

In the context of the global information war, propaganda is disseminated massively and systematically through mass media, news portals and social networks to manipulate public mood and behavior. This is aimed at demoralizing, having a psychological effect and indoctrinating. Social groups, caught in the bubbles of propagandistic information flows, become radicalized and polarized not only in terms of political but also social behavior, which leads to the loss of (i) mutual agreement, (ii) trust in government institutions, (iii) the cohesion of civil society, (iv) the ability to resist external influences, and (v) the overall social capital of society. The objective of this study is to review the latest five-year deep learning applications for automated propaganda detection. We apply PRISMA systemic review research method and corresponding meta-analysis. Obtained insights and observations serve to develop further a decision-support tool to detect propaganda in news articles and social media. For that matter, we explore the best methods to detect propaganda signs, classify them according to types of propaganda and rank them in English and Lithuanian language. In the presented research, we pay particular attention to the best deep learning models capable of fine-tuning the data to the specific language context.

Introduction
The sudden onset of the information age has led to a proliferation of controversial information in the media (social networks, mass media), while societies were not yet, and still are not yet, prepared for it, as they do not have the skills and natural immunity to recognize and filter propaganda, thus creating the perfect environment for the spread of propaganda. The media and social networks are two of the main tools used to systematically manipulate, divide, or psychologically influence society. In the context of information warfare, groups of society caught in the bubbles of propaganda information flows are radicalized and reduce the cohesion of society and, more generally, its social capital. In this respect, an automated propaganda identification and classification tool based on AI could make it easier for media and social network users to identify and stop the spread of propaganda messages. A systematic review of the research will identify the most effective deep learning approaches for automated propaganda detection. The results of this systematic review will be used to develop an AI-based automated propaganda detection tool.

Goals
The application of deep learning methods to propaganda detection will be explored under three themes: propaganda author detection, propaganda text detection, and the social impact of propaganda. First, three research reviews will be made:

1. A systematic research review of deep learning methods for propaganda author detection. The goals of this review are to identify which deep learning methods are applied and show the best results for propaganda identification from author characteristics, as well as what data is used for such models' training.
2. A systematic research review of deep learning methods for propaganda content detection. The goals of this review are to identify which deep learning methods are applied and show the best results for propaganda identification from text, as well as what data is used for such models' training.
3. A systematic research review of deep learning methods for propaganda social impact analysis. The goals of this review are to identify which deep learning methods are applied and show the best results for propaganda social impact analysis, as well as what data is used for such models' training.

These three reviews will be merged into one, focusing on the application of deep learning techniques to propaganda detection by using different data.

Related work

According to [3], hybrid DL models can better capture the syntactic representation of text, extract multiple feature maps, and improve text classification results. A systematic review of deep learning methods for fake news detection [4] revealed that various neural networks (RNN, LSTM, GRU, GNN, GAN, BERT) methods are already used for such tasks. According to [4], most deep learning approaches use similar architectures, where NLP techniques are used for data preprocessing and neural network models for fake news detection, more details in figure 2. Yi-Ju Lu, Cheng-Te Li combined three different approaches: user-based, content-based, structure – based into one GCAN (Graph – aware Co – Attention network) model (more details in figure 3) and outperformed state – of – the – art methods by 16% in accuracy on average. T. Zhang, D. Wang, H. Chen, Z. Zeng, W. Guo, C. Miao, and L. Cui [2] presented a BERT-based (BDANN) domain – adaption neural network for multimodal false news, pre – trained BERT (figure 4) model was used to extract text features and pre-trained VGG-19 model was used to extract image features, more details in figure 5. There are numerous [1], [2], [6], [7], [8], [9], and [10] hybrid deep learning approaches to detect fake news, with the majority of them developed in the last three years, indicating that interest in such research is growing.

