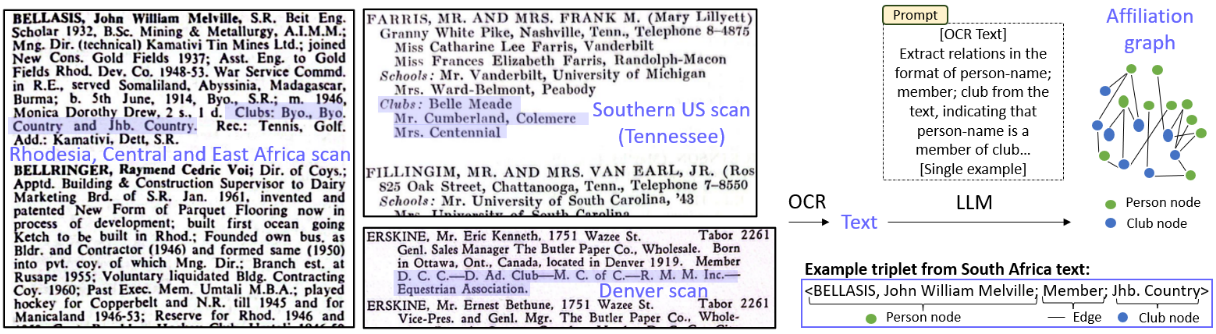

Figure 1: Pipeline of converting scans of historical text to affiliation graphs, where highlighted text indicates clubs.

FILKG[1], the first collection of datasets that pair full scanned books—along with their text extracted via optical character recognition (OCR)—with large, labeled knowledge graphs. Each of the six datasets in AFFILKG provides an affiliation graph, which is a bipartite KG where edges denote MEMBER relationships between PERSON and CLUB entities. These graphs are particularly valuable for studying community cohesion, interaction dynamics (Prem Sankar et al., 2015; Tisch and Ischinsky, 2023), and broader topics such as social stratification and network evolution (Corradi et al., 2024; Bichler et al., 2015; Young and Ready, 2015; Grund and Densley, 2015). Three datasets are further expanded to include triplets of 4 to 11 additional relation types, such as COLLEGE and YACHT NAME, enabling richer relational analyses.

The texts in AFFILKG are unique archival repositories of information on elite families in specific locations. Known as 'social registers', these texts document elite individuals from diverse geographic regions and time periods, including the Southern US (Whitmarsh, 1951), Denver (1931), Rhodesia (1963), Boston (1903), Miami (Jackson, 1965), and New York City (1892). Social register texts served to connect elite families with one another in a common directory of an area, typically listing family members, addresses, club memberships and educational backgrounds. As a historical text, a given social register thus provides a comprehensive record of a local social network, making it ideal for evaluating RE methods their downstream impact on graph analyses. Affiliation graph triplets were generated either through manual annotation or by applying regular expressions and rule-based algorithms, which were rigorously validated against manually constructed ground truth.

AFFILKG addresses critical limitations in existing datasets and provides a foundation for studying how RE errors impact downstream graph analyses. We present preliminary results using both widely used NLP evaluation metrics and higher-level, graph-based analyses relevant to real-world applications. Our findings indicate that, while extraction performance can be high enough to yield meaningful insights, there is also substantial variation across settings. This suggests that further investigation into the propagation of RE errors within graph analyses could uncover important patterns and provide actionable knowledge for practical use.

2. AFFILKG Dataset Construction

AFFILKG contains six datasets pairing scanned historical social register books with large, labeled knowledge graphs. Each individual dataset includes an affiliation graph, and three provide expanded KGs with additional relation types. All datasets include high-quality scans, OCR text, and graph annotations, enabling robust evaluation of relation extraction and higher-level graph analyses (Fig. 1).

Texts. The books in AFFILKG are archival social registers, which serve as historical "who's who" directories that catalog elite individuals and their affiliations within a locality at a specific time, and have been used by social scientists aiming to represent elite populations and their networks (O'Brien, 2024; Broad, 2020; Baltzell, 1958). These sources connect prominent figures to clubs, educational institutions, and other organizations, and often include additional details such as household members, life events, and possessions, offering a rich context for studying elite community structure and interaction (O'Brien, 2025; Broad, 1996). Each register serves as a comprehensive record of a local social network, defined according to the authors' notion of prominence. Despite widely recognized value of social register texts for over half a century (O'Brien, 2024; Broad, 2020; Baltzell, 1958), they have been used sparingly because they are very laborious to collect and use for applied research. A vast amount of knowledge about elite communities has thus been missing from social science, despite

