

Construction of a relevance knowledge graph with application to the LOCAL news angle

Bjørnar Tessem*, Marc Gallofré Ocaña and Andreas L. Opdahl

Dept. of Information Science and Media Studies, University of Bergen, Norway

Abstract

News angles are approaches to journalism content often used to provide a way to present a new report from an event. One particular type of news angle is the LOCAL news angle where a local news outlet focuses on an event by emphasising a local connection. Knowledge graphs are most often used to represent knowledge about a particular entity in the form of relationships to other entities. In this paper we see how we can extract a knowledge sub graph containing entities and relevant relationships that are connected to the locality of a news outlet. The purpose of this graph is to use it for automated journalism or as an aid for the journalist to find local connections to an event, as well as how the local connection relate to the event. We call such a graph a relevance knowledge graph. An algorithm for extracting such a graph from a linked data source like DBpedia is presented and examples of the use of a relevance graph in a LOCAL news angle context are provided.

Keywords

News automation, knowledge graphs, news angles, sub graph extraction

1. Introduction

Knowledge graphs have found many applications lately [1]. A central idea is to represent knowledge about a particular entity in the form of relationships to other entities. Examples include personal knowledge graphs [2], company knowledge graphs [3, 4], and topical knowledge graphs applied for example in recommendation systems [5, 6]. Knowledge graphs have thus become a main information structure used in many domains, including analysis and production of news [7, 8, 9, 10].

Here we explore the knowledge graphs in relation to the concept of news angles. A news angle is a concept from journalism, emphasising a view on an event that address and focus on particular aspects of that event [11, pp.781–795]. News angles has been formalised for computational purposes in [12]. One such news angle is the Local or Proximity news angle [12, 13]. This is a news angle often applied by local news papers when relating to events outside of their normal geographical scope. The idea is that the particular event becomes interesting from a local perspective due to its reference to an entity that has local relevance.

NAIS 2023: The 2023 symposium of the Norwegian AI Society, June 14-15, 2023, Bergen, Norway.

*Corresponding author.

✉ bjornar.tessem@uib.no (B. Tessem); marc.gallofre@uib.no (M. Gallofré Ocaña); andreas.opdahl@uib.no (A. L. Opdahl)

ORCID [0000-0003-2623-2689](https://orcid.org/0000-0003-2623-2689) (B. Tessem); [0000-0001-7637-3303](https://orcid.org/0000-0001-7637-3303) (M. Gallofré Ocaña); [0000-0002-3141-1385](https://orcid.org/0000-0002-3141-1385) (A. L. Opdahl)



© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

Motta et al. [12] focuses on the place of an event in their description of the local/proximity angle. Thus only a place indicates the event’s relevance for some local news outlet. In their example formulations in common logic, events that occur in locally relevant local places, but also locally relevant remote places are included. However, they also suggest that events that include local actors could be locally relevant. We suggest that all sorts of entities that are somehow locally relevant would justify a local/proximity angle on an event. This includes places, persons, organisations, natural phenomena, cultural artifacts etc.

Motta et al. suggest using spatial reasoning on places involved in events to enable the local angle. They, however, did not specify how this could be done. Performing a relevance search for each and other event at the time needed would be costly without pre-existing information structures. The semantic web is such an information structure, but also this is too large to enable efficient searches. We suggest an approach where the search for a local news angle should use a knowledge graph that includes places and other entities that have relevance for the local news paper’s area of interest as well as known relationships between them. We call such a graph a relevance graph. The local news paper’s relevance graph may be used when checking external news items for entities involved in external events.

An example would be a local singer mentioned as participating in the choir in a Eurovision Song Contest performance. That would permit a local news angle on the song contest. If the local singer was included in the news paper’s relevance graph, external news items including the singer’s name should be flagged for the journalists. In addition the journalist could be informed about the relationships path from the news paper’s central entity to the singer.

One may of course construct relevance graph manually, but a less costly way is perhaps to use existing semantic web resources to construct this knowledge graph, at least for an initial version. Later on, journalists may themselves be able to extend the graph by supporting tools, combined with repeated construction to include the latest information from the semantic web.

In the remaining of this paper we address the automatic construction of such a relevance graph, using entities and relationships from DBpedia. We suggest various approaches to measuring relevance of an entity for another related entity, formalising what it means to be of relevance for the local news angle. Then we present an algorithm for constructing the relevance graph. We present some examples and discuss the experiences with those, particularly addressing the problems with using a source like DBpedia. We then describe related work relating to sub graph extraction and graph embeddings, before we conclude with ideas for some future work.