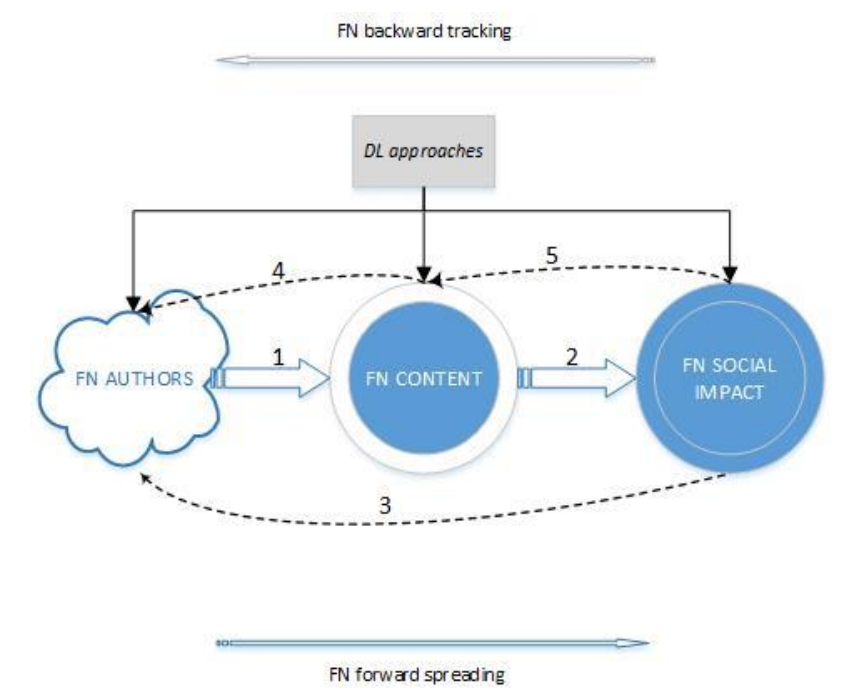


Figure 1. Systematic Review Research framework

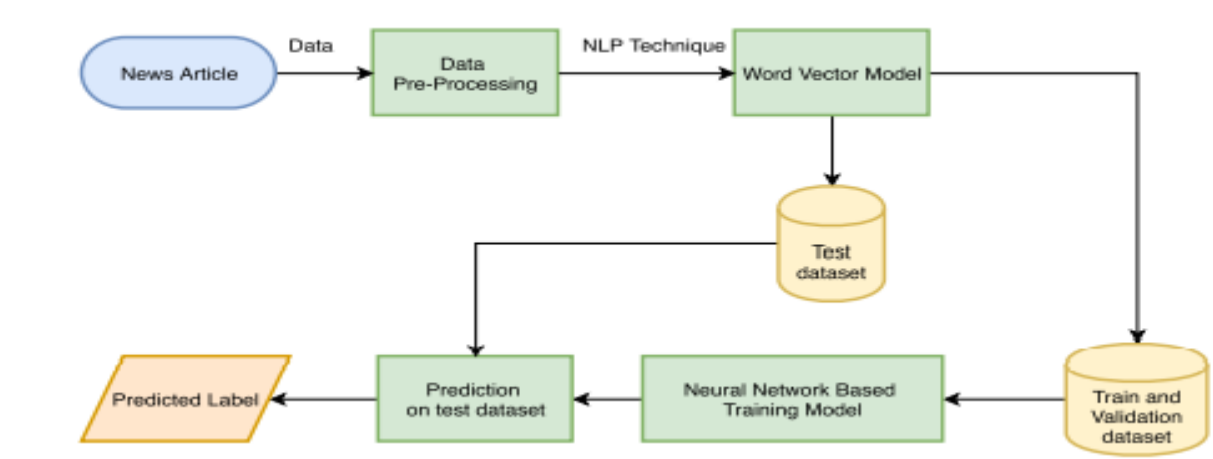


Figure 2. General DL-based architecture for fake news detection [4].

