Which entities are relevant for the story?

Pooja H. Oza
University of New Hampshire
Durham, NH 03824
pho1003@wildcats.unh.edu

Laura Dietz
University of New Hampshire
Durham, NH 03824
dietz@cs.unh.edu

Abstract

A crucial step in the construction of any event story or news report is to identify entities involved in the story, such entities can come from a larger background knowledge graph or from a text corpus with entity links. Along with recognizing which entities are relevant to the story, it is also important to select entities that are relevant to all aspects of the story. In this work, we model and study different types of links between the entities with the goal of identifying which link type is most useful for the entity retrieval task. Our approach demonstrates the efficacy of using entity links in relevant passages to identify relevant entities and relations. This approach outperforms relational information from knowledge graph links and bibliographic coupling. We evaluate different approaches on a benchmark of open-domain queries. We find that the co-occurrence of entities in the relevant text is one of the strongest indicators (0.14 MAP), and hence should always be considered when constructing a story.

1 Introduction

Entities are ubiquitous and central units to describe stories. Identifying which entities are relevant for any story is at the core of the tasks, such as event identification or news reporting. The goal of entity retrieval is to identify and retrieve entities that are relevant for an information need. In this work, we study the first crucial step of a story construction i.e., identification and retrieval of a list of entities relevant to a story. We demonstrate how a text corpus is central for obtaining good retrieval performance. As a byproduct of our approach, we also provide a set of relevant text passages.

Previous work of entity retrieval focuses on the link-structure of knowledge graphs, text surrounding the entities, entity attributes such as type, category, name by matching the entity target type with the query target type [GB17] or by applying language models to evaluate the category matching between queries and entities [BBDR10]. To identify and retrieve entities relevant to a story, we explore three different types of links between entities; i) relevant co-occurrence graph ii) unweighted link patterns iii) relevance-weighted bibliographic coupling. In the relevant co-occurrence graph, we study the effectiveness of the co-occurrence of entities, whereas in unweighted link patterns we examine the impact of knowledge graph link structure in entity retrieval performance. In the third link type, relevance-weighted bibliographic coupling, we investigate the effect of the combination of knowledge graph link structure with the relevance of the text documents. We provide further explanation of these three link types in Section 3.

In: R. Campos, A. Jorge, A. Jatowt, S. Bhatia, M. Finlayson (eds.): Proceedings of the Text2Story'21 Workshop, Online, 1-April-2021, published at http://ceur-ws.org/

We envision that a story request is specified through texts such as 'Radiocarbon dating background history', 'Antibiotics side-effects', 'Antibiotics resistance', 'Deforestation causes'. A user may also be interested in retrieving entities relevant to a larger theme that diversifies over multiple stories, such as 'Antibiotics'. For such an information need, we envision user to specify a multi-story request. We define multi-story request as a list of multiple story requests connected through a common theme. For example, a multi-story query such as ['Antibiotics side-effects', 'Antibiotics resistance'] gives us entities relevant to the theme common among these stories such as 'Antibiotics'.

Task (Story-specific Entity Retrieval):

Given a story request as a text query, (1) identify which entities in the knowledge base are relevant, and (2) associate each entity with a score of relative relevance. We assume access to a general-purpose knowledge base and large text corpus.

The formulation of our task is related to the task of entity retrieval in information retrieval [ZKN15], semantic search [BFC$^+$14], and question answering [BCFL13]. To construct a story, we need to retrieve a diverse set of relevant entities that cover all the facets of a story. We interpret a multi-story request as an overarching set of several stories sharing a common theme.

The main challenge in entity retrieval is that it is not sufficient to rely on exact keyword matches as in TF-IDF. We show that the keyword matching for entity retrieval is not an efficient approach in Section 4. Instead, we find it beneficial to combine several weak indicators of relevance to identify a relevant set of entities.

Contributions:

To find a set of relevant entities, most related work either retrieves entities through matches in the knowledge base entry and/or by examining link patterns in the knowledge base. We demonstrate, that it is much more effective to retrieve relevant text passages and use the occurrence of entities as well as the co-occurrence of two entities as an indicator for the relevance of entities.

While a full exploration of all prior approaches is out of scope for this work, our study suggests that more attention to the textual context in which entities occur is a worthwhile endeavor.

2 Related Work

We review below related work in question answering, information retrieval, and semantic search.

