An analysis of approach to the fake news assessment based on the graph neural networks

Ihor A. Pilkevych[1], Dmytro L. Fedorchuk[1], Mykola P. Romanchuk[1] and Olena M. Naumchak[1]

[1]*Korolyov Zhytomyr Military Institute, 22 Myru Ave., Zhytomyr, 10004, Ukraine*

Abstract

The experience of Russia's war against Ukraine demonstrates the relevance and necessity of understanding the problems of constant disinformation, the spread of propaganda, and the implementation of destructive negative psychological influence. The issue of dissemination in online media informational messages containing negative psychological influence was researched. Ways of improving the system of monitoring online media using the graph neural networks are considered. The methods of automated fake news detection, based on graph neural networks, were reviewed. The purpose of the article is the analysis of existing approaches that allow identifying destructive signs of influence in text data. It is found that the best way to automate the content analysis process is to use the latest machine learning methods. It was determined and substantiated that graph neural networks are the most reliable and effective solution for the specified task. An approach to automating this procedure based on graph neural networks has been designed and analyzed, which will allow timely and efficient detection and analysis of fake news in the information space of our country. During the research, the process of detecting fake news was simulated. The obtained results showed that the described models of graph neural networks can provide good results in solving the tasks of timely detection and response to threats posed by fake news spread by Russia.

Keywords

graph neural networks, psychological influences, fake news, knowledge graph, information messages, online media, information war

1. Introduction

There is more than one definition of the war waged by Russia against Ukraine, in particular: "hybrid war", "new generation war", "subversive war", "information war". Each of these concepts focuses on the use of non-military means in modern warfare. The importance of the information sphere of confrontation in modern wars has grown significantly in recent years. Information technologies are becoming one of the most promising types of weapons. Every year, the scope of its application increases primarily due to its ease of use.

doors-2023: 3rd Edge Computing Workshop, April 7, 2023, Zhytomyr, Ukraine

igor.pilkevich@meta.ua (I. A. Pilkevych); fedor4uk.d@gmail.com (D. L. Fedorchuk); romannik@ukr.net (M. P. Romanchuk); olenanau@gmail.com (O. M. Naumchak)

https://ieeexplore.ieee.org/author/37089181628 (I. A. Pilkevych); https://ieeexplore.ieee.org/author/37089179622 (D. L. Fedorchuk); https://ieeexplore.ieee.org/author/37087013658 (M. P. Romanchuk); https://ieeexplore.ieee.org/author/37089181640 (O. M. Naumchak)

0000-0001-5064-3272 (I. A. Pilkevych); 0000-0003-2896-3522 (D. L. Fedorchuk); 0000-0002-0087-8994 (M. P. Romanchuk); 0000-0003-3336-1032 (O. M. Naumchak)

The official military doctrine of the Russian Federation calls for "simultaneous pressure on the enemy throughout its territory in the global information space". The Internet is used to spread propaganda, misinformation, manipulation of facts, including fake news, etc. The experience of the war of the Russian Federation against Ukraine showed that the enemy widely uses the capabilities of the global network to spread negative psychological influences as a means of waging a hybrid war [1].

From the first day of its independence, our country became the object of Russian propaganda and the direction of concentrated and powerful destructive psychological influences [2]. In particular, Russia's special units widely use the Internet to distribute negative psychological influences to target audiences [3] in distributed special materials of negative psychological influences which have the form of text messages. Therefore, the search for ways to counteract the aggressor's special operations is a relevant research direction.

Special information operations of the Russian Federation are aimed at key democratic institutions (in particular, electoral ones), and special services of the aggressor state are trying to intensify internal contradictions in Ukraine and other democratic states. The Russian hybrid warfare technologies against Ukraine, including information intervention models and mechanisms, are spreading to other states, quickly adapting to local contexts and regulatory policies [4]. Restrictive measures (sanctions) and responsibility for their violation and an effective mechanism for monitoring the information space are one of the effective mechanisms for responding to disinformation and propaganda activity in the Russian Federation [5].