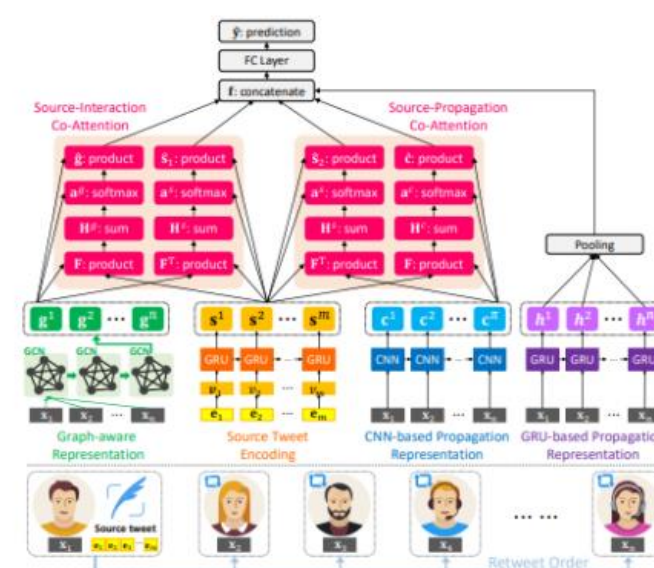


Figure 3. GCAN model [1].

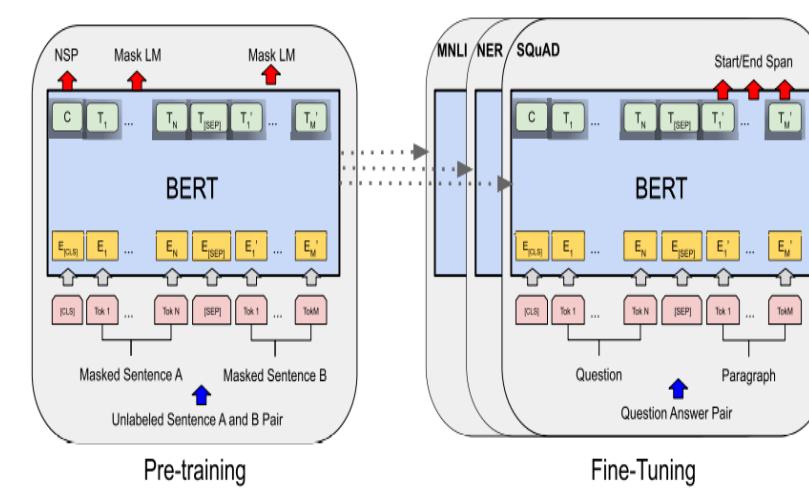


Figure 4. BERT model [5].

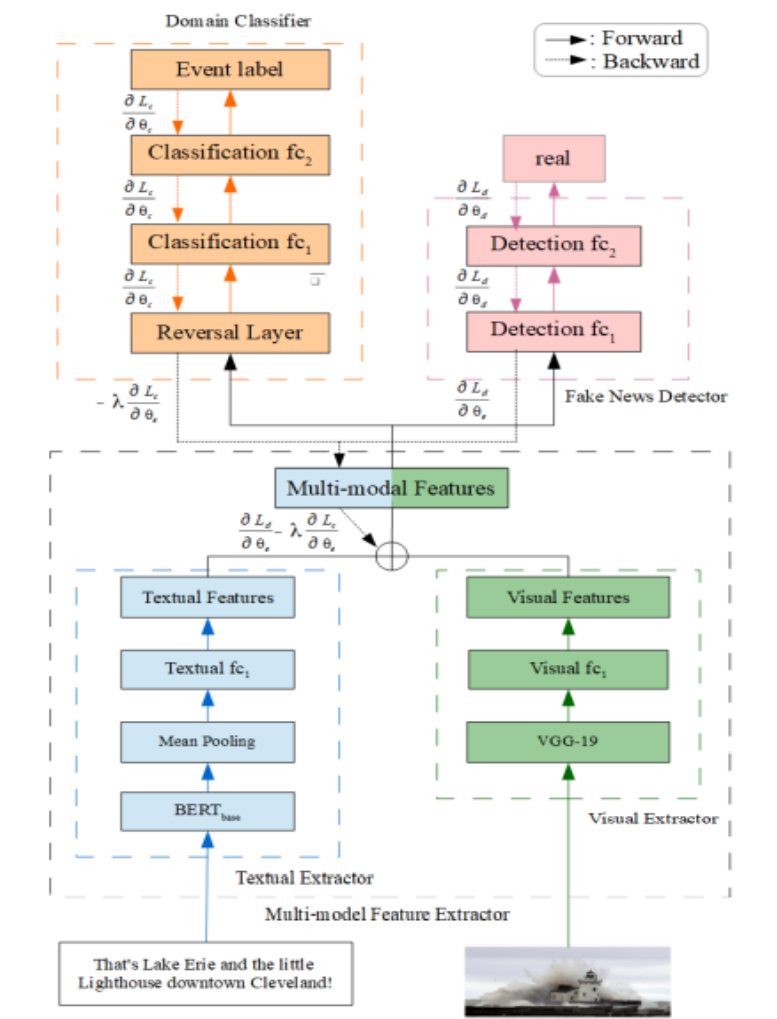


Figure 5 BDANN model [2].