2. Relevance measures

We start out with some suggested relevance measures. To define the various measures we need some notation. Assume we have a knowledge graph with entities $E = \{e_1, e_2, \dots, e_n\}$. Further assume we have a collection of binary relations between the entities in E , $\mathcal{R} = \{R_1, R_2, \dots, R_m\}$ where $R_i \subseteq E \times E, i = 1, \dots, m$.

2.1. Simple relevance

The first approach to finding the most important neighbours in the knowledge graph is based on the assumption that neighbours that are connected through many relations should have a

connectivity value relative to the total number of connections. Now assume that the set $F_l(e_i) = \cup_{k=1}^m \{e_j | (e_j, e_i) \in R_k\}$ is the left neighbour set of e_i , and $F_r(e_i) = \cup_{k=1}^m \{e_j | (e_i, e_j) \in R_k\}$ is the right neighbour set of e_i . Then the neighbour set $F(e_i) = F_l(e_i) \cup F_r(e_i)$. For each of the elements $e_j \in F(e_i)$ we can assign a connectivity score

$$c(e_j|e_i) = |\{k | (e_j, e_i) \in R_k, k = 1, \dots, m\}| + |\{k | (e_i, e_j) \in R_k, k = 1, \dots, m\}| \quad (1)$$

A normalized relevance score is further given by

$$\widehat{c}(e_j|e_i) = \frac{c(e_j|e_i)}{\max_{l=1}^n c(e_l|e_i)} \quad (2)$$

The remaining e_j will have a relevance score of 0.0 to e_i .

2.2. Weighted relevance

One way of considering relevance is that relations through which e_i has fewer neighbours should contribute more to the total score for those neighbours. Relationships with fewer examples may convey more information, and may therefore be seen as more important. This can be done by weighting the score from a neighbour from a particular relation by 1 divided by the number of neighbours originating from that relation, i.e., $1/|\{e_j | (e_j, e_i) \in R_k\}|$. An alternative approach is to weigh by $1/\log_2(|\{e_j | (e_j, e_i) \in R_k\}| + 1)$. This gives the following weighted relevance scores wc_1 and wc_2 for entities in $F(e_i)$

$$wc_1(e_j|e_i) = \sum_{k=1}^m \frac{\mathbb{1}_{R_k}((e_j, e_i))}{|\{e_l | (e_l, e_i) \in R_k\}|} + \sum_{k=1}^m \frac{\mathbb{1}_{R_k}((e_i, e_j))}{|\{e_l | (e_i, e_l) \in R_k\}|} \quad (3)$$

$$wc_2(e_j|e_i) = \sum_{k=1}^m \frac{\mathbb{1}_{R_k}((e_j, e_i))}{\log_2(|\{e_l | (e_l, e_i) \in R_k\}| + 1)} + \sum_{k=1}^m \frac{\mathbb{1}_{R_k}((e_i, e_j))}{\log_2(|\{e_l | (e_i, e_l) \in R_k\}| + 1)} \quad (4)$$

and their normalized versions

$$\widehat{wc}_1(e_j|e_i) = \frac{wc_1(e_j|e_i)}{\max_{l=1}^n wc_1(e_l|e_i)} \quad (5)$$

$$\widehat{wc}_2(e_j|e_i) = \frac{wc_2(e_j|e_i)}{\max_{l=1}^n wc_2(e_l|e_i)} \quad (6)$$

2.3. Reciprocal relevance

As we know, relevance among entities in the real world is not a symmetric property. For example, in a local news paper, the local buildings, places, and persons born in the community are important, but the country of the locality may not be so interesting to cover. This even though the country and the locality may have a lot of connections in the overarching knowledge graph. One way of handling this is by considering what we call (neighbour) reciprocal relevance

between the two entities. We consider two such dependent connectivity measures rc_1 and rc_2 . They are based on the normalized expressions found in equations 2, 5 and 6:

$$rc_1(e_j|e_i) = \widehat{c}(e_j|e_i) \cdot \widehat{wc_1}(e_i|e_j) \quad (7)$$

$$rc_2(e_j|e_i) = \widehat{c}(e_j|e_i) \cdot \widehat{wc_2}(e_i|e_j) \quad (8)$$

For practical use the normalised versions are

$$\widehat{rc_1}(e_j|e_i) = \frac{rc_1(e_j|e_i)}{\max_{l=1}^n rc_1(e_l|e_i)} \quad (9)$$

$$\widehat{rc_2}(e_j|e_i) = \frac{rc_2(e_j|e_i)}{\max_{l=1}^n rc_2(e_l|e_i)} \quad (10)$$

These expressions have the effect that the relevance is higher if there are relations in which e_i is one of relatively few entities related to e_j . That is, as compared to the situation where e_i is one of many entities related to e_j .

