

# SERAG: Semantic Entity Retrieval from Arabic knowledge Graphs

Saher Esmeir

Bloomberg L.P.

London, United Kingdom

sesmeir2@bloomberg.net

## Abstract

Knowledge graphs (KGs) are widely used to store and access information about entities and their relationships. Given a query, the task of entity retrieval from a KG aims at presenting a ranked list of entities relevant to the query. Lately, an increasing number of models for entity retrieval have shown a significant improvement over traditional methods. These models, however, were developed for English KGs. In this work, we build on one such system, named KEWER, to propose SERAG (Semantic Entity Retrieval from Arabic knowledge Graphs). Like KEWER, SERAG uses random walks to generate entity embeddings. DBpedia-Entity v2 is considered the standard test collection for entity retrieval. We discuss the challenges of using it for non-English languages in general and Arabic in particular. We provide an Arabic version of this standard collection, and use it to evaluate SERAG. SERAG is shown to significantly outperform the popular BM25 model thanks to its multi-hop reasoning.

## 1 Introduction

Recent years have witnessed a rapid growth in the use of KGs. Both large-scale graphs that connect millions of entities through billions of relationships and smaller domain-specific KGs have become widely available. One way to access the information stored in KGs is through a dedicated query language, such as SPARQL. This, however, requires proficiency with the syntax of the language and knowledge of the exact schema of the database. Given a query in natural language, the task of entity retrieval aims at presenting a ranked list of entities relevant to the query. Consider, for example, the query “museums in Arab capitals”. Figure 1 shows how the subgraph around the entity  $E3$ , labelled “Egyptian Museum”, supports its relevance to the query. This complex, indirect relationship is not represented textually in any single document.

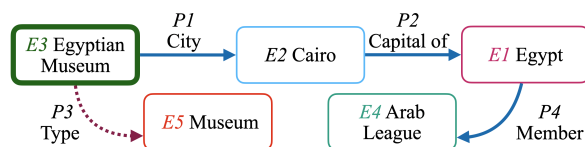


Figure 1: Given the query “museums in Arab capitals”, the relevance of the entity  $E3$  (Egyptian Museum) is supported by the information along two paths.

The current standard test collection for entity search is DBpedia-Entity v2 (DEv2). It is provided with the performance results of a dozen baseline methods (Hasibi et al., 2017). Lately, Nikolaev and Kotov (2020) presented KEWER, a system that significantly outperformed previous methods on DEv2. KEWER, nevertheless, was developed for and tested on English KGs. In this work, we build on it and propose SERAG,<sup>1</sup> a method for Semantic Entity Retrieval from Arabic knowledge Graphs. To evaluate SERAG, we create and share a Modern Standard Arabic (MSA) version of DEv2. To our knowledge, this is the first attempt to offer DEv2 in a non-English language. SERAG is shown to perform well on Arabic DEv2 and to significantly improve on the classic BM25 retriever.

**Contributions** The notable contributions of this work are: (1) introducing a new Arabic test collection for entity retrieval by translating queries from DEv2 and leveraging DBpedia’s inter-language links,<sup>2</sup> and (2) proposing a method for entity retrieval from Arabic KGs and evaluating it on the new test collection.

## 2 Background and Related Work

There are several attempts in literature to define KGs (Ehrlinger and Wöb, 2016; Fensel et al., 2020).

<sup>1</sup>سراج (sirāj) is Arabic for a lantern.

<sup>2</sup>The translated queries and usage directions can be found at <https://doi.org/10.5281/zenodo.4560653>.

In this work, we follow Färber et al. (2018) and view a KG as a finite set of RDF triples in the form of subject-predicate-object ( $s, p, o$ ).

**Graph Embedding** Graph embeddings transform nodes and edges into a low dimension vector space while preserving structural information, an idea that was adopted from language modeling. Cai et al. (2018) survey various graph embedding methods. These methods, however, focus on the relational structure of the graph and do not leverage the information present in the surface forms of entities and predicates. Consider again the graph in Figure 1.  $P1$  links  $E3$  and  $E2$ , but there is additional information in the labels attached to those.