2.1 Question Answering and Semantic Search

Many question answering tasks ask to retrieve one (of few) entities in response to a concise description. The Dr. QA system [CFWB17] demonstrates that using a combination of information retrieval and neural network, many open-domain entity-centric questions get answers from free text. An alternative is semantic parsing, in which the question translates into a SPARQL query to select the relevant entities [BCFL13]. Many issues revolve around matching terms in the question to properties, types, and how to use the link structure most effectively [ZHW$^+$14]. Combined knowledge graph and corpus approaches have demonstrated potential [SA16].

2.2 Entity Retrieval

Several information retrieval models are modified to search in structured entries of knowledge graphs, using the information of entity such as name, type, category, outlinks/inlinks [PMZ10, RCK12]. For example, the Fielded Sequential Dependence Model model [ZKN15] uses unigram and sparse bigram matches of the query in the knowledge graph fields name, category, and neighbors. Garigliotti and Balog [GB17] explore the use of entity type (i.e. taxonomic) information to exploit the type-based similarity between queries and entities. While only rarely used, text-based indicators from the context of an entity were found to be a strong feature for entity ranking tasks [SDPP15, Die19].

2.3 Entity Set Expansion and Graph Walks

When applied to topically focused knowledge graphs, such as ConceptNet, random walks and other graph-based methods are effective in finding relevant entities [KZ12]. However, in application to large general-purpose

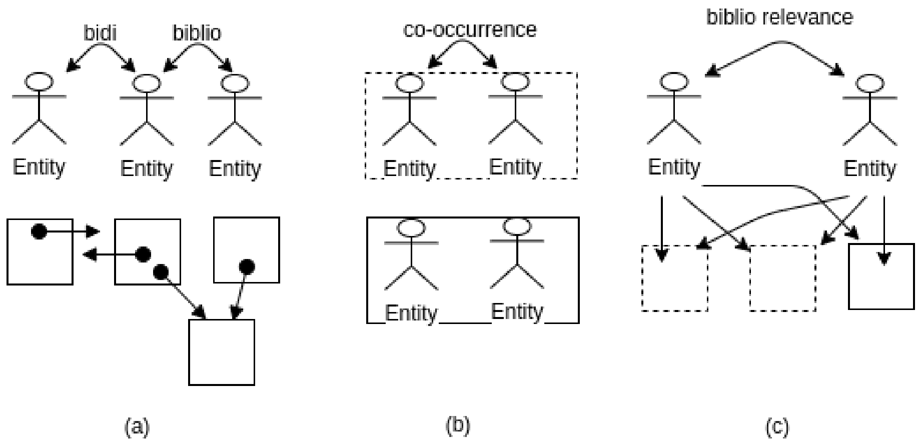

Figure 1: Depiction of three link types that are based on (a) link-patterns in the knowledge base such as bidirectional links or unweighted bibliographic coupling (b) co-occurrence of entities in relevant passages (c) bibliographic / co-coupling measures that incorporate the relevance of shared entities. Logical entities and entity links are represented as stick figures, which are associated with the knowledge base metadata (depicted as rectangles with links) or can be contained in passages. Documents with dash lines are of higher relevance.

knowledge graphs derived from Wikipedia, indicators from the link structure offer only disappointing performance [BFC+14].

A common approach is to start with a set of seed entities, then try to expand the set to additional entities that are similar to the seed set. For example, Pantel et al. [PCB+09] represent entities through a vector representation, then include the most similar entities. This method can be refined with neural networks [RPLVD19].

Most of the related work focuses on identifying a small set of correct entities, which explains the effectiveness of exact keyword matching and semantic parsing. In contrast, we focus on obtaining a relevant set of entities that are relevant to different degrees, for which we show that a keyword matching approach is not sufficient. In contrast, we find that one of the most useful indicators is to use the context of the entity mentions in relevant text passages.

3 Approach: Identifying Entities Relevant for a Story

Previous work relies on keyword matches in entity names and categories, or graph walks on related entities. We hypothesize that one of the most effective indicators for the relevance of entities lies in the text surrounding the entity mentions. We argue that entities which are co-mentioned in relevant text passages are likely to have a relevant connection. We demonstrate that co-coupling patterns defined on the co-occurrence graph are more effective than co-coupling patterns on the general-purpose knowledge base. We also explore a hybrid approach, where links for co coupling and bibliographic coupling are weighted by relevance using text retrieval indicators.