3. Constructing the relevance graph

As mentioned, to build a relevance graph for a local news paper could be a manual task, but also quite demanding for a single journalist. The web, and in particular Wikipedia, consists of exactly the kind of knowledge about relationships between entities that is of importance in the news context. However, these web pages suffer from

- incompleteness of relevance links
- the emphasis on particular content domains
- links to irrelevant or marginally relevant entities
- erroneous links

These problems is of course due to the fact that these web pages are produced and edited on a voluntary basis. Anyhow, the content and links found in Wikipedia and further structured in DBpedia, is a good starting point for exploring algorithms for constructing relevance graphs.

The algorithm, found in Algorithm 1, for building the graphs starts out with the central entity, the node we want to build a relevance graph for. We expand the graph by first adding entities one link away from the central entity, then those that are two links away, etc., etc. At each new level we first remove entities that are irrelevant by certain heuristics (unspecified in algorithm, see page 6). Next, we compute the relevance value for the entity in mind by computing its relevance to the originating entity at the previous level (Eq 9 or Eq 10) and multiply with that entity's relevance to the central entity. The central entity has relevance 1.0. At each level we expand only those entities that have relevance value above some boundary b_{exp} . If there are several links to an entity from the level below the entity's relevance is increased slightly combining scores from all links (done by combining two scores with the commutative and transitive formula $s = 1.0 - (1.0 - s_1)(1.0 - s_2)$). To keep the graph at a reasonable size we take away those entities that have a score lower than b_{keep} . After having chosen the entities to be included in the graph as nodes, we add to the graph all relationships found in DBpedia among entities in the graph.

Algorithm 1 Get Relevance Graph Nodes

Input c_ent : a DBpedia entity to be the centre of the relevance graph.

```
 $C \leftarrow \text{new Node}(c\_ent, \text{level} = 0, \text{score} = 1.0)$  ▷ Initial node
 $G \leftarrow \{C\}$  ▷ Initialise graph
 $i \leftarrow 1$ 
while  $i \leq \text{max\_level}$  do ▷ Augment graph to level  $i$ 
   $A \leftarrow \emptyset$  ▷ The augmenting set
   $G_{i-1} \leftarrow \{m : m \in G \text{ such that } \text{level}(m) = i - 1 \wedge \text{score}(m) \geq b_{exp}\}$  ▷ Nodes at previous level that are relevant enough to be expanded

  for  $m \in G_{i-1}$  do ▷ Find nodes at level  $i$ 
     $E \leftarrow \{n : (n, r, m) \in \text{DBpedia} \vee (m, r, n) \in \text{DBpedia}\}$ 

    ▷ Remove immediately irrelevant entities. See page 6.

    for  $e \in E$  do ▷ for all candidate entities
       $v = \text{reciprocal relevance score for } e \text{ in relation to } m \text{ (Eq 9 or Eq 10)}$ 
       $w = v \cdot \text{score}(m)$ 

      if  $\text{Node}(e, \text{level}(e), \text{score}(e)) \in A$  ▷ has  $e$  already been created
         $\text{score}(e) \leftarrow \text{adjust}(\text{score}(e), w)$  ▷ adjust relevance score for  $e$ 
      else
         $A \leftarrow A \cup \text{new Node}(e, \text{level} = i, \text{score} = w)$  ▷ new node
      end if

    end for
  end for
   $G \leftarrow G \cup A$ 
end while

for  $\text{Node}(e, \text{level}(e), \text{score}(e)) \in G$  do ▷ remove nodes with too small score
  if  $\text{score}(e) < b_{keep}$ 
     $G \leftarrow G \setminus \{\text{Node}(e, \text{level}(e), \text{rel}(e))\}$ 
  end if
end for
return  $G$ 
```

4. Examples

We have explored this algorithm on two locations. One is Valdres in Norway which is a rural, and mountainous district with six municipalities (Norwegian: kommune) and a population of about 12,000. We experimented with various relevance boundaries for including entities in the graph, and with $b_{exp} = 0.15$ and $b_{keep} = 0.05$ (see Algorithm 1) we got a graph with 929 entities. The graph was visualised with a rdf-visualiser¹. The size and complexity of the graph is indicated in Figure 1. The central entity is found amplified in the middle of the figure. We