The availability of online media, the rapidly growing number of sources of information (such as news sites, social networks, blogs, websites, etc.) and the ease with which they can be used to spread information quickly lead to the problem of the viral spread of fake news. The popularization of social networks has exacerbated this long-standing problem [6]. Now, fake news has become a major problem for society and individuals, as well as for organizations and governments fighting disinformation and propaganda [7].

It should be noted that at the current stage, scientific interest is not the amount of information and its constant growth, but the structure of distributed data and their relationship. That is why one of the urgent tasks is the creation of a unique collection of knowledge. For this, first of all, it is necessary to automate the processes of collecting, analyzing, and summarizing data from the network. And the requirements for knowledge will be: the ability to read and understand them both by an automated system and by a person, their structure and sequence.

A modern tool for presenting and preserving knowledge is knowledge graphs (KG). KG is a graph in which vertices are unique entities, and edges are connections between them and their attributes. The advantages of KG include: the ability to model both abstract concepts and real objects; the ability to think about new connections between existing entities; the ability to generate new knowledge based on existing knowledge (creation of new entities).

KG are somewhat similar to relational databases (DBs), but their main difference is semi-structuredness and underlying logical apparatus. (DBs are completely structured and therefore not "flexible" and not suitable for solving a large number of tasks). For example, KG are currently used in such fields as information search, natural language processing; semantic technologies that allow using the semantic load of data in the analysis; machine learning, generation of new knowledge, etc.

The use of KG in the field of processing natural language texts can allow automating the process of monitoring the information space. The purpose of the study is to analyze the approaches and choose the most effective one for building a knowledge graph for detecting fake news (informational messages containing negative psychological influences.

The first knowledge base, on the basis of which the KG was implemented, was DBpedia, which contains about 6 billion related entities, created on the basis of semantic processing of articles from Wikipedia [8]. The most famous example is the Google Knowledge Graph. Other implementations are YAGO [9], WordNet, NELL [10], Freebase (since 2014 as part of Google Knowledge Graph), Wikidata graph [11], LOD Cloud [12] and other.

Wikidata is an open, collaboratively edited knowledge base created to present information in a compatible machine-readable format. The actual information from Wikidata conforms to the RDF data model, where entities are represented as triplets (s, p, o). Other information can be added to the entity description. In [13] other formats were also considered. In particular, they use a variant of the RDF format – named graphs in the form of quads, where a fourth element is added to the usual triplet (s, p, o, i). Where i is additional identifier.

Named graphs extend the RDF ternary model and consider sets of pairs in the form $G(n)$, where G is RDF-graph, n is IRI or an empty node in some cases, or maybe even for the default graph. We can smooth this representation by concatenating $G \cdot \{n\}$ for each such pair, resulting in fours. Thus, we can encode the quad (s, p, o, i) directly using N-Quads.

KG accumulate knowledge not only in a human-friendly form, like Wikipedia, but also in a machine-intelligible form, creating a basis for machine learning and solving intellectual tasks in various fields.

For the research being conducted, GIS can be an effective tool in solving the task of automating the process of collecting and analyzing data from the information space. Namely, the processing of text data from social Internet services for the purpose of identifying signs of negative psychological influence and, if possible, finding its original source, author, determining the purpose of distribution, target audience, to which the psychological influence is directed, etc.

2. Method

An example of the construction of a KG when solving the problem of analyzing natural language texts.

Having a certain text at the input, the first task is to highlight the named entities and the connections between them, combining the received facts into a graph. For visualization, we will use the metafactory platform, which uses the Wikidata knowledge graph. For example, let's take an article from Wikipedia about Ukraine. Several key points can be identified from the text. For example, language, neighbors, population and start building a graph (figure 1).

We select the predicate "shared border with..." and select the entities corresponding to it. The platform allows you to select all predicates connecting the selected entities for visualization at once. Particular attention is drawn to the size of the graph containing only a few entities and the predicates connecting them.