PRISMA methodology

We apply PRISMA systemic review research method and corresponding meta-analysis. According to [11], The Preferred Reporting Items for Systematic reviews and Meta – Analyses (PRISMA) statement was developed to facilitate transparent and complete reporting of systematic reviews and has been updated (to PRISMA 2020) to reflect recent advances in systematic review methodology and terminology, we will use the latest PRISMA 2020 version, more details in flow diagram (figure 6). The PRISMA 2020 statement includes a checklist of 27 items to guide reporting of systematic reviews. In our systematic review, the meta-analysis will consist of eight main criteria: creator analysis, type of news analysis, news content analysis, DL methods analysis, learning strategies, social context analysis, target audience analysis, more details are provided in figure 7.

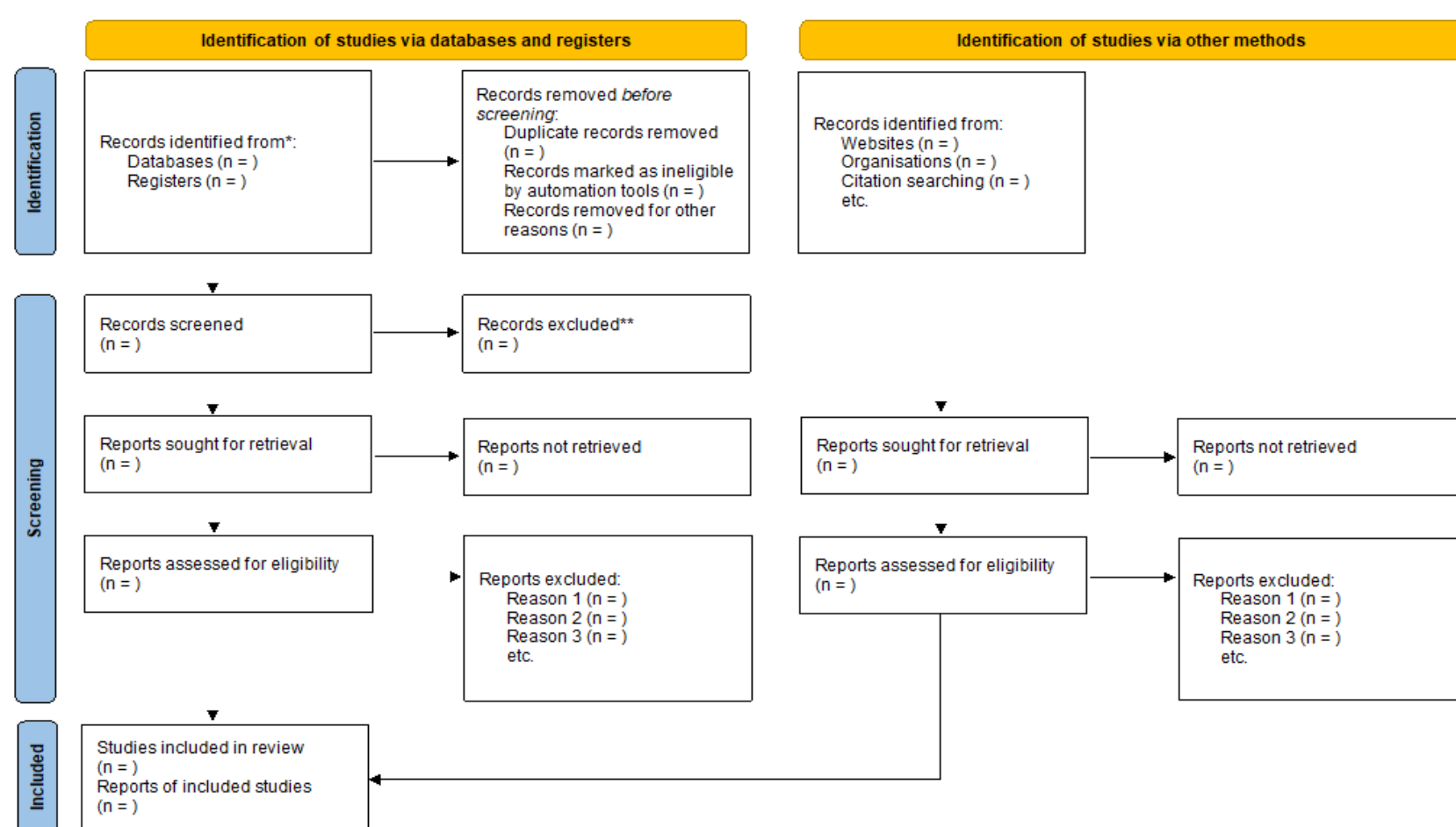


Figure 6 PRISMA 2020 flow diagram for new systematic reviews which included searches of databases, registers and other source [11].