[1]https://doi.org/10.5281/zenodo.15427977

Loc./Year	# pages	# nodes (person / org)	# edges
NYC/1892	299	5553 / 178	14061
Boston/1903	150	2534 / 188	6228
Denver/1931	34	358 / 248	1352
S. US/1951			
All (13 states)	826	6188 / 3107	13262
Alabama	49	401 / 236	753
Arkansas	22	167 / 83	292
Florida	164	1395 / 728	3208
Georgia	100	685 / 289	1275
Kentucky	28	227 / 121	501
Louisiana	67	544 / 206	1328
Mississippi	26	149 / 98	293
North Carolina	51	375 / 224	747
South Carolina	33	235 / 145	515
Tennessee	47	399 / 174	872
Texas	153	993 / 394	2026
Virginia	55	401 / 274	995
West Virginia	31	217 / 135	457
Rhod./1963	169	898 / 767	2367
Miami/1965	503	3775 / 1471	11830

Table 1: Details about *Denver*, *Southern US* (13 states aggregated and separate), individual southern states, *Rhodesia*, *Miami*, *Boston*, and *New York City* text/affiliation graph pairs.

Loc./Year	# nodes (person / other)	# edges	# rel types
NYC/1892	13054 / 5182	29263	4
Boston/1903	6748 / 4619	18945	5
Miami/1965	6340 / 8065	30845	11

Table 2: Details about *Miami*, *Boston*, and *New York City* expanded KGs.

the widely recognized importance of club membership to the study of elite communities (Bond, 2012; Weber, 1947; Mills, 1956; Keller, 1963).

The six registers are selected to support ongoing research into historical elite networks, drawing from several U.S. (and one non-U.S.) locations, over the late 19th to mid-20th centuries (Table 1).[2] The texts themselves represent some of the variety in layout and writing styles used for social registers (Fig. 1).

Bipartite affiliation graphs (Table 1). Each book yields a bipartite affiliation graph of MEMBER relational edges between PERSON and CLUB entity nodes (Breiger, 1974; Lattanzi and Sivakumar, 2009), represented as a collection of ⟨PERSON; MEMBER; CLUB⟩ triplets: for example, ⟨"*BELLASIS, John William Melville*"; MEMBER; "*Byo.*"⟩ in Fig. 1's *Rhodesia* snippet. Affiliation graphs are widely used to model memberships in elite groups, terrorist organizations, and other social networks, and are well-suited for studying diverse social science questions (Prem Sankar et al., 2015; Tisch and Ischinsky, 2023; Corradi et al., 2024; Isah et al., 2015; Holme et al., 2007; Marotta et al., 2014; Duxbury, 2020).

Expanded knowledge graphs (Table 2). For the Miami, Boston, and New York City datasets, we further expanded the KGs by including additional relation types, such as COLLEGE, EVENT, PLANE NAME, and YACHT NAME, reflecting the complexity commonly found in NLP research (Doddington et al., 2004; Riedel et al., 2010; Zhang et al., 2017; Han et al., 2018; Yao et al., 2019; Gardent et al., 2017; Huguet Cabot et al., 2023). Table 2 summarizes the number of relation types incorporated into the expanded knowledge graphs for the *Miami*, *Boston*, and *NYC* social register texts. Specifically, the relation types for *Miami* include: Address, Address - Northern, Address - Summer, Address - Permanent, Address, Yacht, Clubs, College, School, Cruiser, Yacht, Plane, Sailer, and Racer. In the *Boston* data, we included Address, Club, Phone, College, and Event relation types. For *NYC*, the relation types expanded to Address, Club, College, and Event.

Construction of ground truth triplets. To construct ground truth triplets, we developed a semi-automatic approach based on the observation that each book, despite containing relatively free-form text, followed several complicated patterns when presenting information. These patterns varied significantly across social register texts, requiring tailored strategies for each dataset. To capture the patterns, one author designed regular expressions and rule-based algorithms to extract triplets of ⟨PERSON, MEMBER, CLUB⟩ relations for affiliation graphs, or of multiple relationship types for expanded KGs. For validation, two additional authors independently annotated four to six randomly selected pages from different sections of each book, resulting in approximately 200 ground truth triplets per book. To ensure consistency across all datasets, all authors first discussed and agreed on a standardized labeling scheme prior to annotation.

Because the structure and conventions for encoding affiliations and relationships varied considerably between datasets, our extraction rules were customized for each book. For the Southern US and Miami texts, we linked adults to clubs based on the presence of titles (e.g., "Mr.", "Mrs."); in Rhodesia, we identified clubs following specific string markers and parsed lists with varied delimiters. Boston and NYC data relied on abbreviations and line formatting, and required context-sensitive parsing for more complex cases such as couples or

[2]Our dataset, drawn from six books, actually includes 18 separate text/network pairs, since the *The Southern Social Register* (Whitmarsh, 1951) contains 13 individual sections for each state, which we separate.

household groupings. Addresses, colleges, phone numbers, and events were extracted using pattern matching for characteristic markers and by resolving ambiguities in name and household references, such as differentiating households by last name appearance or adult status. When matching was ambiguous, we used default associations consistent with social register conventions.