**Entity Retrieval** Since the introduction of DEv2, several works have reported significant progress on entity retrieval. Naseri et al. (2018) explored enriching entity representations using information from related entities. Kadilierakis et al. (2020) adapted Elasticsearch<sup>3</sup> for supporting keyword search over RDF datasets. Gerritse et al. (2020) studied utilizing Wikipedia2Vec (Yamada et al., 2020) for entity reranking. Nikolaev and Kotov (2020) proposed KEWER, a system that employs joint word and entity embeddings to rank entities. KEWER was evaluated on DEv2 but can be applied to any KG and does not require a large textual corpus.

**Question Answering (QA)** QA systems are concerned with automatically providing answers to natural language questions. Such systems have been integrated in search engines and virtual assistants. Answers may be constructed by querying a structured database or pulled from an unstructured collection of documents. Inspired by DrQa (Chen et al., 2017), Mozannar et al. (2019) presented SOQAL, a system for open domain factual Arabic QA using Wikipedia. SOQAL was evaluated on two novel datasets: Arabic Reading Comprehension Dataset and a machine translated version of Stanford QA Dataset. Samy et al. (2019) overview the work done in Arabic QA and recommend tools and linguistic resources for future systems.

**Question Answering over Knowledge Graphs** A special case of QA is KGQA, where the information source is a KG. Unlike entity retrieval, the input for a KGQA system must be structured as a question, and the output is typically one or more answers formed as text. Chakraborty et al. (2019)

<sup>3</sup><https://www.elastic.co>

provides a recent survey on neural methods for KGQA. Abu Taha (2015) and Albarghothi et al. (2017) designed systems for domain specific Arabic KGQA. The tasks of entity retrieval and KGQA may overlap when the query is a question, and the relevant answers are named entities from a KG. DEv2, our chosen test collection, has a category focusing on this special case.

### 3 Dataset

**Queries** DEv2 consists of 467 natural language queries, each with a list of DBpedia entities, scored based on relevance (0: irrelevant, 1: relevant, and 2: highly relevant). The scores are collected using crowdsourcing. In total, there are 49280 query-entity pairs. Table 1 details the four different groups the queries are categorized into. The set creators also offer a “stopped” version<sup>4</sup> which includes the same queries, with stop patterns and punctuation marks removed. Similarly to previous works, we use the stopped version.

**DBpedia** DBpedia provides structured content from the information available on Wikipedia. The test collection is based on DBpedia version 2015-10, released in 2016.<sup>5</sup> The English edition describes 6.2M entities, of which 4.6M have abstracts (required by DEv2), and has 1.1B RDF triples.

**DBpedia Arabic** An Arabic chapter was added to DBpedia in 2015.<sup>6</sup> Ismail et al. (2016) introduced an Arabic endpoint for the chapter and provided a comparison between the Arabic and English editions. While there are later releases, we restrict ourselves to the same release as DEv2 due to the time sensitivity of some queries. The number of entities with abstracts in the Arabic chapter of 2015-10 is 368K (8% of the English chapter). This reflects the gap in coverage between the English and corresponding Arabic editions of Wikipedia. The gap has narrowed since and is expected to narrow further thanks to the increasing popularity of Wikipedia Arabic and to the efforts by the Wikimedia Foundation (Redi et al., 2021).

**DBpedia Arabic Challenges** Prior to its release, several works identified challenges in the path towards an Arabic DBpedia (Al-Feel, 2013; Bahanshal and Al-Khalifa, 2013). Lakshen et al. (2018)

<sup>4</sup><https://github.com/iai-group/DBpedia-Entity>

<sup>5</sup><http://wiki.dbpedia.org/datasets>

<sup>6</sup><http://wiki.dbpedia.org/join/chapters>

Category	Description	Example	#queries	#Arabic
SemSearch-ES	Named entity queries	brooklyn bridge	113	7
INEX-LD	IR-style keyword queries	electronic music genres	99	23
QALD2	Natural language questions	Who is the mayor of Berlin?	140	56
ListSearch	Seek a particular list of entities	Campuses of Indiana University	115	53

Table 1: Query categories in DEv2 and in Arabic DEv2.

discuss challenges in quality assessment of DBpedia Arabic and list issues regarding accuracy, consistency, and relevancy. When working with the DBpedia files, we apply standard text preprocessing, similar to Obeid et al. (2020) and Mohammad et al. (2017). We also replace all instances of Farsi Ye with Arabic Yā.<sup>7</sup> Finally, and due to the inconsistent use of the definite article “al”, we remove it from the beginning of words. Prior to retrieval, we apply the same preprocessing to the queries.