We are making the following claims:

- Claim 1: Finding entities through relevant passages is more effective than searching in the knowledge base for entity entries (using short descriptions).
- Claim 2: To predict the relevance of an entity link, entity co-occurrences in relevant passages are more effective than link patterns on edges of a general-purpose knowledge base.
- Claim 3: Count-based bibliographic and co-coupling similarities can be improved by weighting edges in the knowledge base by the relevance of the shared entities.

We study these claims through the following approaches, both in isolation and in combination. We define three link types between entities, which are based on knowledge base links, co-occurrences, and relevance-based coupling as depicted in Figure 1 and defined in Sections 3.1–3.3.

Our approach takes as an input a ranking of passages P for the relevant co-occurrence graph and for the relevance weighted coupling measures we use knowledge base entries E.

KG-Entity:

We use BM25 retrieval model to generate rankings of the top 1000 entities using the description field of knowledge base entries.

3.1 Relevant Co-occurrence Graph

We define the *graph of relevant co-occurrences* as follows: Using a large corpus of text passages, we retrieve a ranking of passages for the story. Here we use the top 100 passages using BM25 model. After entity linking each passages, we take note of pairs of entities e_i, e_j which co-occur in the passage P.

Relevance-based co-occurrences (Co-occ Relevance):

From the passage ranking we derive a relevance weight of the edge by accumulating reciprocal ranks of passages in which the entities co-occur as in Equation 1.

$$f_{ecr}(e_i, e_j) = \sum_{\forall P: e_i, e_j \text{ mentioned in } P} \frac{1}{rank(P)}, i \neq j \tag{1}$$

We calculate the entity ranking by aggregating the edge weights of the incident links akin to the degree centrality.

Co-occurrence count (Co-occ Count):

We also explore a simpler alternative by simply counting the number of occurrences. We only count occurrences in passages that were retrieved. Similar to Co-occ Relevance, we calculate the entity ranking by accumulating the edge weights of the incident links.

Mention Freq:

We count the number of time an entity e_i is mentioned in retrieved passages. We rank the entities based on the frequency.

Unlike link styles described below, the graph of relevant co-occurrences only uses the knowledge base through the entity linker—it does not incorporate any relations provided in the given general-purpose knowledge base.

3.2 Unweighted Link Patterns

As baselines to support Claim 3, we compare to several widely-used link patterns that can be derived from any general-purpose knowledge base. In this line of work, we are ignoring the semantics of different relationship predicates. Concretely, we include the following approaches:

Direct links:

For every pair of entities (e_i, e_j), we determine if a direct link exists between the entities e_i and e_j. We count the number of different link occurrences, which are used as a relevance score. We consider four types of direct links as follows:

1. Outlinks: Number of outlinks from the entry of entity e_i to entity e_j.
2. Inlinks: Number of the incoming link to the entry of e_i from e_j.
3. Bidi: 1 if a bidirectional link exists and 0 otherwise. For a bidirectional link, both entities e_i and e_j link to each other.
4. Undirected: Sum of outlinks and inlinks between e_i and e_j.

For each entity e_i in Direct links, we derive entity ranking by accumulating the score of the incident links.

Coupling (baseline):

Bibliographic coupling and co-coupling are similarity measures that score a semantic relationship between two entities based on the number of shared in or outlinks.

1. Count-based Bibliographic Coupling (Biblio Count): Number of outlinks (i.e., citations) shared by entities e_i and e_j.
2. Count-based Co-coupling (Co Count): Number of shared inlinks between the entities e_i and e_j.

Similar to other link types, we rank the entities by aggregating the edge weights for each entity e_i.

3.3 Relevance-weighted Bibliographic Coupling

We further study a hybrid between link-based and content-based approaches. Traditionally, coupling measures focus on counting the number of shared pages that link to or from both entities e_i and e_j. We suggest relevance-weighted coupling measures by first retrieving a ranking of knowledge base entries E, then accumulating the relevance scores of shared pages. We retrieve ranking of entries E using BM25 retrieval model.

Relevance-weighted Bibliographic Coupling (Biblio Relevance):

Accumulated rank scores for entries E that link from both entities e_i and e_j to E, as in Equation 2.

$$f_{ebr}(e_i, e_j) = \sum_{\forall E: e_i, e_j \text{ linked to } E} score(E), i \neq j \tag{2}$$

We calculate the entity rankings by aggregating the relevance scores of all the entities associated to the entity e_i.

Relevance-weighted Co-Coupling (Co Relevance):

Accumulating rank scores for entries E that link to both e_i and e_j (analogously to Equation 2). For entity ranking of each entity e_i, we calculate the edge weights of all the incident links of e_i.