Annotation quality and validation. Most AFFILKG annotations were generated using a semi-automatic pipeline: customized rule-based extraction, refined through careful manual review. For each book, the author responsible for rule design examined hundreds of pages of text to develop rules that captured the core structural conventions, of which there were a countable number for each dataset. While the vast majority of entries followed these primary patterns, we also checked for and addressed occasional less-common cases, which were infrequent.

To formally assess extraction quality, we randomly sampled 200 triplets per dataset for manual annotation by two additional authors and compared these to the rule-based output. For all datasets except *NYC*, this spot-checking confirmed 100% accuracy. In the *NYC* dataset, a very small number of rare cases, often involving addresses that spanned many lines or household boundaries, occasionally introduced ambiguity. To supplement the sampling, we also manually inspected many additional samples, until further inspection yielded no unseen cases. For the *Denver* dataset, which is smaller, two authors manually annotated the entire dataset without any automation. This protocol gives us high confidence that the resulting gold data approaches manual-quality annotation across all books.

To enable the application of rule-based algorithms to the text, we first converted the scanned book images into machine-readable text using Claude Sonnet (Anthropic, 2024) as an OCR model. We then manually corrected transcription errors by systematically comparing the OCR output with the original scanned images to ensure accuracy.

Significant manual effort required. The regular expressions and rule-based algorithms were highly complex, with up to 540 lines of code, due to the variability and intricacy of the text formats. In the *Southern US* states and *Miami* datasets, entire families were described in contiguous blocks of text, using complex structural patterns based on titles to identify both people and their club memberships. In the *Rhodesia* dataset, the varying formats for club memberships within text blocks required flexible regular expressions. For *Boston* and *New York City* books, the text formats lacked explicit family delineations. Family members were listed sequentially—adults first, then children, all on separate lines—with addresses, phone numbers, and events spanning multiple lines on the right side of the page. Club information for a person might appear on the same line as their name or on the following line next to another individual's name, linked correctly only through punctuation like brackets.

Constructing these regular expressions and annotating parts of the ground truth demanded significant manual effort, in particular for the older social register texts. This labor-intensive process underscores the need for more efficient methods to reduce manual intervention. We propose using the high quality labels generated here as benchmarks to evaluate alternative approaches that require less manual effort.

3. Benchmarking LLMs for Extracting Local and Global Structure

AFFILKG is the first ground truth resource to support both micro-level evaluation at the edge level—consistent with evaluations in the NLP literature—and macro-level evaluation based on high-level graph structure, which is crucial for real-world applications. In our preliminary evaluation, we demonstrate AffilKG's unique value by assessing both granular extraction accuracy and higher-level graph properties, exemplified by analysis of club membership sizes.

Scope of experimental evaluation. We benchmark three datasets (*Denver*, *Tennessee*, *Rhodesia*) with bipartite PERSON–CLUB affiliation graphs, chosen for their conceptual simplicity, which enables straightforward, interpretable evaluation of extraction methods and their impact on real-world downstream graph analyses. In contrast, our expanded-relationship datasets (Miami, Boston, NYC) present greater annotation complexity and enable more advanced, real-world analyses that require further methodological development. By establishing clear baselines on these simpler cases, we lay the foundation for future work with the complex, multi-relational AffilKG datasets. Rigorous evaluation of high-level graph analyses on these multi-relational datasets is left for future work, as they present new and different methodological challenges than the bipartite case.

We emphasize that our resource release includes all annotated datasets, including those with multi-relational KGs, to foster future research and community-driven benchmarking.

Experimental pipeline. To evaluate extraction methods under these conditions, we use the following experimental pipeline:

Step 1: OCR. *Input: Image scans; Output: Text.* A vast amount of text, encompassing mil-

Dataset	LLM Prompt
Denver	[text entry] Extract relations in the format of person-name; member; club from the text, indicating that person-name is a member of club. In the text, — indicates a different club. List each relation as a bullet point. An example is: [example text entry and outputted relation triplets]
Southern US	[text entry] Extract relations in the format of person-name; member; club from the text, indicating that person-name is a member of club. Schools are not clubs. If no title such as 'Mr.', 'Mrs.', 'Miss', 'Rev.', 'Dr.', etc. is in front of a list of clubs, then the first two persons mentioned are members in each of the listed clubs. Otherwise, only the person with the corresponding title is in each of the listed clubs. List each relation as a bullet point. [example text entry and outputted relation triplets]
Rhodesia	[text entry] Extract relations in the format of person-name; member; club from the text, indicating that person-name is a member of club. Consider only the clubs that are listed after "Club(s):". List each relation as a bullet point. An example is: [example text entry and outputted relation triplets]

Table 3: LLM prompts used for relation extraction in each dataset. Prompts are shown exactly as given to the models.