**Searching DBpedia Arabic vs. English** Given the challenges and smaller coverage of Arabic DBpedia, an alternative would be to translate queries from Arabic into English, search the English KG and map the retrieved entities to Arabic ones using DBpedia’s inter-language links. While this approach may work for some tasks, we view it as limited. At least 30% of DBpedia Arabic entities do not have a counterpart in English and capture unique information of high interest to Arabic users. Furthermore, while Wikipedia English allows only one article per entry, content across chapters varies. The availability of other sources and the different points of view mean that Arabic articles are not necessarily a direct translation of English ones. They may include additional or conflicting information.<sup>8</sup> Forcing a search through English content prevents utilization of this diversity and richness. Finally, our aim in this work is to provide a solution for entity retrieval from any Arabic KG, not only the ones that have an English edition.

**Entity Coverage** For each query, DEv2 provides a list of relevant and irrelevant entities, obtained through pooling and crowdsourcing. It does not, therefore, guarantee coverage of all relevant entities, and no system will be rewarded for ranking high a relevant but unjudged entity. This problem is

<sup>7</sup>The two look identical when connected ٱ, but have different Unicode values.

<sup>8</sup>Consider, for example, the popular dish “Mujaddara”. As of Jan 2021, its place of origin according to the the Arabic article is “Mashriq” but “Persia or India” in the English one.

```
<http://ar.dbpedia.org/resource/ابن_خلدون>
<http://www.w3.org/2002/07/owl#sameAs>
<http://dbpedia.org/resource/Ibn_Khaldun>
```

Figure 2: Entity mapping in DBpedia.

exacerbated when mapping entities across chapters.

**Arabic DEv2 Entities** Because we wanted to keep the relevance scores provided in DEv2, we used DBpedia’s inter-language files to programmatically map English entity IDs to their Arabic counterparts. Figure 2 demonstrates how the mapping between the English and Arabic entities of Ibn Khaldun is defined. Once mapped, English entities were not used for retrieval or ranking, and SERAG only relied on the Arabic graph. Due to lower coverage, however, many entities were not mapped. Out of the 16700 relevant entities, only 3025 were successfully mapped to Arabic and a large number of queries did not have a single mapped relevant entity. A query is included in Arabic DEv2 if the number of Arabic relevant entities is at least 10, or if at least half of the relevant entities were mapped. Consequentially, Arabic DEv2 consists of 139 queries. Table 1 shows the number of queries in each category of the Arabic set.

**Query Translation Considerations** Having chosen the queries, the next step is translation to Arabic. We opted for human translation by native Arabic speakers in our organization, rather than machine translation (MT). Many of the queries are not necessarily structured as valid sentences and present a challenge to MT.<sup>9</sup> We wanted the queries in Arabic to be as close as possible to English, including (mis)structure. Furthermore, MT would have presented a significant gender bias. Arabic uses gender-specific terms, while MT is designed to generate a single output text. Even if a translation model is debiased, a query with a single gender term will result in biased

<sup>9</sup>For example, query INEX\_LD-2012327: “Beloved author African-American Nobel Prize Literature”.