3.4 Combining Weighted Link-graphs

As some approaches might demonstrate their strengths only when combined, we study combinations with a learning-to-rank approach (L2R). Each of our approaches produces a separate feature for each entity. While many different methods are available, we opt for a list-wise learning-to-rank approach implemented in RankLib [DBC13], optimizing with coordinate ascent to optimize mean-average precision (MAP), which has demonstrated robust state-of-the-art performance on small feature spaces in the past.

3.5 Multi-Story Request

We identify the entities relevant to a multi-story request from the results of multiple stories connected to the request. We generate the result of a multi-story request by combining the top 100 ranking entities for each story connected and accumulating the relevance scores of the entities. We keep the top 100 entities from the merged entities.

4 Evaluation

We are evaluating the three claims in Section 3 on a large-scale benchmark from TREC Complex Answer Retrieval (CAR) [DG18].

4.1 Benchmark and Evaluation Paradigm

The evaluation is based on the TREC Complex Answer Retrieval "Y1 test" benchmark for evaluation and "Y1 train" for training the combination with learning-to-rank. The benchmark includes a knowledge base that is derived from a Wikipedia dump that preserves meta-information about names (title, disambiguation, and anchor text) and the first paragraph as well as the entity ids of inlinks and outlinks. The benchmark also includes a

Table 1: Results on selecting entities for 1952 stories.

Features	MAP	Rprecision	F1	In L2R
Co-occ Relevance	**0.1485** ± 0.0052	0.1427 ± 0.0053	0.0689 ± 0.0017	X
Co-occ Count	0.0955 ± 0.0036	0.1033 ± 0.0039	0.0623 ± 0.0016	X
Mention-Freq	0.1111 ± 0.0035	0.1120 ± 0.0038	0.0723 ± 0.0016	X
Outlinks	0.0740 ± 0.0027	0.0731 ± 0.0029	0.0621 ± 0.0014	
Inlinks	0.0724 ± 0.0025	0.0733 ± 0.0029	0.0652 ± 0.0014	
Bidi	0.0407 ± 0.0015	0.0341 ± 0.0018	0.0608 ± 0.0014	
Undirected	0.0788 ± 0.0027	0.0811 ± 0.0030	0.0640 ± 0.0014	X
Biblio Count	0.0712 ± 0.0023	0.0711 ± 0.0028	0.0655 ± 0.0015	
Co-coupling Count	0.0389 ± 0.0015	0.0365 ± 0.0019	0.0587 ± 0.0013	
Biblio Relevance	0.0501 ± 0.0020	0.0497 ± 0.0024	0.0526 ± 0.0012	
Co-coupling Relevance	0.0766 ± 0.0028	0.0777 ± 0.0031	0.0624 ± 0.0014	
KG-Entity	0.0344 ± 0.0028	0.0355 ± 0.0028	0.0112 ± 0.0004	
L2R	**0.1558** ± 0.0053	0.1461 ± 0.0054	0.0707 ± 0.0015	

background corpus, which is derived from passages on Wikipedia. To ensure reproducibility, we use hyperlinks provided with the text instead of an entity linker.

The TREC CAR benchmark includes queries that are derived from the Wikipedia article outlines. Articles for queries are removed from the knowledge base. The benchmark includes section-level queries formed by concatenating the article title and heading text and article-level queries that consist of only the article title. In line with the task statement in Section 1, a story request is a section-level query and the article-level query forms the common theme of the multiple stories. For example, the query "Radiocarbon dating/Use in archaeology/Notable applications/Pleistocene/Holocene boundary in Two Creeks Fossil Forest" forms one of the stories that is associated to a multi-story request "Radiocarbon dating". The benchmark includes automatic assessments for 1952 stories that are associated to 132 multi-story requests. The validity of the automatic assessments was confirmed by manual assessors [DD20].

We evaluate the quality of the resulting entity set using the F1 measure. The quality of the relative relevance scores is evaluated as a ranking via the measures mean-average Precision (MAP) and Rprecision. With R being the number of true entities, Rprecision measures the set precision among the first R entities of a predicted ranking. In contrast, MAP averages the precision of ranks at which relevant entities are found, awarding partial credit for low-ranking entities.