Dataset	OCR	Metric	Gem	4o	ll-70b	ll-8b
Denver	Doc AI	F1 (%)	92.8	90.0	85.5	66.0
		RMAE	0.03	0.07	0.08	0.31
	Gemini	F1 (%)	95.9	94.1	91.5	90.6
		RMAE	0.01	0.03	0.02	0.06
Tenn.	Doc AI	F1 (%)	93.8	91.4	84.2	57.1
		RMAE	0.04	0.13	0.23	0.40
	Gemini	F1 (%)	94.3	91.8	84.2	54.0
		RMAE	0.04	0.12	0.21	0.37
Rhod.	Doc AI	F1 (%)	76.8	79.1	69.9	60.9
		RMAE	0.15	0.21	0.14	0.19
	Gemini	F1 (%)	83.7	86.1	77.5	68.9
		RMAE	0.19	0.19	0.17	0.17

Table 4: Edge-triplet F1 (%) and club-size error (RMAE) for the top 10 largest clubs, across four LLMs on OCR outputs from better (Gemini) and worse (TextRact) OCR models.

lions of historical documents, exists solely in print (Gotscharek et al., 2009; Piotrowski, 2012), making OCR an essential step in extracting KGs from these texts. We compare two OCR models: Google DocumentAI[3], which outperforms prior methods (Hegghammer, 2022), and Gemini-1.5-pro, which we find to also perform well.

Step 2: LLM. *Input: Text; Output: Relation triplets of* ⟨PERSON*;* MEMBER*;* CLUB⟩. To perform RE with LLMs as in existing literature, we use 1-shot in-context learning by providing an LLM with instructions about triplets to extract from an entry of text, plus one illustrative example (see Table 3 for prompts and examples; Fig. 1). We evaluate four LLMs: two proprietary models—Gemini-1.5-pro (Gemini, 2024) and GPT-4o (OpenAI, 2024)—and two open-weight models, Llama 3 (8B and 70B) (AI@Meta, 2024). Each model uses the same prompt.

Step 3: Analysis of the extracted KG. We perform the experiments over the text of *Denver*, *Tennessee*, and *Rhodesia* in AFFILKG and measure each model's output triplets using two metrics:

Triplet precision, recall, and F1 (Edge F1). In line with hundreds of previous papers (Zhao et al., 2024; Detroja et al., 2023), we compute precision, recall, and F1 on relation triplets of ⟨PERSON; MEMBER; CLUB⟩ across 4 LLMs, over texts outputted by the two OCR models. We refer to these as "edge correctness" metrics because each triplet uniquely defines an edge in the affiliation graph, where PERSON and CLUB entities correspond to nodes and MEMBER relations correspond to edges. A true positive requires that all three components of an LLM triplet—PERSON, MEMBER, and CLUB—match that of a ground truth triplet. To account for minor spelling or naming differences (such as omitted middle initials or suffixes in person names, or LLM outputs with abbreviated or parenthetical club names), we allowed a flexible entity match: either an exact match, or both strings exceeding six characters and one is a substring of the other. We also counted matches between common abbreviations and their expanded forms (e.g., "Bulawayo" and "Byo"; "Assn" and "Association"), after normalizing to remove spaces and punctuation. For PERSON entities, titles such as "Mr." or "Mrs." were required to match if present in the text. This approach was validated by manually reviewing 150 entity pair comparisons per dataset. Table 4 shares F1 performance across four LLMs on text produced by two OCR models, revealing substantial variability: differences between LLMs on the same OCR output range from 5 to 40 points, with the highest F1 reaching 95.9.

Relative Mean Absolute Error (RMAE) for number of people in a club. Among many graph analyses that are useful for real applications, we select a simple one that captures higher-level KG structure: number of people in a club. Club membership is deemed to be vital to how elite communities socialize and cohere as a social group, as they convey status, facilitate socialization and the development of norms within elite communities (Accominotti et al., 2018; Cousin and Chauvin, 2014). To compute the number of people in a club c on an affiliation graph, we count the number of neighbors

[3]https://aws.amazon.com/textract/

each club node has, as: $n_c = \sum_{neighbors(c)} 1$. To quantify errors, we use the following error formula:

$$\text{RMAE} = (1/|C|) \sum_{c \in C} |\hat{n}_c - n_c|/n_c$$

where C is the gold-standard set of clubs, and n_c and $\hat{n}_c$ are the true and predicted sizes of club c respectively. Our analysis focuses on the ten largest clubs, as they typically play the most central roles in elite social networks—serving as hubs of influence, status, and social cohesion—and most strongly influence downstream applications. Details about these clubs, including their names, full descriptions, and membership counts in *Denver*, *Tennessee*, and *Rhodesia* appear in Tables 5, 6, and 7. Table 4 shows that club-size RMAE may be as low as 0.01 and as high as 0.37.

Implications for future work. Our preliminary results on three AFFILKG datasets show that OCR and LLM combinations can construct meaningful KGs for real-world use, but performance varies significantly by model and dataset. Higher edge accuracy corresponds to lower club-size errors in *Denver* and *Tennessee*, but not in *Rhodesia*, suggesting other factors at play.