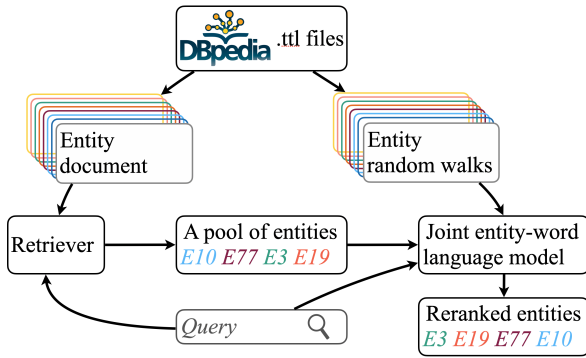


Figure 3: Entity retrieval and re-ranking framework



Figure 4: A sample random walk (top) and its corresponding generated sentence (bottom). Some, but not all, of the IDs are replaced with their surface form.

retrieval. Habash et al. (2019) proposed to extend MT’s output with gender-specific re-inflections. We believe their work is very promising, but it is currently limited to first-person sentences. To avoid such bias, our translators were asked to concatenate both forms where relevant. This, however, comes at a price: the queries are less natural and the duplication may introduce a new bias. An alternative approach, which we plan to explore in future work, is to choose one form at random and rely on the retriever to escape gender bias by leveraging morphological analyzers and language models.

## 4 Method

SERAG, our proposed framework for Arabic entity retrieval, is inspired by KEWER. As illustrated in Figure 3, for each entity, DBpedia is used to generate (1) a document formed of directly linked textual information (recipe provided with DEV2), and (2) a set of random walks starting at that entity. Given a query, BM25 scores entities based on the relevance of their documents to the query. It does not, however, utilize information about the graph structure or entities indirectly connected.

To enhance the ranking, a word2vec (Mikolov et al., 2013) model is built using the random walks.

Each walk consists of entities and predicates along a path in the Arabic DBpedia graph. Entities appear in their entity ID form (e.g., E3) with a probability  $p$  and replaced with their surface form (e.g., Egyptian Museum<sup>10</sup>) otherwise. Each walk is considered a sentence and all sentences are concatenated into one corpus, used to train the word2vec model. Hence, the vocabulary of the model includes both words and entity IDs. Figure 4 illustrates how random walks and sentences are formed.

We use a standard, two-stage ranking approach: given a query  $q$ , we first select 1000 entities using BM25, then we rerank them using embeddings. Let  $e$  be an entity retrieved by BM25,  $v_e$  its embedding, and  $v_q$  the embedding of  $q$ . The final score of  $e$  with respect to  $q$  is based on a mix of its BM25 score and the embeddings’ cosine similarity:

$$MM(q, e) = \beta \cdot \cos(v_q, v_e) + (1 - \beta) \cdot BM25(e)$$

In our implementation, we modified KEWER’s code to allow for Arabic KGs. To keep everything in Python, we reimplemented the document generator and used gensim (Řehůřek and Sojka, 2010) for BM25. While KEWER also uses BM25F (Zhiltsov et al., 2015), a fielded version of BM25, in this work we only consider BM25. KEWER tunes the parameters of BM25 and  $\beta$ , the mixture weight, using a 5-fold split. Our data, however, is prohibitively small. Therefore, SERAG is tested in an unsupervised setup using the default values of gensim, a fixed  $\beta = 0.9$ , chosen to give random walks a higher impact, and the default values KEWER uses for the number of walks (100), their length (10), and  $p$  (0.1). We are confident that SERAG will benefit from parameter tuning and hope to include it when a larger dataset becomes available.

## 5 Experiments and Discussion

The evaluation metric of DEV2 is Normalized Discounted Cumulative Gain (nDCG) at ranks 10 and 100. Table 2 compares the performance of BM25 and SERAG. SERAG yields better results across all categories. In total, its advantage is found to be statistically significant (paired t-test with  $\alpha = 0.05$ ).

In Section 3 we noted that DEV2 in general, and Arabic DEV2 in particular, do not offer judgements for all relevant entities. Buckley and Voorhees (2004) introduced bpref, a retrieval measure designed specifically for incomplete judgment sets.

<sup>10</sup>SERAG uses the Arabic surface form: المتحف المصري



Cat	nDCG@10		nDCG@100	
	BM25	SERAG	BM25	SERAG
Sem	0.186	<b>0.324</b>	0.321	<b>0.463*</b>
INEX	0.215	<b>0.284*</b>	0.293	<b>0.365*</b>
QALD	0.149	<b>0.198*</b>	0.213	<b>0.267*</b>
List	0.174	<b>0.216*</b>	0.256	<b>0.293*</b>
<b>Total</b>	0.172	<b>0.226*</b>	0.248	<b>0.303*</b>

Table 2: Normalized Discounted Cumulative Gain. Best results are in **bold**. \* indicates a statistically significant improvements (paired t-test with  $\alpha = 0.05$ ).