¹<https://issemantic.net/rdf-visualizer>

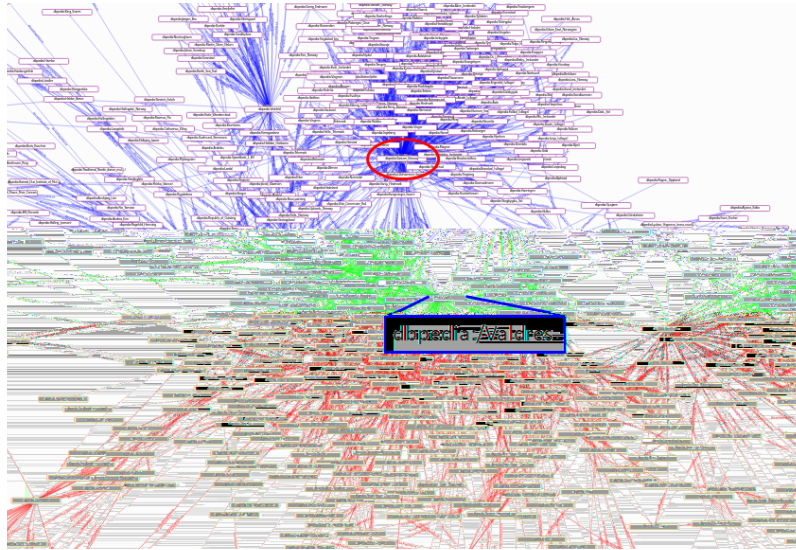


Figure 1: A visualisation of a relevance graph for the DBpedia entity Valdres

also note that there is a hub structure in the graph. For example, there is a hub around the node with a ring, i.e., the node for the entity `dbpedia:Eastern_Norway`.

The inclusion of the entity for Eastern Norway illustrates an issue with use of Wikipedia as a source. The relevance measures would find some relevance to Eastern Norway. But if you know more about Norway, you will understand that this entity is not relevant for the local newspaper of Valdres. Eastern Norway includes Valdres, but like **Scandinavia** or **Norway** it relates to a too big area to be considered an important key word for Valdres. For a large area like **Norway** this is solved as it removed because it passes the limit of 1.000 in going or out going links. This heuristic is used in Algorithm 1 together with some other heuristics to immediately exclude entities from the construction of the relevance graph. The heuristics include

- Entities which have more than 1.000 in going or out going relationships
- Entities that indicates a collection, ending in for example “_of_Norway” (Municipalities_of_Norway)
- Files (start with “File:”)
- Timelines (start with “Timeline_of”)
- Years

Notice that we have to, to some extent, specialize these heuristics to a Norwegian context.

As we see, our heuristics does not exclude Eastern_Norway, as its number of links are fewer than 1.000. To find a useful boundary will be a matter of trial and error, or we need to create more complex methods for assessing immediate exclusion of an entity.

We also explored the small town of Clifden, Ireland as a the central entity. Figure 2 shows the central node and the highest scoring nodes at level one.

We have suggested the use of such graphs to support automatic journalism algorithms that detect events that may have a local connection. Further, we envision that the detected event

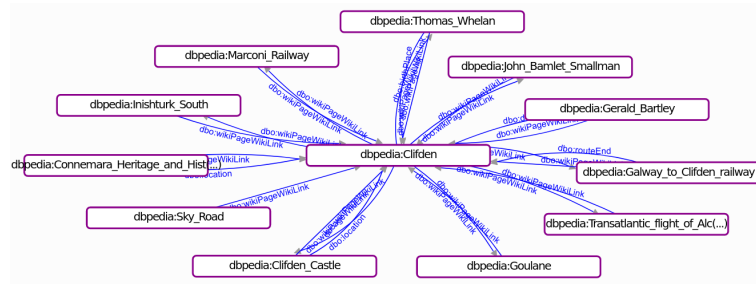


Figure 2: The central entity and close nodes of a relevance graph for Clifden, Ireland