Therefore, "Ukraine" is the essence of the KG, which is connected with other entities in the form of triplets (s, p, o) or (h, r, t), where s ad o represent entities, p – connection between them. In the case of a built-in GK, examples of linked triplets for the entity "Ukraine" would

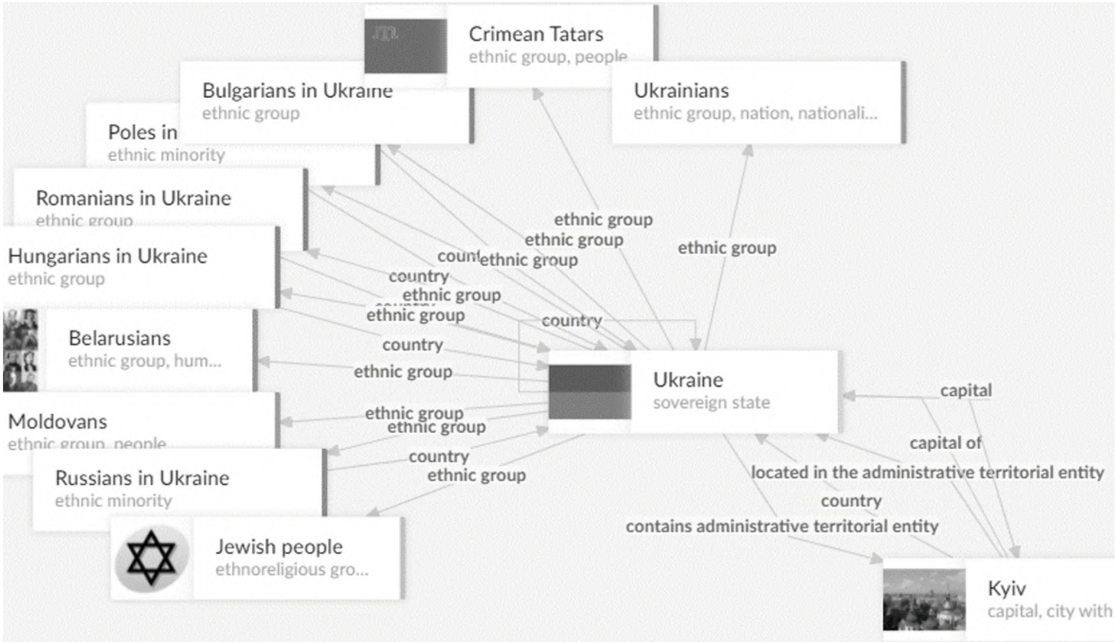

Figure 1: Visualization of the constructed knowledge graph.

be (Ukraine, capital, Kyiv) and (Ukraine, ethnic_groups, Ukrainians), (Ukraine, ethnic_groups, Crimean Tatars), etc.

The use of the KG as a basis for the encoder of entities is effective for several reasons: the distribution of information within the graph allows combining information about the object itself and about its neighbors in the representation of the object; there are several large-scale open source KG.

As mentioned earlier, KG can be presented in two ways. The first is an ontological representation based on formal logic and semantics. The second – vector representation – uses statistical mechanisms to minimize the distances between close entities in multidimensional spaces.

A comparison of the approaches is presented in table 1.

The main difference between the considered approaches is that the symbolic representation

Table 1
Comparison of the ontological and vector representation of the KG.

Representation	Ontological	Vector
What is it based on?	formal logic (propositional, predicate logic, modal, first-order logic, etc.); semantics	statistics; vector distances
Approaches (standards)	RDF, OWL_1, OWL_2, etc.	GCN, GNN, GAN, TextGCN, etc.
Presentation of data	XML, Turtle, RDFa, JSON-LD, etc.	Embeddings
Formal description	(s, p, o), $p(s, o)$, s, p, o	$s, p, o \in \mathbb{R}^d$

implies the recording of facts using symbols (for example, RDF triplets), while in the vector representation the essence and predicates are projected into some d-dimensional space (embedding space).