A core motivation for AffilKG is to enable rigorous evaluation not just of extraction performance, but of the impact of extraction errors on downstream graph analyses, including tasks such as community detection, centrality, clustering, and connectivity. For example, Cai and O'Connor (2025) use AffilKG to systematically quantify how real-world extraction errors propagate to these analyses, empirically establish performance thresholds for reliable results, and reveal new patterns of error bias not captured by existing datasets. By providing large-scale, text-derived ground truth networks, AffilKG uniquely enables this kind of end-to-end evaluation, addressing shortcomings of prior resources that were limited to disconnected, small, or low-accuracy graphs.

Future research should continue to test additional OCR models and LLMs, expand coverage to all AffilKG datasets and relation types, and further explore how various extraction errors influence a range of graph-based analyses in real-world contexts.

4. Conclusion

AFFILKG establishes a new benchmark for evaluating knowledge graph extraction, as the first dataset collection that pairs complete, scanned books with large, high quality annotations of affiliation and knowledge graphs. Unlike prior resources, AFFILKG captures coherent, connected networks from real-world historical texts, enabling the first rigorous investigations of how extraction errors affect downstream graph analyses, which is an important concern for practical and scientific applications. By covering diverse geographical regions and time periods, our datasets support robust benchmarking of KG extraction methods and foster research into both methodological advancements in relation extraction and substantive questions about social network analysis. We release AFFILKG to the community with the aim of driving forward comprehensive, application-relevant evaluation of KG extraction systems, and we encourage future work to build upon this resource to advance both NLP and social science research.

5. Limitations

We did not carry out extensive prompt engineering for LLMs because our goal was to explore the impact of LLM errors on higher-level graph analyses, rather than to push performance. Early tests showed that rewording instructions produced similar results. Our prompts included detailed instructions and a carefully chosen example.

While semi-structured text (i.e., text that may have a few patterns, such as names on the first line) may seem easier to extract affiliation graphs from, its prevalence in real applications (O'Brien, 2024; Broad, 2020; Baltzell, 1958) and the numerous errors by all LLMs on all texts in AFFILKG highlight the need for its study. In particular, performance on *Rhodesia* struggles significantly more overall.

6. Ethical considerations

The collected manual annotations (for the Denver book) were conducted by two of the coauthors. No human subject study was performed.

The collection of these social register books is part of a larger project to map out and analyze elite social networks, which may shed light on social questions about the dynamics of power and community, especially among individuals at the top of a social hierarchy. The books are historical, focusing on elite adults from 80 to 110 years ago, and do not uniquely identify living individuals.

While this paper's focus is on historical social networks, accurate inference of contemporary social networks has numerous uses not only in academic social science, but also applications with more complex ethical considerations such as marketing, law enforcement, etc.

Acknowledgments

We thank anonymous reviewers, members of the UMass Natural Language Processing group, and participants in the UMass Computational Social Science Institute seminar, for feedback. Thanks

Club name in text	Full club name	Description	# members
D. C. C.	Denver Country Club	An exclusive organization devoted principally to golf and outdoor recreation.	174
D. C.	Denver Club	One of Denver's most exclusive social organizations. Membership composed of financiers, representative business and professional men.	150
M. C. of C.	Motor Club of Colorado	Devoted to the motoring interest. Membership composed of auto owners.	110
D. A. C.	Denver Athletic Club	Devoted to athletics. Membership composed of prominent business and professional men.	104
C. H. C.	Cherry Hill Club	A country club devoted to outdoor recreation. Membership composed of representative men and women.	98
U. C.	University Club	An exclusive organization composed of representative business and professional men who are college graduates.	64
M. H. C.	Mile High Club	An exclusive organization to entertain national and international notables at dinner.	43
L. C. C.	Lakewood Country Club	A representative organization devoted to golf and outdoor recreation.	42
R. C.	Rotary Club	An organization composed of prominent merchants from every line of trade. Formed for commercial and civic betterment.	36
C. C.	Cactus Club	An organization for literary and dramatic purposes. Largely used as a downtown lunch and dinner club by professional and business men.	35

Table 5: 10 largest clubs in *Denver* (1931), with descriptions and membership counts, where descriptions are copied from the Denver 1931 Social Register 1931.

Club name in text	# members
Junior League	93
Fairyland	73
Belle Mead Country	65
Memphis Country	63
Mountain City	38
Cumberland	26
Centennial	25
Cotillion	24
Daughters of the American Revolution	23
Colonial Dames	22

Table 6: 10 largest clubs in *Tennessee* (1951) with membership counts.