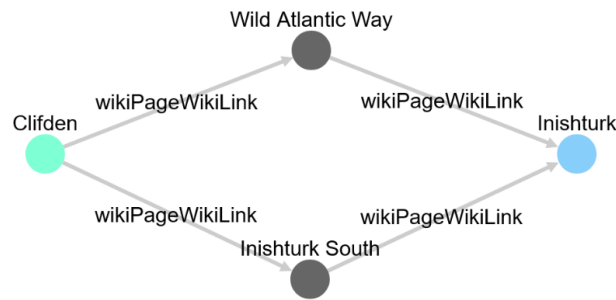


Figure 3: The connection from Clifden to Inishturk in our constructed relevance graph. In this type of visualisation the green node shows the central node, and the blue the entity found identified in a text.

can be framed and presented in the context of a local news paper (using a LOCAL news angle). In our first version we detect external events by performing entity linking on texts from news papers that are collected continuously. To find a relevant event is rare for such rural areas as Clifden or Valdres, but it does happen. In a newspaper in a neighbouring county (“The Connaught Telegraph”) to Clifden we discovered this story:

Achill RNLI assist with medical evacuation on Inishturk. *Achill Island RNLI responded to a request for assistance with the medical evacuation of a patient on Inishturk on bank holiday Monday. ... etc.*

The connection from Clifden is through the “Wild Atlantic Way” as both Clifden and Inishturk are included in the DBpedia description of this tourist attraction (Figure 3). It is moreover interesting to observe that there is also a connection through an entity named “South Inishturk”. This is a small island close to Clifden that also has the name Inishturk. The description of this island mentions that it has the same name as the larger Inishturk, including a link to its Wikipedia page. This illustrates yet an important problem with using DBpedia for such a purpose. The links are general, as most of them are `wikiPageWikiLink` relationships which DBpedia uses to represent links from one Wikipedia page to another. Further, the authors of Wikipedia are not good at categorising the entities, so entities that have no useful meaning in this context do also show up. We already saw this with all the entities in DBpedia that are mainly lists of other entities of a particular type, although no type from the Wikipedia ontology is used neither for the list or the listed entities.

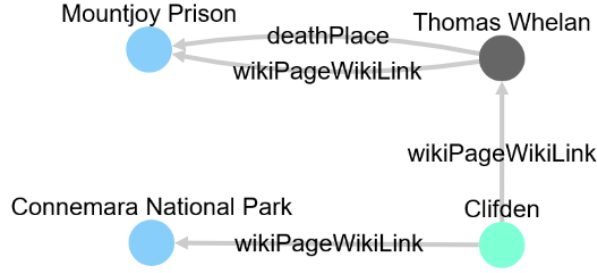


Figure 4: The discovered connections from a Clifden relevance graph are extracted from this tweet: *The Mountjoy Prison is going to be renovated after decades of protests, the minister for justice said during his visit at the Connemara National Park.*

Any texts, not only news reports, could be used as source for identifying local connections. For example, this imagined tweet text does result in the connections from the relevance graph as seen in Figure 4:

The Mountjoy Prison is going to be renovated after decades of protests, the minister for justice said during his visit at the Connemara National Park.

5. Related work

5.1. Sub-graph extraction for news

The idea of extracting subgraphs from open KGs like Wikidata for news-related purposes has already been explored in the literature, although none of the proposals address news angles and locality. An early example is [14], which represents news messages as small KGs with edges extracted from Freebase, in order to support content-based news recommendation.

[15] also represents news articles as small anchor graphs, with edges and K-hop neighbourhoods extracted from Wikidata by a sub-graph extractor trained jointly with a news recommender. The purpose is to provide explainable recommendations in real time.

Context-Aware Graph Embeddings (CAGE) [16] also extracts Wikidata sub-graphs to represent news texts. News-text graph embeddings are concatenated with embeddings that represent user behaviours to offer new recommendations that take short-term user preferences into account.

[17] instead extracts entities and edges from Wikidata to enrich graphs that represent users short- and long-term interests. Entities extracted from candidate news texts are then compared with entities in interest graphs to predict which articles a user may find interesting.

The aim of KLG-GAT [18] is to enhance fact checking and verification by connecting nodes that represent the claim and evidence sentences with triples from Wikidata5M, a subset of Wikidata. A multi-head graph attention network (GAT) is trained to provide input to a claim classifier.

Finally, *NewsLink* [19] aims to support robust and explainable query answering by representing both news texts and user queries as connected sub-graphs extracted from an open KG.

A broader and more detailed review of these and other uses of KGs for the news is presented in [20]. In relation to the work presented here, some of them (e.g., [15, 16, 17]) demonstrate *learned* sub-graph extraction with Graph Neural Networks (GNNs) [21, 22] pointing forward to alternative ways to suggest relevance graphs in future work.