The main idea of the vector representation is to search for a graph vertices mapping function in a vector space of a certain dimension. That is, a network is taken, fed to the input of a parametric function-encoder, and at the output we get vector representations.

The disadvantage of methods based on shallow learning is transductivity – the model learns vector representations for vertices once and must be retrained every time the graph changes. Also, the disadvantage is that wandering around the graph is random, so the model will produce different results (representations) each time.

Deep models – graph neural networks (GNN) – are free from the mentioned shortcomings. The main idea of which is to build a computational graph for each vertex, the features of which are determined by the features of its neighbors through a non-linear aggregator. GNN are capable of processing graphically structured data. Other types of neural networks work with tabular data, image data (pixel grid), or text data.

In table 2 shows examples of existing models of graph neural networks and areas (problems) in which they are used.

The application of GNN allows prediction to be performed both at the level of nodes and at the level of connections (edges). This allows us to predict certain properties of unlabeled nodes based on other nodes and their edges. As for the edges, the prediction of the occurrence of connections between the vertices in the future can be performed. GNNs can classify nodes or predict connections in a network by studying the embedding of nodes. These embeddings are low-dimensional vectors that summarize the positions of nodes in the network as well as the structure of their local neighborhood. It is also possible to perform graph-level prediction based on the structural properties of these graphs when the input data is the complete graph. Such a model can be used, for example, to solve the problem of detecting fake news. Fake news is a phenomenon of modern propaganda and disinformation, which is widely used by the Russian Federation in conducting hybrid warfare.

In [14] a three-stage approach to the analysis of fake news using KG is proposed:

Stage 1 – Encoder of news – coding of the title.
Stage 2 – Encoder of entities – identification of named entities, coding of individual objects using KG.
Stage 3 – Classification of news – final study and classification of news (using, for example, GNN).

Based on this and [15], we have the following steps of the GNN model:

1) embedding nodes is done using several rounds of message passing:
2) combining node embeddings into a single graph embedding (called a reading layer, for example: global mean pool);
3) classifier training based on graph embedding.

The architecture of the GNN model is shown in the figure 2.

Table 2
Existing models of graph neural networks and areas in which they are applied.

Field of application	Tasks	Algorithm	Model
Text	Text classification	GCN	Graph Convolutional Network
		GAT	Graph Attention Network
		DGCNN Text GCN	Graph Convolutional Network
		Sentence LSTM	Graph LSTM
	Marking sequences	GraphSAGE	GraphSAGE
		GAT	Graph Attention Network
		Sentence LSTM Tree LSTM	Graph LSTM
	Classification by tonality	GraphSAGE	GraphSAGE
	Neural machine translation	GAT	Graph Attention Network
		Syntatic GCN	Graph Convolutional Network
		GGNN	Gated Graph Neural Network
	Edge extraction	Tree LSTM Graph LSTM	Graph LSTM
		GCN	Graph Convolutional Network
	Event extraction	Syntatic GCN	Graph Neural Network
		GraphSAGE	GraphSAGE
		GAT	Graph Attention Network
	Text generation	GGNN	Gated Graph Neural Network
		Sentence LSTM	Graph LSTM
	Reading comprehension	GraphSAGE	GraphSAGE
		GAT	Graph Attention Network
	Relational thinking	RNN	MLP Reccurent Neural Network
Image	Image classification	GCN DGP	Graph Convolutional Network
		GSNN GGNN	Gated Graph Neural Network
	Visual answers to questions		
	Interaction detection	GPNN Strucrural-RNN	Graph Neural Network
	Region classification	GNN DGCNN	Graph Convolution Network
	Semantic segmentation	GGNN	Gated Graph Neural Network
		Graph LSTM	Graph LSTM
		3DGNN	Graph Neural Network
Knowledge Graphs	Completed knowledge bases	GNN	Graph Neural Network
	Alignment of knowledge graphs	GCN	Graph Convolutional Network