4.2 Story Results

For each story, we evaluate the quality of the retrieved entities through results presented in Table 1. We find that the relevance-based co-occurrence approach is the most effective approach for entity ranking. In fact, the best three approaches are all based on a passage ranking (supporting Claim 2). This improvement can be noticed in the quality of the selected set, where relevance-weighted co-occurrence obtains a significantly higher F1. Overall the F1 scores are low because we choose to select the top 100 entities, but for most queries, fewer than 100 entities are relevant. Rprecision reflects the quality when selecting the right number of entities.

In Section 1 we claim that keyword matching is not sufficient for identifying a set of relevant entities. To support our claim, we evaluate the ranking of entities via their knowledge base entries as described in Section 3 (KG-Entity). However, this is one of the worst-performing approaches. This also supports our Claim 1 that the relevance-based entity mentions count approach is more effective in giving relevant entities than searching for entity entries in the knowledge base.

For Claim 3, we study whether using a retrieval over knowledge base entries as a relevance-weight in bibliographic and co-coupling patterns, but obtain mixed results: While weighting shared inlinks by relevance (co-coupling relevance), it decreases the performance when weighting shared outlinks (biblio relevance).

To analyze how multiple approaches work in combination, we use learning-to-rank using the four best approaches: three context-based approaches and undirected links in the knowledge base. In this combination, we obtain a small, but not a significant improvement. The combination of all features performs worse, as some

Table 2: Results on selecting entities for 132 multi-story requests.

Features	MAP	Rprecision	F1	In L2R
Co-occ Relevance	**0.2140** ± 0.0111	0.3092 ± 0.0106	0.3106 ± 0.0096	X
Co-occ Count	0.1669 ± 0.0085	0.2636 ± 0.0095	0.2714 ± 0.0087	X
Mention-Freq	0.1971 ± 0.0085	0.3032 ± 0.0087	0.3189 ± 0.0086	X
Outlinks	0.1509 ± 0.0071	0.2568 ± 0.0078	0.2722 ± 0.0078	
Inlinks	0.1566 ± 0.0065	0.2630 ± 0.0078	0.2839 ± 0.0079	
Bidi	0.0538 ± 0.0043	0.1357 ± 0.0072	0.1665 ± 0.0077	
Undirected	0.1640 ± 0.0071	0.2707 ± 0.0078	0.2856 ± 0.0081	X
Biblio Count	0.1469 ± 0.0068	0.2527 ± 0.0081	0.2700 ± 0.0079	
Co-coupling Count	0.0812 ± 0.0052	0.1844 ± 0.0075	0.2129 ± 0.0077	
Biblio Relevance	0.1099 ± 0.0064	0.2135 ± 0.0075	0.2320 ± 0.0080	
Co-coupling Relevance	0.1339 ± 0.0068	0.2320 ± 0.0083	0.2448 ± 0.0086	
KG-Entity	0.0172 ± 0.0020	0.0519 ± 0.0037	0.0499 ± 0.0031	
L2R	**0.2534** ± 0.0116	0.3552 ± 0.0104	0.3175 ± 0.0092	

features work extremely well for some cases, but not consistently.

The relevance-based co-occurrence approach also performs better in terms of MAP and Rprecision (with an exception of co-coupling relevance).

4.3 Multi-story Request Results

We interpret a multi-story request as a list of stories connected to it. In the context of the TREC CAR benchmark, all queries that are derived from the same article form one multi-story request. We evaluate the performance obtained by each approach when using multiple stories, as described in Section 3. The results are given in Table 2.

To demonstrate our Claim 1, we include one baseline: Using each story to retrieve from knowledge base entries (KG-Entity), then combine the results as in Section 3.5. With a MAP of 0.0172 and F1 of 0.05, the performance of the baseline is dramatically lower than any of the alternative approaches in our study—supporting Claim 1.

We see that approaches that consider the relevance of passage contexts are still superior (supporting Claim 2). However, while in the story results, co-occurrence relevance improved over undirected links by 80% in terms of MAP, the advantage is only 30% for multi-story requests.

In general, we see the same pattern for multi-story requests as in the story analysis: relevance-based weighting of shared nodes improves the performance of co-coupling but not bibliographic coupling. We suspect that one of the reasons why no convincing improvement is seen with relevance-weighted bibliographic coupling is that the quality of the underlying knowledge base ranking (KG-Entity) is inferior.