Club name in text	Full club name	# members
Sby.	Salisbury Club	299
Byo.	Bulawayo Club	146
Royals Byo. Golf	Royal Bulawayo Golf Club	88
Byo. Country	Bulawayo Country Club	55
Ruwa Country	Ruwa Country Club	51
Sby. Sports	Salisbury Sports Club	48
Umtali	Umtali Club	44
New	New Club	42
Lon.	London Club	35
Rotary	Rotary Club	30

Table 7: 10 largest clubs in *Rhodesia, Central and East Africa* (1963) with abbreviated/full names and membership counts.

also to Sol Doeleman, Jia Sharma, Izaura Solipa and Shih-Yen Pan for additional comments and assistance. This material is based in part upon work supported by National Science Foundation award 1845576. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

7. Bibliographical References

1892. *Social Register, New York*, volume VI. Social Register Association, 35 LIBERTY STREET NEW YORK CITY.

1903. *Social Register, Boston*, volume XVII. Social Register Association, 261 BROADWAY NEW YORK CITY.

1931. *Social Record and Club Annual Denver Colorado, Volume XXIII 1930-1931*. Geo V. Richards, Denver.

1963. *Who's Who of Southern Africa 1963: Incorporating South African Who's Who and the Who's Who of the Federation of Rhodesia and Nyasaland, Central and East Africa*. Wootton and Gibson, Johannesburg.

Fabien Accominotti, Shamus Khan, and Adam Storer. 2018. How cultural capital emerged in Gilded Age America: Musical purification and cross-class inclusion at the new york philharmonic. *American Journal of Sociology*, 123:1743–1783.

AI@Meta. 2024. The Llama 3 herd of models.

Anthropic. 2024. The Claude 3 model family: Opus, sonnet, haiku.

E. Digby Baltzell. 1958. Philadelphia gentlemen: The making of a national upper class. *(New York: The Free Press)*, 123:1743–1783.

Gisela Bichler, Andrea Schoepfer, and Stacy Bush. 2015. White collars and black ties: Interlocking social circles of elite corporate offenders. *Journal of Contemporary Criminal Justice*, 31(3).

Matthew Bond. 2012. The bases of elite social behaviour: Patterns of club affiliation among members of the house of lords. *Sociology*, 46(4):613–632.

Ronald L. Breiger. 1974. The duality of persons and groups. *Social Forces*, 53(2):181–190.

David B. Broad. 1996. The social register: Directory of America's upper class. *Sociological Spectrum*, 16(2):173–181.

David B. Broad. 2020. The social register: Staying relevant in the post-industrial age. *Journal of Public and Professional Sociology*, 12(1):Article 4.

Erica Cai and Brendan O'Connor. 2025. Understanding the effect of knowledge graph extraction error on downstream graph analyses: A case study on affiliation graphs.

Erica Cai, Olga Simek, Benjamin A. Miller, Danielle Sullivan, Evan Young, and Christopher L. Smith. 2024. From low resource information extraction to identifying influential nodes in knowledge graphs. In *Complex Networks XV*, pages 15–27, Cham. Springer Nature Switzerland.

Jhih-wei Chen, Tsu-Jui Fu, Chen-Kang Lee, and Wei-Yun Ma. 2021. H-FND: Hierarchical false-negative denoising for distant supervision relation extraction. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 2579–2593, Online. Association for Computational Linguistics.

Joseph A Clougherty, Klaus Gugler, Lars Sørgard, and Florian W Szücs. 2014. Cross-border mergers and domestic-firm wages: Integrating "spillover effects" and "bargaining effects". *Journal of International Business Studies*, 45(4):450–470.

Alessandro Corradi, Diane H. Felmlee, and Scott S. Gartner. 2024. A bipartite network analysis of terrorism in Nigeria. *Crime & Delinquency*, 71(3):795–829.

Bruno Cousin and Sébastien Chauvin. 2014. Globalizing forms of elite sociability: Varieties of cosmopolitanism in Paris social clubs. *Ethnic and Racial Studies*, 37(12):2209–25.

Kartik Detroja, C.K. Bhensdadia, and Brijesh S. Bhatt. 2023. A survey on relation extraction. *Intelligent Systems with Applications*, 19:200244.

Ning Ding, Xiaobin Wang, Yao Fu, Guangwei Xu, Rui Wang, Pengjun Xie, Ying Shen, Fei Huang, Haitao Zheng, and Rui Zhang. 2021. Prototypical representation learning for relation extraction. *ICLR*.

George Doddington, Alexis Mitchell, Mark Przybocki, Lance Ramshaw, Stephanie Strassel, and Ralph Weischedel. 2004. The automatic content extraction (ACE) program – tasks, data, and evaluation. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*, Lisbon, Portugal. European Language Resources Association (ELRA).

Brett Drury, Samuel Morais Drury, Md Arafatur Rahman, and Ihsan Ullah. 2022. A social network of crime: A review of the use of social networks for crime and the detection of crime. *Online Social Networks and Media*, 30:100211.

Scott W. Duxbury. 2020. Identifying key players in bipartite networks. *Network Science*, 8(1).