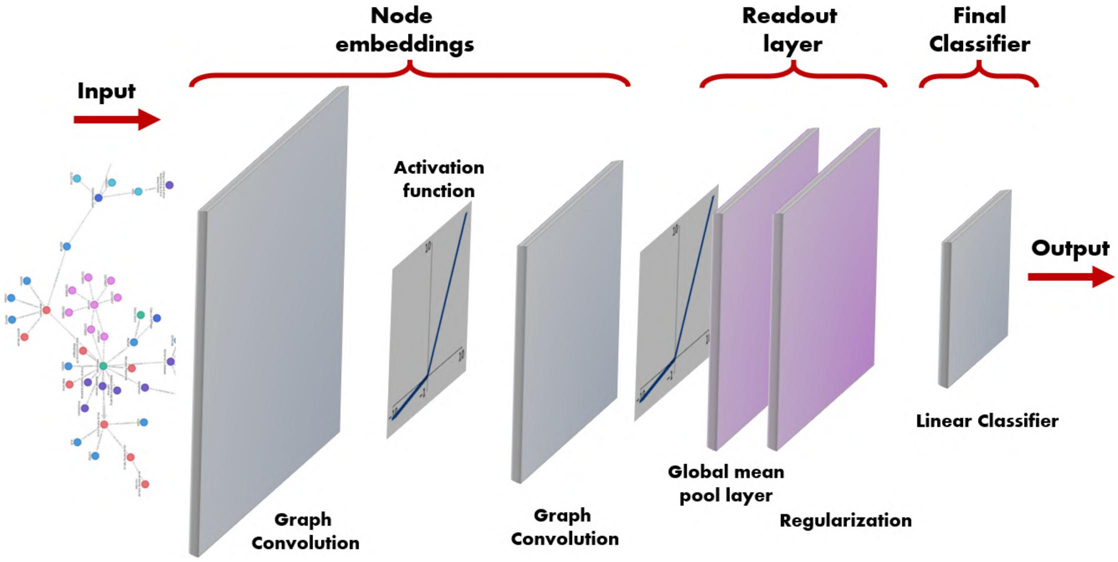

Figure 2: Architecture of the GNN model.

3. Results

The User Preference-aware Fake News Detection (UPFD) data set was used to study the application of the proposed GNN model [16]. This dataset consists of fact-checked fake and real news stories received and distributed on Twitter by Politifact and GossipCop [17]. About 20 million messages from users involved in spreading fake news were processed. Nodes of the data set are characterized by four types of features, held due to the use of pre-trained models of the transformer, word2vec and from the profile of the Twitter account, its comments. The data was split into two datasets: the training set, which contains about 70% of the total dataset, and the test set, which contains the rest of the dataset.

The solution was built on the basis of GCN, GAT [18] and GraphSAGE [19] models. Models

Table 3
The results were obtained during model training.

	GCN		GAT		GraphSAGE	
Politifact	Loss	Accuracy	Loss	Accuracy	Loss	Accuracy
profile	0.2587	0.7873	0.1544	0.7557	0.0476	0.8009
spaCy	0.0417	0.7907	0.0415	0.7919	0.0266	0.8100
BERT	0.0079	0.8371	0.0071	0.8326	0.0013	0.8462
content	0.0560	0.8869	0.0363	0.8959	0.0180	0.8978
Gossipcop	Loss	Accuracy	Loss	Accuracy	Loss	Accuracy
profile	0.2441	0.9038	0.1890	0.9140	0.1633	0.9258
spaCy	0.1010	0.9634	0.1129	0.9597	0.0584	0.9681
BERT	0.0347	0.9660	0.0170	0.9698	0.0135	0.9757
content	0.1082	0.9663	0.0822	0.9773	0.0698	0.9801

were trained using cross-entropy losses with class weights. They are evaluated according to the average accuracy measured on the test sets. The selection of hyperparameters consisted of the type and number of GNN convolutions used for node embedding, the activation function, and the learning rate. GNN models were trained for 100 epochs. The results obtained during model training are shown in table 3.