5.2. Graph embeddings

Many of the above sub-graph embedding techniques use graph-embedding techniques to make the structural and symbolic content in knowledge graphs content available for numerical and sub-symbolic deep learning. Graph embeddings represent graph nodes or sub-graphs as vectors in low-dimensional semantic spaces, so that semantic similarity is reflected in spatial distance between nodes and/or that semantic relations between nodes are reflected in spatial directions and distances.

Early graph-embedding approaches exploited the translation characteristic of knowledge graphs to generate node embeddings, so that an edge type used to connect nodes should be represented as a translation in semantic space from its head (or source) node to its tail (or target) [23]. Subsequent approaches were proposed, e.g., to account for one-to-many and many-to-many relations between nodes and for semantically different relations between nodes [24, 25, 26].

Inspired by word embeddings [27], other approaches used parameterised random walks to represent graphs as sets of walks [28, 29], similar to NL sentences, used as inputs to word embedding models like CBOW, Skip-Gram [30] or GloVe [27].

More recently, dedicated Graph Neural Networks (GNNs) have been proposed [21, 22], inspired by deep neural networks used for image and text analysis, including Graph Attention Networks (GATs), Convolutional Graph Networks (CGMs) and Recurrent Graph Networks (RCNs). For example, in an RCN, each node is assigned an initial vector which is updated in subsequent iterations based on the node’s neighbouring vectors until the network converges. The final node vectors can be fed, for example into a node classifier whose loss is used to train the RCN.

In relation to the work presented here, graph embeddings offers an alternative way of discovering local connections to events. Computationally, graph embeddings can be combined with scaling similarity search algorithms such as Meta’s FAISS [31] and HSNW [32] to find connections efficiently. However, they are unlikely to be faster than simple entity look-up in pre-computed relevance graphs, as we propose in this paper. And, unlike the graph-traversal based technique we have described, embedding approaches are unable to *explain* local connections by showing *how* an event is related to the location of interest. They nevertheless offer a way of suggestion connections that are not represented in existing KGs and that should therefore be explored in future work.

6. Conclusions and Further Work

In this paper we have shown how to construct a relevance graph, i.e., a knowledge graph intended to support identification of textual information sources that may be of use for automated journalism. The relevance graph extracts from a larger linked data set, here DBpedia, a smaller

graph of entities that is considered to be of relevance for a news outlet, for example, a smaller local news paper.

The algorithm has been shown to work on a couple of examples, but its application still has to face the problems caused by the lack of structure in Wikipedia, and with that DBpedia. To solve the challenges of DBpedia, one may go to a different source, like Wikidata. This could be supplied with various approaches from network analysis. For example, measures of centrality could perhaps be used to identify irrelevant entities that are highly connected.

It is also pertinent to explore different graph embedding techniques. They can be used to assess similarity between entities. Techniques like e-walks and p-walks [33] can be used as alternatives or complement to the relevance measures described in Section 2. A similar approach can be used to identify new connections. Instead of computing the distance between two known nodes, we can search for the closest nodes in the low-dimensional space to build the location graph. However, some close nodes may lack an explicit connection in the knowledge base, which could be resolved by link prediction models like TransE [23].

Acknowledgments This research is funded by the Norwegian Research Council's IKTPLUSS programme as part of the News Angler project (grant number 275872) and by MediaFutures partners and the Research Council of Norway as part of MediaFutures: Research Centre for Responsible Media Technology & Innovation (grant number 309339).

References

- [1] A. Hogan, C. Gutierrez, M. Cochez, G. d. Melo, S. Kirrane, A. Polleres, R. Navigli, A.-C. N. Ngomo, S. M. Rashid, L. Schmiedl, S. Staab, E. Blomqvist, C. d'Amato, J. E. L. Gayo, S. Numaic, A. Rula, J. Scudà, A. Zimmermann, Knowledge Graphs, Synthesis Lectures on Data, Semantics, and Knowledge, Springer International Publishing, Cham, 2022. doi:10.1007/978-3-031-01918-0.
- [2] K. Balog, T. Kenter, Personal Knowledge Graphs: A Research Agenda, in: Proceedings of the ACM SIGIR International Conference on the Theory of Information Retrieval (ICTIR), 2019, p. 217–220. doi:10.1145/3341981.3344241.
- [3] A. Singhal, Introducing the Knowledge Graph: things, not strings, 2012. URL: <https://blog.google/products/search/introducing-knowledge-graph-things-not/>.
- [4] D. Sullivan, A reintroduction to our Knowledge Graph and knowledge panels, 2020. URL: <https://blog.google/products/search/about-knowledge-graph-and-knowledge-panels/>.
- [5] I. Cantador, P. Castells, A. Bellogín, An enhanced semantic layer for hybrid recommender systems: Application to news recommendation, *Int. J. Semantic Web Inf. Syst.* 7 (2011) 44–78.
- [6] E. Brocken, A. Hartveld, E. de Koning, T. van Noort, F. Hogenboom, F. Frasincar, T. Robal, Bing-CF-IDF+: A Semantics-Driven News Recommender System, in: P. Giorgini, B. Weber (Eds.), *Advanced Information Systems Engineering*, Springer International Publishing, Cham, 2019, pp. 32–47. doi:10.1007/978-3-030-21290-2_3.
- [7] M. Rospocher, M. van Erp, P. Vossen, A. Fokkens, I. Aldabe, G. Rigau, A. Soroa, T. Ploeger,