Claire Gardent, Anastasia Shimorina, Shashi Narayan, and Laura Perez-Beltrachini. 2017.

Creating training corpora for NLG micro-planners. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 179–188, Vancouver, Canada. Association for Computational Linguistics.

Gemini. 2024. Gemini: A family of highly capable multimodal models.

Annette Gotscharek, Ulrich Reffle, Christoph Ringlstetter, and Klaus U. Schulz. 2009. On lexical resources for digitization of historical documents. In *Proceedings of the 9th ACM Symposium on Document Engineering*, DocEng '09, page 193–200, New York, NY, USA. Association for Computing Machinery.

Thomas U. Grund and James A. Densley. 2015. Ethnic homophily and triad closure: Mapping internal gang structure using exponential random graph models. *Journal of Contemporary Criminal Justice*, 31:354–370.

Klaus Gugler, Dennis C Mueller, B.Burcin Yurtoglu, and Christine Zulehner. 2003. The effects of mergers: an international comparison. *International Journal of Industrial Organization*, 21(5):625–653.

Jiale Han, Bo Cheng, and Wei Lu. 2021. Exploring task difficulty for few-shot relation extraction. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 2605–2616, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Xu Han, Hao Zhu, Pengfei Yu, Ziyun Wang, Yuan Yao, Zhiyuan Liu, and Maosong Sun. 2018. FewRel: A large-scale supervised few-shot relation classification dataset with state-of-the-art evaluation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 4803–4809, Brussels, Belgium. Association for Computational Linguistics.

Thomas Hegghammer. 2022. OCR with Tesseract, Amazon Textract, and Google Document AI: a benchmarking experiment. *Journal of Computational Social Science*, 5(1):861–882.

Petter Holme, Sung Min Park, Beom Jun Kim, and Christofer R. Edling. 2007. Korean university life in a network perspective: Dynamics of a large affiliation network. *Physica A: Statistical Mechanics and its Applications*, 373(C):821–830.

Pere-Lluis Huguet Cabot, Simone Tedeschi, Axel-Cyrille Ngonga Ngomo, and Roberto Navigli. 2023. RED$^{\text{fm}}$: a filtered and multilingual relation extraction dataset. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4326–4343, Toronto, Canada. Association for Computational Linguistics.

Haruna Isah, Daniel Neagu, and Paul Trundle. 2015. Bipartite network model for inferring hidden ties in crime data. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*, ASONAM '15, page 994–1001, New York, NY, USA. Association for Computing Machinery.

Mrs. John Northen Jackson, editor. 1965. *Social Register of Greater Miami, 1965*. Blue Book Publishing Company, 1830 West 24th Street, Sunset Island #3, Miami Beach, Florida. Publisher–Editor; Mrs. Lino Sertel, Founder–Advisor.

Sarthak Jain, Madeleine van Zuylen, Hannaneh Hajishirzi, and Iz Beltagy. 2020. SciREX: A challenge dataset for document-level information extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7506–7516, Online. Association for Computational Linguistics.

Bernal Jimenez Gutierrez, Nikolas McNeal, Clayton Washington, You Chen, Lang Li, Huan Sun, and Yu Su. 2022. Thinking about GPT-3 in-context learning for biomedical IE? think again. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pages 4497–4512, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Suzanne Keller. 1963. *Beyond the Ruling Class*. Random House, New York.

Gueorgi Kossinets and Duncan J. Watts. 2006. Empirical analysis of an evolving social network. *Science*, 311(5757):88–90.

Silvio Lattanzi and D. Sivakumar. 2009. Affiliation networks. In *Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing*, STOC '09, page 427–434, New York, NY, USA. Association for Computing Machinery.

Yi Luan, Luheng He, Mari Ostendorf, and Hannaneh Hajishirzi. 2018. Multi-task identification of entities, relations, and coreference for scientific knowledge graph construction. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3219–3232, Brussels, Belgium. Association for Computational Linguistics.

P. Mac Carron and Ralph Kenna. 2013. Viking sagas: Six degrees of icelandic separation social networks from the viking era. *Significance*, 10(6):12–17.

Luca Marotta, Salvatore Miccichè, Yoshi Fujiwara, Hiroshi Iyetomi, Hideaki Aoyama, Mauro Gallegati, and Rosario N. Mantegna. 2014. Bank-firm credit network in japan: An analysis of a bipartite network. *PLoS ONE*, 10.

C. Wright Mills. 1956. *The Power Elite*. Oxford University Press, Oxford.

Abhishek Nadgeri, Anson Bastos, Kuldeep Singh, Isaiah Onando Mulang', Johannes Hoffart, Saeedeh Shekarpour, and Vijay Saraswat. 2021. KGPool: Dynamic knowledge graph context selection for relation extraction. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 535–548, Online. Association for Computational Linguistics.

OpenAI. 2024. Gpt-4o system card.