As can be seen from the results of model training, the best results were obtained when using the GraphSAGE model. The advantage of the GraphSAGE model compared to other GNN models is that it uses only a set of fixed size formed by uniform sampling for aggregation.

Therefore, to solve the problem, it is advisable to use the GraphSAGE model trained on selected text data containing signs of negative psychological influence. Such a model will be able to analyze and detect textual data containing destructive content with signs of negative psychological impact in the process of online media monitoring. An important condition is the availability of a significant amount of training data for training the model.

4. Conclusions and future work

Therefore, the issue of analyzing messages from online mass media for the purpose of detecting fake news remains relevant and has become more acute in the conditions of a large-scale war. In order to timely identify and respond to the negative impact that spreads through such messages, it is necessary to improve monitoring systems. The article developed and analyzed an approach to the automation of this process based on graph neural networks, which will allow timely and qualitative detection and analysis of fake news in the information space of our country.

KG can be used to supplement training samples for machine learning algorithms, which allows improving the performance of applications with a limited amount of training data – for example, systems for analyzing the tonality (sentiment analysis) of messages to determine the level of negative impact; vocal expressions. Since the KG contains auxiliary factual information about the elements contained in the training samples (entities from the texts on which the model is trained), it helps to expand its functionality. This addition increases the accuracy of classification when detecting fake news.

A perspective direction for further research is to increasing the level of automation of content analysis, in particular textual information, by developing and implementing methods of automatic semantic analysis of texts and determining their content based on neural networks, in particular, using graph classification, regression, and clustering.

References

[1] T. Vakaliuk, I. Pilkevych, D. Fedorchuk, V. Osadchyi, A. Tokar, O. Naumchak, Methodology of monitoring negative psychological influences in online media, Educational Technology Quarterly 2022 (2022) 143–151. doi:10.55056/etq.1.

[2] O. V. Levchenko, O. M. Kosogov, Methods of identification of events of negative information influence based on the analysis of open sources, Systemy obrobky informatsii 1 (2016) 100–102. URL: http://nbuv.gov.ua/UJRN/soi_2016_1_22.

[3] G. E. Howard, M. Czekaj (Eds.), Russia's Military Strategy and Doctrine, Jamestown Foundation, 2019, pp. 159–184.

[4] S. Woolley, P. Howard, Automation, Algorithms, and Politics| Political Communication, Computational Propaganda, and Autonomous Agents — Introduction, International Journal of Communication 10 (2016). URL: https://ijoc.org/index.php/ijoc/article/view/6298.

[5] Stratehiia informatsiinoi bezpeky, 2021. URL: https://www.president.gov.ua/documents/6852021-41069.

[6] H. Allcott, M. Gentzkow, Social Media and Fake News in the 2016 Election, Journal of Economic Perspectives 31 (2017) 211–36. doi:10.1257/jep.31.2.211.

[7] A. J. Berinsky, Rumors and Health Care Reform: Experiments in Political Misinformation, British Journal of Political Science 47 (2017) 241–262. doi:10.1017/S0007123415000186.

[8] P. N. Mendes, M. Jakob, A. García-Silva, C. Bizer, DBpedia Spotlight: Shedding Light on the Web of Documents, in: Proceedings of the 7th International Conference on Semantic Systems, I-Semantics '11, Association for Computing Machinery, New York, NY, USA, 2011, p. 1–8. doi:10.1145/2063518.2063519.