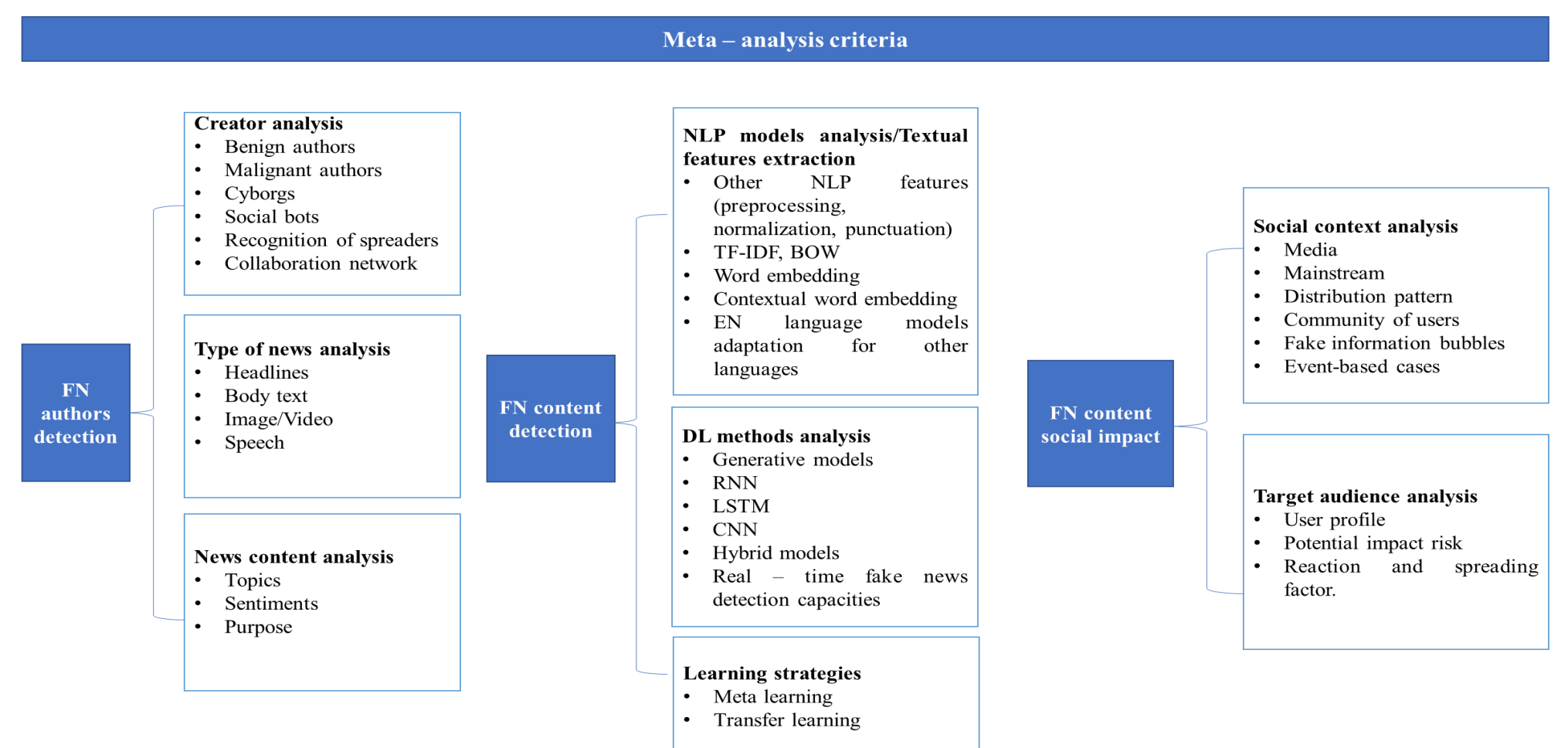


Figure 7 Systematic Review Research meta-analysis criteria

Systematic Review Research process

For a systematic research review of propaganda detection using author characteristics, deep learning approaches that use data on author characteristics will be analyzed. It is expected to find such deep learning approaches that can learn the differences between the characteristics of users who spread propaganda and those who post information.

For a systematic research review of propaganda detection using textual data, deep learning approaches that use textual data will be analyzed. It is expected to find what deep learning methods can learn the differences between propagandic and non-propagandic text content.

For a systematic research review of social impact identification using deep learning methods, it will be analyzed deep learning approaches that identifies target audience and modeling interactions between the propaganda spread methods and society's reactions. It will be determined how deep learning methods can learn to predict interactions between propaganda spread methods and social impact.