Shay O'Brien. 2024. The family web: Multigenerational class persistence in elite populations. *Socio-Economic Review*, 22(1):1–27.

Shay O'Brien. 2025. The gender of inheritance in an upper-class family network: Dallas, 1895–1945. *Socio-Economic Review*, page Forthcoming.

Michael Piotrowski. 2012. *Natural Language Processing for Historical Texts*. Morgan & Claypool Publishers.

C. Prem Sankar, K. Asokan, and K. Satheesh Kumar. 2015. Exploratory social network analysis of affiliation networks of indian listed companies. *Social Networks*, 43:113–120.

Pawan Rajpoot and Ankur Parikh. 2023. GPT-FinRE: In-context learning for financial relation extraction using large language models. In *Proceedings of the Sixth Workshop on Financial Technology and Natural Language Processing*, pages 42–45, Bali, Indonesia. Association for Computational Linguistics.

Sebastian Riedel, Limin Yao, and Andrew McCallum. 2010. Modeling relations and their mentions without labeled text. In *ECML/PKDD*.

Yicheng Tao, Yiqun Wang, and Longju Bai. 2024. Graphical reasoning: Llm-based semi-open relation extraction.

Daria Tisch and Emma Ischinsky. 2023. Top wealth and its historical origins: Identifying entrenched fortunes by linking rich lists over 100 years. *Socius*, 9:23780231231192774.

Bayu Distiawan Trisedya, Gerhard Weikum, Jianzhong Qi, and Rui Zhang. 2019. Neural relation extraction for knowledge base enrichment. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 229–240, Florence, Italy. Association for Computational Linguistics.

Somin Wadhwa, Silvio Amir, and Byron Wallace. 2023. Revisiting relation extraction in the era of large language models. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15566–15589, Toronto, Canada. Association for Computational Linguistics.

Zhen Wan, Fei Cheng, Zhuoyuan Mao, Qianying Liu, Haiyue Song, Jiwei Li, and Sadao Kurohashi. 2023. GPT-RE: In-context learning for relation extraction using large language models. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 3534–3547, Singapore. Association for Computational Linguistics.

Max Weber. 1947. *The Theory of Social and Political Organization*. William Hodge, London.

Barry Wellman and Charles Wetherell. 1996. Social network analysis of historical communities: Some questions from the present for the past. *The History of the Family*, 1(1):97–121.

Stuart Whitmarsh. 1951. *The Southern Social Register*. The Southern Social Register Foundation, New York.

Peng Xu and Denilson Barbosa. 2019. Connecting language and knowledge with heterogeneous representations for neural relation extraction. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 3201–3206, Minneapolis, Minnesota. Association for Computational Linguistics.

Lilong Xue, Dan Zhang, Yuxiao Dong, and Jie Tang. 2024. AutoRE: Document-level relation extraction with large language models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)*, pages 211–220, Bangkok, Thailand. Association for Computational Linguistics.

Yuan Yao, Deming Ye, Peng Li, Xu Han, Yankai Lin, Zhenghao Liu, Zhiyuan Liu, Lixin Huang, Jie Zhou, and Maosong Sun. 2019. DocRED: A large-scale document-level relation extraction dataset. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 764–777, Florence, Italy. Association for Computational Linguistics.

Jacob T. N. Young and Justin T. Ready. 2015. Diffusion of ideas and technology: The role of networks in influencing the endorsement and use of on-officer video cameras. *Journal of Contemporary Criminal Justice*, 31:243–261.

Bowen Zhang and Harold Soh. 2024. Extract, define, canonicalize: An LLM-based framework for knowledge graph construction. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 9820–9836, Miami, Florida, USA. Association for Computational Linguistics.

Jeffrey Zhang, Maxwell Wibert, Huixue Zhou, Xueqing Peng, Qingyu Chen, Vipina K Keloth, Yan Hu, Rui Zhang, Hua Xu, and Kalpana Raja. 2024. A study of biomedical relation extraction using GPT models. *AMIA Jt Summits Transl Sci Proc*, 2024:391–400.

Yuhao Zhang, Victor Zhong, Danqi Chen, Gabor Angeli, and Christopher D. Manning. 2017. Position-aware attention and supervised data improve slot filling. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 35–45, Copenhagen, Denmark. Association for Computational Linguistics.

Xiaoyan Zhao, Yang Deng, Min Yang, Lingzhi Wang, Rui Zhang, Hong Cheng, Wai Lam, Ying Shen, and Ruifeng Xu. 2024. A comprehensive survey on relation extraction: Recent advances and new frontiers. *ACM Comput. Surv.*, 56(11).

Huixue Zhou, Mingchen Li, Yongkang Xiao, Han Yang, and Rui Zhang. 2024. LEAP: LLM instruction-example adaptive prompting framework for biomedical relation extraction. *Journal of the American Medical Informatics Association*, 31(9):2010–2018.