[9] T. Rebele, F. Suchanek, J. Hoffart, J. Biega, E. Kuzey, G. Weikum, YAGO: A Multilingual Knowledge Base from Wikipedia, Wordnet, and Geonames, in: The Semantic Web – ISWC 2016: 15th International Semantic Web Conference, Kobe, Japan, October 17–21, 2016, Proceedings, Part II, Springer-Verlag, Berlin, Heidelberg, 2016, p. 177–185. URL: https://suchanek.name/work/publications/iswc-2016-yago.pdf. doi:10.1007/978-3-319-46547-0_19.

[10] J. M. Giménez-García, M. C. Duarte, A. Zimmermann, C. Gravier, E. R. H. Jr., P. Maret, NELL2RDF: Reading the Web, Tracking the Provenance, and Publishing it as Linked Data, in: S. Capadisli, F. Cotton, J. M. Giménez-García, A. Haller, E. Kalampokis, V. Nguyen, A. P. Sheth, R. Troncy (Eds.), Joint Proceedings of the International Workshops on Contextualized Knowledge Graphs, and Semantic Statistics co-located with 17th International Semantic Web Conference (ISWC 2018), volume 2317 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2018. URL: https://ceur-ws.org/Vol-2317/article-02.pdf.

[11] A. Waagmeester, G. Stupp, S. Burgstaller-Muehlbacher, B. M. Good, M. Griffith, O. L. Griffith, K. Hanspers, H. Hermjakob, T. S. Hudson, K. Hybiske, S. M. Keating, M. Manske, M. Mayers, D. Mietchen, E. Mitraka, A. R. Pico, T. Putman, A. Riutta, N. Queralt-Rosinach, L. M. Schriml, T. Shafee, D. Slenter, R. Stephan, K. Thornton, G. Tsueng, R. Tu, S. Ul-Hasan, E. Willighagen, C. Wu, A. I. Su, Science Forum: Wikidata as a knowledge graph for the life sciences, eLife 9 (2020) e52614. doi:10.7554/eLife.52614.

[12] A. Jentzsch, Linked Open Data Cloud, in: T. Pellegrini, H. Sack, S. Auer (Eds.), Linked Enterprise Data: Management und Bewirtschaftung vernetzter Unternehmensdaten mit Semantic Web Technologien, X.media.press, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, pp. 209–219. doi:10.1007/978-3-642-30274-9_10.

[13] D. Hernández, A. Hogan, M. Krötzsch, Reifying RDF: What Works Well With Wikidata?, in: T. Liebig, A. Fokoue (Eds.), Proceedings of the 11th International Workshop on Scalable Semantic Web Knowledge Base Systems co-located with 14th International Semantic Web Conference (ISWC 2015), Bethlehem, PA, USA, October 11, 2015, volume 1457 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2015, pp. 32–47. URL: https://ceur-ws.org/Vol-1457/

SSWS2015_paper3.pdf.
[14] M. Mayank, S. Sharma, R. Sharma, DEAP-FAKED: Knowledge Graph based Approach for Fake News Detection, in: 2022 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2022, pp. 47–51. doi:10.1109/ASONAM55673.2022.10068653.
[15] I. Pilkevych, D. Fedorchuk, O. Naumchak, M. Romanchuk, Fake News Detection in the Framework of Decision-Making System through Graph Neural Network, in: 2021 IEEE 4th International Conference on Advanced Information and Communication Technologies (AICT), 2021, pp. 153–157. doi:10.1109/AICT52120.2021.9628907.
[16] Y. Dou, K. Shu, C. Xia, P. S. Yu, L. Sun, User Preference-Aware Fake News Detection, in: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 2051–2055. doi:10.1145/3404835.3462990.
[17] N. Vo, K. Lee, PolitiFact, 2020. URL: https://paperswithcode.com/dataset/politifact.
[18] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, Y. Bengio, Graph Attention Networks, 2018. arXiv:1710.10903.
[19] W. L. Hamilton, R. Ying, J. Leskovec, Inductive Representation Learning on Large Graphs, in: Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17, Curran Associates Inc., Red Hook, NY, USA, 2017, p. 1025–1035. doi:10.5555/3294771.3294869.