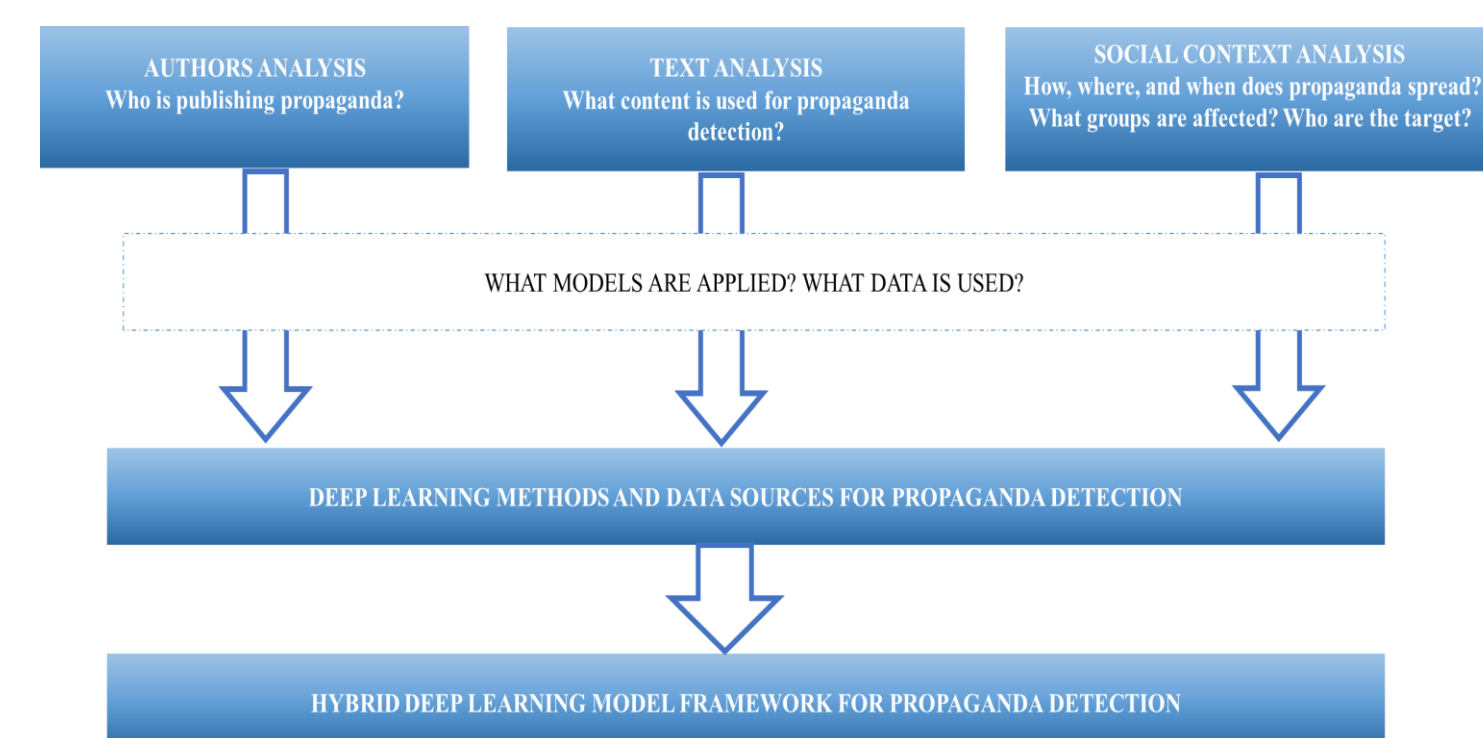


Figure 8. Systematic Review Research process

Results

- It is expected that after this systematic research review, we will identify:
- deep learning approaches, which show the best results in propaganda detection tasks by using different data sources;
 - data sources, which can be used to train deep learning methods for propaganda detection.

The results of this research will be applied to our artificial intelligence-based automatic propaganda detection tool, which we are planning to apply for Lithuanian language.

References

1. Yi-Ju Lu, Cheng-Te Li, “GCAN: Graph-aware Co-Attention Networks for Explainable Fake News Detection on Social Media”, 2020, arXiv:2004.11648.
2. T. Zhang, D. Wang, H. Chen, Z. Zeng, W. Guo, C. Miao, and L. Cui, ‘BDANN: BERT-based domain adaptation neural network for multimodal fake news detection.’ in Proc. Int. Joint Conf. Neural Netw. (IJCNN), Jul. 2020, pp. 1–8.
3. Akpatsa, Samuel & Li, Xiaoyu & Lei, Hang. (2021). A Survey and Future Perspectives of Hybrid Deep Learning Models for Text Classification. 10.1007/978-3-030-78609-0_31.
4. M. F. Mridha, A. J. Keya, M. A. Hamid, M. M. Monowar and M. S. Rahman, "A Comprehensive Review on Fake News Detection With Deep Learning," in IEEE Access, vol. 9, pp. 156151-156170, 2021, doi: 10.1109/ACCESS.2021.3129329.
5. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” 2018, arXiv:1810.04805.
6. Ge, X.; Hao, S.; Li, Y.; Wei, B.; Zhang, M. Hierarchical Co-Attention Selection Network for Interpretable Fake News Detection. Big Data Cogn. Comput. 2022, 6, 93. <https://doi.org/10.3390/bdcc6030093>.
7. Wei-Yao Wang, Wen-Chih Peng, Team Yao at Factly, “Utilizing Pre-trained Models and Co-attention Networks for Multi-Modal Fact Verification”, 2022, arXiv:2201.11664v1.
8. S. Ni, J. Li and H. -Y. Kao, “MVAN: Multi-View Attention Networks for Fake News Detection on Social Media,” in IEEE Access, vol. 9, pp. 106907-106917, 2021, doi: 10.1109/ACCESS.2021.3100245.
9. Mudith Dhawan, Shakshi Sharma, Aditya Kadam, Rajesh Sharma, Ponnurangam Kumaraguru, “GAME-ON: Graph Attention Network based Multimodal Fusion for Fake News Detection”, 2022, arXiv:2202.12478.
10. Xinyu Cui and Yang Li, “Fake News Detection in Social Media based on Multi-Modal Multi-Task Learning” International Journal of Advanced Computer Science and Applications(IJACSA), 13(7), 2022. <http://dx.doi.org/10.14569/IJACSA.2022.01307106>
11. Page M J, Moher D, Bossuyt P M, Boutron I, Hoffmann T C, Mulrow C D et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews BMJ 2021; 372 :n160 doi:10.1136/bmj.n160