5 Conclusion

We study the problem of selecting a set of relevant entities for a story construction. We find that links in the knowledge base are not effective in selecting relevant entities, possibly because the majority of links are not relevant for the topic at hand. In contrast, using entity co-occurrences in retrieved text passages—especially when the relative relevance of passages is incorporated as link strength—is between 80% and 30% more effective than the best link-based approach. Even combining link-based approach with relevance co-occurrence gives no significant improvement in the result.

While incorporating the relevance of shared entities can improve coupling approaches, this requires a high-quality ranking of entities. We found that one of the most commonly used approaches, namely retrieving entities through names and abstract of their knowledge base entry, is an ineffective approach for a story construction.

6 Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1846017. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and

do not necessarily reflect the views of the National Science Foundation.

References

[BBDR10] Krisztian Balog, Marc Bron, and Maarten De Rijke. Category-based query modeling for entity search. In *European Conference on Information Retrieval*, pages 319–331. Springer, 2010.

[BCFL13] Jonathan Berant, Andrew Chou, Roy Frostig, and Percy Liang. Semantic parsing on freebase from question-answer pairs. In *In Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2013.

[BFC+14] Christopher Boston, Hui Fang, Sandra Carberry, Hao Wu, and Xitong Liu. Wikimantic: Toward effective disambiguation and expansion of queries. *Data & Knowledge Engineering*, 90:22–37, 2014.

[CFWB17] Danqi Chen, Adam Fisch, Jason Weston, and Antoine Bordes. Reading wikipedia to answer open-domain questions. In *Proceedings of the 55rd Annual Meeting of the Association for Computational Linguistics (ACL)*, 2017.

[DBC13] Van Dang, Michael Bendersky, and W Bruce Croft. Two-stage learning to rank for information retrieval. In *ECIR*, pages 423–434. Springer, 2013.

[DD20] Laura Dietz and Jeff Dalton. Humans optional? automatic large-scale test collections for entity, passage, and entity-passage retrieval. *Datenbank-Spektrum*, 2020.

[DG18] Laura Dietz and Ben Gamari. Trec car 2.0: A data set for complex answer retrieval, 2018.

[Die19] Laura Dietz. Ent rank: Retrieving entities for topical information needs through entity-neighbor-text relations. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 215–224, 2019.

[GB17] Darío Garigliotti and Krisztian Balog. On type-aware entity retrieval. In *Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval*, pages 27–34. ACM, 2017.

[KZ12] Alexander Kotov and ChengXiang Zhai. Tapping into knowledge base for concept feedback: leveraging conceptnet to improve search results for difficult queries. In *Proceedings of the fifth ACM international conference on Web search and data mining*, pages 403–412. ACM, 2012.

[PCB+09] Patrick Pantel, Eric Crestan, Arkady Borkovsky, Ana-Maria Popescu, and Vishnu Vyas. Web-scale distributional similarity and entity set expansion. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 2-Volume 2*, pages 938–947. Association for Computational Linguistics, 2009.

[PMZ10] Jeffrey Pound, Peter Mika, and Hugo Zaragoza. Ad-hoc object retrieval in the web of data. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, page 771–780, New York, NY, USA, 2010. Association for Computing Machinery.

[RCK12] Hadas Raviv, David Carmel, and Oren Kurland. A ranking framework for entity oriented search using markov random fields. In *Proceedings of the 1st Joint International Workshop on Entity-Oriented and Semantic Search*, page 1. ACM, 2012.

[RPLVD19] Pushpendre Rastogi, Adam Poliak, Vince Lyzinski, and Benjamin Van Durme. Neural variational entity set expansion for automatically populated knowledge graphs. *Information Retrieval Journal*, 22(3-4):232–255, 2019.

[SA16] Denis Savenkov and Eugene Agichtein. When a knowledge base is not enough: Question answering over knowledge bases with external text data. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 235–244, 2016.

[SDPP15] Michael Schuhmacher, Laura Dietz, and Simone Paolo Ponzetto. Ranking entities for web queries through text and knowledge. In *Proceedings of the 24th ACM international on conference on information and knowledge management*, pages 1461–1470. ACM, 2015.

[ZHW+14] Lei Zou, Ruizhe Huang, Haixun Wang, Jeffrey Xu Yu, Wenqiang He, and Dongyan Zhao. Natural language question answering over rdf: a graph data driven approach. In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, pages 313–324. ACM, 2014.

[ZKN15] Nikita Zhiltsov, Alexander Kotov, and Fedor Nikolaev. Fielded sequential dependence model for ad-hoc entity retrieval in the web of data. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 253–262. ACM, 2015.