- T. Bogaard, Building event-centric knowledge graphs from news, *Journal of Web Semantics* 37-38 (2016) 132–151. doi:10.1016/j.websem.2015.12.004.
- [8] C. Rudnik, T. Ehrhart, O. Ferret, D. Teyssou, R. Troncy, X. Tannier, Searching News Articles Using an Event Knowledge Graph Leveraged by Wikidata, in: *Companion Proceedings of The 2019 World Wide Web Conference on - WWW '19*, ACM Press, San Francisco, USA, 2019, pp. 1232–1239. doi:10.1145/3308560.3316761.
- [9] M. Gallofré Ocaña, L. Nyre, A. L. Opdahl, B. Tessem, C. Trattner, C. Veres, Towards a Big Data Platform for News Angles, in: *Proceedings of the 4th Norwegian Big Data Symposium (NOBIDS 2018)*, volume 2316, CEUR-WS.org, Aachen, Germany, 2018, pp. 17–29.
- [10] D. Fernàndez-Cañellas, J. Espadaler, D. Rodriguez, B. Garolera, G. Canet, A. Colom, J. M. Rimmek, X. Giro-i Nieto, E. Bou, J. C. Riveiro, VLX-Stories: Building an Online Event Knowledge Base with Emerging Entity Detection, in: C. Ghidini, O. Hartig, M. Maleshkova, V. Svátek, I. Cruz, A. Hogan, J. Song, M. Lefrançois, F. Gandon (Eds.), *The Semantic Web - ISWC 2019*, Springer International Publishing, Cham, 2019, pp. 382–399. doi:10.1007/978-3-030-30796-7_24.
- [11] P. J. Shoemaker, S. D. Reese, *Mediating the message: theories of influences on mass media content*, 2nd ed., Longmann, White Plains, NY, 1995.
- [12] E. Motta, E. Daga, A. L. Opdahl, B. Tessem, Analysis and Design of Computational News Angles, *IEEE Access* 8 (2020) 120613–120626. doi:10.1109/ACCESS.2020.3005513.
- [13] A. L. Opdahl, B. Tessem, Ontologies for finding journalistic angles, *Software and Systems Modeling* 20 (2020) 71–87.
- [14] K. Joseph, H. Jiang, Content based News Recommendation via Shortest Entity Distance over Knowledge Graphs, in: *Companion Proceedings of The 2019 World Wide Web Conference*, ACM, San Francisco USA, 2019, pp. 690–699. doi:10.1145/3308560.3317703.
- [15] D. Liu, J. Lian, Z. Liu, X. Wang, G. Sun, X. Xie, Reinforced Anchor Knowledge Graph Generation for News Recommendation Reasoning, in: *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, ACM, Virtual Event Singapore, 2021, pp. 1055–1065. doi:10.1145/3447548.3467315.
- [16] H.-S. Sheu, Z. Chu, D. Qi, S. Li, Knowledge-Guided Article Embedding Refinement for Session-Based News Recommendation, *IEEE Transactions on Neural Networks and Learning Systems* (2021) 1–7. doi:10.1109/TNNLS.2021.3084958.
- [17] Y. Tian, Y. Yang, X. Ren, P. Wang, F. Wu, Q. Wang, C. Li, Joint Knowledge Pruning and Recurrent Graph Convolution for News Recommendation, in: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, ACM, Virtual Event Canada, 2021, pp. 51–60. doi:10.1145/3404835.3462912.
- [18] B. Zhu, X. Zhang, M. Gu, Y. Deng, Knowledge Enhanced Fact Checking and Verification, *IEEE/ACM Transactions on Audio, Speech, and Language Processing* (2021) 3132–3143. doi:10.1109/TASLP.2021.3120636.
- [19] Y. Yang, Y. Li, A. K. H. Tung, NewsLink: Empowering Intuitive News Search with Knowledge Graphs, in: *2021 IEEE 37th International Conference on Data Engineering (ICDE)*, IEEE, Chania, Greece, 2021, pp. 876–887. doi:10.1109/ICDE51399.2021.00081.
- [20] A. L. Opdahl, T. Al-Moslmi, D.-T. Dang-Nguyen, M. Gallofré Ocaña, B. Tessem, C. Veres, Semantic Knowledge Graphs for the News: A Review, *ACM Computing Surveys* 55 (2022) 140:1–140:38. doi:10.1145/3543508.