It only considers the relative ranks of relevant and nonrelevant judged entities. We applied bpref to the first 10 and 100 retrieved entities. In both cases, SERAG outperformed BM25 (0.17 vs. 0.142 for top 10 and 0.336 vs. 0.314 for top 100). To the best of our knowledge, bpref has not been used to evaluate English DEv2 systems.

**Query Analysis** To demonstrate the effectiveness of SERAG, we analyze several sample queries. Consider query `SemSearch_LS-50` “wonders of the ancient world”. Of the highly relevant entities (score 2), BM25 ranks only one in the top 10, namely “Hanging Gardens of Babylon” (ranked 8<sup>th</sup>). SERAG ranks this entity higher (3<sup>rd</sup>) and also lists “Temple of Artemis” as 10<sup>th</sup>. Other highly relevant entities outside the top 10 were also ranked higher by SERAG. “Great Pyramid of Giza”, for instance, was ranked 30<sup>th</sup> by SERAG and 59<sup>th</sup> by BM25. As another example, consider query `INEX_LD-2012383` “famous computer scientists disappeared at sea”. In this case there is only one relevant entity, namely “Jim Gray (computer scientist)”. SERAG ranks it first, while BM25 ranks it 5<sup>th</sup>.

**Arabic DEv2 Challenges** Compared to the results reported in literature for English DEv2, Arabic DEv2 proved to be more challenging, and the retrieval effectiveness, as measured by nDCG, was generally lower. Recall the query “wonders of the ancient world”. While SERAG listed only two highly relevant entities in the top 10, KEWER listed five, including, for example, “Great Pyramid of Giza”. Overall, for about a third of the queries, both BM25 and SERAG failed to rank a relevant entity in the top 10. We believe this is because:

- As discussed in Section 3, Arabic DBpedia is much smaller. Some relevant information is missing or mismatched.

- Arabic’s morphological richness presents a challenge to NLP in general and IR in particular (Habash, 2010; Shaalan et al., 2018). This richness leads to MSA verbs with upwards of 5,400 forms, making the task of word-level representation extremely difficult.
- Diacritics are typically omitted in Wikipedia, resulting in morphological ambiguity.<sup>11</sup> Fadel et al. (2019) proposed Translation over Diacritization to assist NLP tasks in such cases.
- Inconsistent choices of synonyms make retrieval much harder. For example, Nobel prize winners are referred to in three ways.<sup>12</sup>
- Spelling of named entities vary. For instance, Google is transcribed in three forms.<sup>13</sup>
- DEv2 was created for English DBpedia. In the pooling stage, entities were obtained using runs of previous methods, optimized for English documents. In the crowdsourcing stage, annotators only considered English content.

## 6 Conclusions

There has been significant progress in offering tools and resources for Arabic NLP applications. In this work, we addressed the task of entity retrieval from an Arabic KG. Our contribution is twofold. We introduced an Arabic version of DEv2, the standard test collection for entity retrieval and then proposed SERAG, a system for end-to-end entity retrieval from Arabic KGs.

We believe Arabic DEv2 can already be used as a benchmark for KG entity retrieval. It is, however, only the first step. We hope to collaborate with the community to expand both the queries and relevant entities. A fresher dump of Wikipedia Arabic should be considered in order to increase coverage. We also plan to translate the collection to other languages and create a multi-lingual set.

While SERAG was shown to improve over BM25, we observed that Arabic entity retrieval faced unique challenges. To address those, we plan to employ the recently introduced Arabic NLP toolkit (Obeid et al., 2020) and utilize pretrained language models such as ArabicBert (Safaya et al., 2020) and AraBert (Antoun et al., 2020).

<sup>11</sup>The word عقد, for instance, can take over a dozen combinations of diacritics, each with multiple meanings.