- [21] P. Goyal, E. Ferrara, Graph embedding techniques, applications, and performance: A survey, *Knowledge-Based Systems* 151 (2018) 78–94. doi:10.1016/j.knosys.2018.03.022.
- [22] H. Cai, V. W. Zheng, K. C.-C. Chang, A Comprehensive Survey of Graph Embedding: Problems, Techniques, and Applications, *IEEE Transactions on Knowledge and Data Engineering* 30 (2018) 1616–1637. doi:10.1109/TKDE.2018.2807452.
- [23] A. Bordes, N. Usunier, A. Garcia-Durán, J. Weston, O. Yakhnenko, Translating embeddings for modeling multi-relational data, in: *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2, NIPS'13*, Curran Associates Inc., Red Hook, NY, USA, 2013, pp. 2787–2795.
- [24] Y. Lin, Z. Liu, M. Sun, Y. Liu, X. Zhu, Learning Entity and Relation Embeddings for Knowledge Graph Completion, *Proceedings of the AAAI Conference on Artificial Intelligence* 29 (2015). doi:10.1609/aaai.v29i1.9491.
- [25] Z. Wang, J. Zhang, J. Feng, Z. Chen, Knowledge graph embedding by translating on hyperplanes., in: *Aaai*, volume 14, Citeseer, 2014, pp. 1112–1119. Issue: 2014.
- [26] G. Ji, S. He, L. Xu, K. Liu, J. Zhao, Knowledge graph embedding via dynamic mapping matrix, in: *Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing (volume 1: Long papers)*, 2015, pp. 687–696.
- [27] J. Pennington, R. Socher, C. Manning, Glove: Global Vectors for Word Representation, in: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Association for Computational Linguistics, Doha, Qatar, 2014, pp. 1532–1543. doi:10.3115/v1/D14-1162.
- [28] A. Grover, J. Leskovec, node2vec: Scalable Feature Learning for Networks, in: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*, Association for Computing Machinery, New York, NY, USA, 2016, pp. 855–864. doi:10.1145/2939672.2939754.
- [29] A. Narayanan, M. Chandramohan, R. Venkatesan, L. Chen, Y. Liu, S. Jaiswal, graph2vec: Learning Distributed Representations of Graphs, *arXiv:1707.05005 [cs]* (2017). URL: <http://arxiv.org/abs/1707.05005>, arXiv: 1707.05005.
- [30] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, J. Dean, Distributed Representations of Words and Phrases and their Compositionality, in: C. J. Burges, L. Bottou, M. Welling, Z. Ghahramani, K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems*, volume 26, Curran Associates, Inc., 2013, pp. 3111–3119. URL: https://proceedings.neurips.cc/paper_files/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf.
- [31] J. Johnson, M. Douze, H. Jegou, Billion-Scale Similarity Search with GPUs, *IEEE Transactions on Big Data* 7 (2021) 535–547. doi:10.1109/TBDATA.2019.2921572.
- [32] Y. A. Malkov, D. A. Yashunin, Efficient and Robust Approximate Nearest Neighbor Search Using Hierarchical Navigable Small World Graphs, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42 (2020) 824–836. doi:10.1109/TPAMI.2018.2889473.
- [33] J. Portisch, H. Paulheim, Walk This Way!, in: P. Groth, A. Rula, J. Schneider, I. Tiddi, E. Simperl, P. Alexopoulos, R. Hoekstra, M. Alam, A. Dimou, M. Tamper (Eds.), *The Semantic Web: ESWC 2022 Satellite Events*, Springer International Publishing, Cham, 2022, pp. 133–137. doi:10.1007/978-3-031-11609-4_25.