<sup>12</sup>حاصلون، حائزون، فائزون

<sup>13</sup>جوجل، غوغل، قوقل

## References

- Alaa W. Abu Taha. 2015. [An ontology-based arabic question answering system submitted](#). Master's thesis, Faculty of Information Technology, Islamic University of Gaza.
- Haytham Al-Feel. 2013. [A step towards the arabic dbpedia](#). *International Journal of Computer Applications*, 80:27–33.
- Ali Albarghothi, Feras Khater, and Khaled Shaalan. 2017. [Arabic question answering using ontology](#). *Procedia Computer Science*, 117:183–191.
- Wissam Antoun, Fady Baly, and Hazem Hajj. 2020. [AraBERT: Transformer-based model for Arabic language understanding](#). In *Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools, with a Shared Task on Offensive Language Detection*, pages 9–15, Marseille, France. European Language Resource Association.
- Alia O. Bahanshal and Hend S. Al-Khalifa. 2013. [Toward recipes for arabic dbpedia](#). In *Proceedings of International Conference on Information Integration and Web-Based Applications & Services, II-WAS '13*, pages 331–335, New York, NY, USA. Association for Computing Machinery.
- Chris Buckley and Ellen M. Voorhees. 2004. [Retrieval evaluation with incomplete information](#). In *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '04*, pages 25–32, New York, NY, USA. Association for Computing Machinery.
- H. Cai, V. W. Zheng, and K. Chang. 2018. [A comprehensive survey of graph embedding: Problems, techniques, and applications](#). *IEEE Transactions on Knowledge & Data Engineering*, 30(09):1616–1637.
- Nilesh Chakraborty, D. Lukovnikov, Gaurav Maheshwari, Priyansh Trivedi, Jens Lehmann, and Asja Fischer. 2019. [Introduction to neural network based approaches for question answering over knowledge graphs](#). *ArXiv*, abs/1907.09361.
- Danqi Chen, Adam Fisch, Jason Weston, and Antoine Bordes. 2017. [Reading Wikipedia to answer open-domain questions](#). In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1870–1879, Vancouver, Canada. Association for Computational Linguistics.
- Lisa Ehrlinger and Wolfram Wöb. 2016. [Towards a definition of knowledge graphs](#). In *Joint Proceedings of the Posters and Demos Track of the 12th International Conference on Semantic Systems - SEMANTICS2016 and the 1st International Workshop on Semantic Change & Evolving Semantics (SuCESS'16) co-located with the 12th International Conference on Semantic Systems (SEMANTICS 2016)*, Leipzig, Germany, September 12-15, 2016, volume 1695 of *CEUR Workshop Proceedings*. CEUR-WS.org.
- Ali Fadel, Ibraheem Tuffaha, Bara' Al-Jawarneh, and Mahmoud Al-Ayyoub. 2019. [Arabic text diacritization using deep neural networks](#). *CoRR*, abs/1905.01965.
- Michael Färber, Frederic Bartscherer, Carsten Menne, and Achim Rettinger. 2018. [Linked data quality of dbpedia, freebase, opencyc, wikidata, and yago](#). *Semantic Web*, 9:77–129.
- Dieter Fensel, Umutcan Şimşek, Kevin Angele, Elwin Huaman, Elias Kärle, Oleksandra Panasiuk, Ioan Toma, Jürgen Umbrich, and Alexander Wahler. 2020. [Introduction: What Is a Knowledge Graph?](#), pages 1–10. Springer International Publishing, Cham.
- Emma J. Gerritse, Faegheh Hasibi, and Arjen P. de Vries. 2020. [Graph-embedding empowered entity retrieval](#). *Advances in Information Retrieval*, pages 97–110.
- Nizar Habash, Houda Bouamor, and Christine Chung. 2019. [Automatic gender identification and reinflection in Arabic](#). In *Proceedings of the First Workshop on Gender Bias in Natural Language Processing*, pages 155–165, Florence, Italy. Association for Computational Linguistics.
- Nizar Y. Habash. 2010. [Introduction to arabic natural language processing](#). *Synthesis Lectures on Human Language Technologies*, 3(1):1–187.
- Faegheh Hasibi, Fedor Nikolaev, Chenyan Xiong, Krisztian Balog, Svein Erik Bratsberg, Alexander Kotov, and Jamie Callan. 2017. [Dbpedia-entity v2: A test collection for entity search](#). In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '17*, pages 1265–1268. ACM.
- Ahmed Salama Ismail, Haytham Al-Feel, and Hoda M. O. Mokhtar. 2016. [Introducing a new arabic endpoint for dbpedia internationalization project](#). In *Proceedings of the 20th International Database Engineering & Applications Symposium, IDEAS '16*, pages 284–289, New York, NY, USA. Association for Computing Machinery.
- Giorgos Kadilierakis, Pavlos Fafalios, Panagiotis Papadakis, and Yannis Tzitzikas. 2020. [Keyword search over rdf using document-centric information retrieval systems](#). In *The Semantic Web*, pages 121–137, Cham. Springer International Publishing.
- Guma Abdulkhader Lakshen, Valentina Janev, and Sanja Vraneš. 2018. [Challenges in quality assessment of arabic dbpedia](#). In *Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics, WIMS '18*, New York, NY, USA. Association for Computing Machinery.

- Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. [Distributed representations of words and phrases and their compositionality](#). In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2, NIPS'13*, pages 3111–3119, Red Hook, NY, USA. Curran Associates Inc.
- Abu Bakr Mohammad, Kareem Eissa, and Samhaa El-Beltagy. 2017. [Aravec: A set of arabic word embedding models for use in arabic nlp](#). *Procedia Computer Science*, 117:256–265.
- Hussein Mozannar, Elie Maamary, Karl El Hajal, and Hazem Hajj. 2019. [Neural Arabic question answering](#). In *Proceedings of the Fourth Arabic Natural Language Processing Workshop*, pages 108–118, Florence, Italy. Association for Computational Linguistics.
- Shahrzad Naseri, John Foley, J. Allan, and Brendan T. O'Connor. 2018. Exploring summary-expanded entity embeddings for entity retrieval. In *CIKM Workshops*.
- Fedor Nikolaev and Alexander Kotov. 2020. [Joint word and entity embeddings for entity retrieval from a knowledge graph](#). In *Advances in Information Retrieval*, pages 141–155, Cham. Springer International Publishing.
- Ossama Obeid, Nasser Zalmout, Salam Khalifa, Dima Taji, Mai Oudah, Bashar Alhafni, Go Inoue, Fadhl Eryani, Alexander Erdmann, and Nizar Habash. 2020. [CAMEL tools: An open source python toolkit for Arabic natural language processing](#). In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 7022–7032, Marseille, France. European Language Resources Association.
- Miriam Redi, Martin Gerlach, Isaac Johnson, Jonathan Morgan, and Leila Zia. 2021. [A taxonomy of knowledge gaps for wikimedia projects \(second draft\)](#). *ArXiv*.
- Radim Řehůřek and Petr Sojka. 2010. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pages 45–50, Valletta, Malta. ELRA.
- Ali Safaya, Moutasem Abdullatif, and Deniz Yuret. 2020. [Kuisail at semeval-2020 task 12: Bert-cnn for offensive speech identification in social media](#). *ArXiv*, abs/2007.13184.
- Heba Samy, Ehab E. Hassanein, and Khaled Shaalan. 2019. [Arabic question answering: A study on challenges, systems, and techniques](#). *International Journal of Computer Applications*, 181(44):6–14.
- Khaled Shaalan, Sanjeera Siddiqui, Manar Alkhatib, and Azza Monem. 2018. [Challenges in Arabic Natural Language Processing](#), chapter Chapter 3. World Scientific Publishing Company.
- Ikuya Yamada, Akari Asai, Jin Sakuma, Hiroyuki Shindo, Hideaki Takeda, Yoshiyasu Takefuji, and Yuji Matsumoto. 2020. [Wikipedia2Vec: An efficient toolkit for learning and visualizing the embeddings of words and entities from Wikipedia](#). In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 23–30, Online. Association for Computational Linguistics.
- Nikita Zhiltsov, Alexander Kotov, and Fedor Nikolaev. 2015. [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, SIGIR '15*, pages 253–262, New York, NY, USA. Association for Computing Machinery.