Сырым Жақыпбеков – магистр, «Халықаралық ақпараттық технологиялар университеті» АҚ, Алматы қ., Қазақстан; e-mail: syrymzhakypbekov@gmail.com. ORCID: https://orcid.org/0000-0001-9112-5922.

Камила Бакенова – PhD студент, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан; e-mail: bakenova.k@bk.ru. ORCID: https://orcid.org/0009-0004-2567-173X.

Information about the authors

Arailym Kenzhebay* – doctoral student, Alikhan Bokeikhan University, Semey, Kazakhstan; e-mail: salik2000@icloud.com. ORCID: https://orcid.org/0009-0002-5465-0634.

Assemgul Tynykulova – PhD, acting Associate Professor, Astana International University, Astana, Kazakhstan; e-mail: asem_110981@mail.ru. ORCID: https://orcid.org/0000-0002-4557-6869.

Gilshat Yensebayeva – Candidate of Pedagogical Sciences senior teacher department «Information and technical sciences», Alikhan Bokeikhan University, Semey, Republic of Kazakhstan; email: enss_gulshat@mail. ru. ORCID: https://orcid.org/0009-0003-4625-3898.

Syrym Zhakypbekov – Master of Computer Science, International Information Technology University, Almaty, Kazakhstan; e-mail: syrymzhakypbekov@gmail.com. ORCID: https://orcid.org/0000-0001-9112-5922.

Kamila Bakenova – PhD student. L.N. Gumilyov Eurasian National University, Astana, Kazakhstan; e-mail: bakenova.k@bk.ru. ORCID: https://orcid.org/0009-0004-2567-173X.

Received 06.01.2026
Revised 05.03.2026
Accepted 10.03.2026

https://doi.org/10.53360/2788-7995-2026-1(21)-24

IRSTI: 20.23.25

Zh.B. Lamasheva[1,2], M.A. Sambetbayeva[1,2,*], A.N. Nekessova[1,2,3], B.Kh. Abdygalym[1,2,3], N. Tasbolatuly[1,3]

[1]International Science Complex Astana,
010000, Republic of Kazakhstan, Astana, avenue. Kabanbai Batyr 8
[2]L.N. Gumilyov Eurasian National University,
010008, Republic of Kazakhstan, Astana, Satpayev str. 2.
[3]Astana International University,
010000, Republic of Kazakhstan, Astana, avenue. Kabanbai Batyr 8
*e-mail: sambetbayeva_ma_1@enu.kz

INTELLIGENT FAKE NEWS DETECTION SYSTEM BASED ON ONTOLOGICAL MODEL AND SEMANTIC MARKUP OF MEDIA TEXTS

Abstract: *This paper presents a framework for fake news detection based on ontological modeling and semantic annotation of media texts. The study includes a bibliometric analysis of Scopus publications from 2018 to 2026 to identify trends in fake news detection and semantic approaches. The results show a shift from traditional machine learning methods toward transformer-based, graph-based, and semantic models. An adaptive system architecture is proposed, covering data collection, preprocessing, multi-level annotation in Label Studio, knowledge graph construction, model training, inference, and analytics. A formal ontological model was developed to structure the key elements of news texts, including claim, source, evidence, author intent, target audience, and disinformation techniques. The framework supports multilingual processing in Kazakh and Russian. The dataset consists of 5,000 news articles, evenly distributed between fake and real categories and balanced across both languages. Annotation quality was evaluated using Cohen's Kappa, with values ranging from 0.72 to 0.81, indicating consistent inter-annotator agreement. The proposed approach provides a structured basis for the further development and evaluation of automated fake news detection systems in multilingual environments.*

Key words*: Fake news, fake news detection, ontological model, semantic markup, semantic analysis.*

Introduction

In recent years, fake news detection has evolved from a *classic* textual classification problem to an interdisciplinary direction at the intersection of computational linguistics, computer vision, and graph theory. This was facilitated by the explosive growth of user content, the acceleration of disinformation cycles and the spread of large language/multimodal models, which increased the requirements for methods: they are expected to be able to identify semantic inconsistencies, consider visual signals, and rely on structural knowledge about sources and distribution. At the same time, the field remains methodologically heterogeneous: deep representations coexist with the features of the *old school*, and graph and multimodal solutions place increased demands on data, annotations, and validation.

The study [1] proposes an approach to increasing the reliability of automated fact-checking by filtering leaks and unreliable sources. The authors presented the CREDULE dataset, which includes more than 90 thousand news articles classified as reliable, unreliable and fact-checking, and also developed a neural network model EVER-Net, capable of automatically detecting *leaky* and unreliable evidence. The results demonstrate a significant improvement in the accuracy of automatic fact-verification systems and highlight the need for a comprehensive assessment of evidence sources in the fight against misinformation. The FactCheck Editor system is presented in [2], which provides automated fact checking in a multilingual environment. The developed tool integrates the functions of identifying statements requiring verification, generating search queries and extracting relevant evidence from open sources. Verification is carried out using Natural Language Inference (NLI) models and leverages LLM, which allow not only for determining the validity of statements but also for suggesting their correct formulations. The authors have shown that the XLM-Roberta-large model, which has been refined on multilingual data, surpasses the large language models (GPT-3,5-Turbo, Mistral-7b) in solving problems related to statements detection and truth prediction. The work emphasizes the potential of combining transformational and generative models to create of full-fledged automated fact-checking systems.

The study [3] introduces FactCheck Editor, which implements multilingual automated fact checking based on transformational models and LLMs, allowing the system to identify, verify, and correct false statements in the text. In modern research, the methods of machine (mL) and deep learning (DL) are actively used to solve the problems of information reliability analysis [4-6]. In the reviewed works, it is shown that the use of neural network architectures, especially models based on transformers (BERT, Roberta), significantly increases the effectiveness of detection of fake news and automatic fact-checking. These approaches demonstrate high accuracy and stability in the processing of texts of various subjects.

Related works

Modern research in the field of fake news detection focuses on a combination of machine learning techniques, natural language analysis, and graph analysis [7-8]. Particular attention is paid to the integration of ontological models, which allow the formalization of subject areas and increase the interpretability of the results. The use of ontologies contributes to a deeper understanding of the context, the identification of hidden semantic connections between objects and the refinement of semantic features, which increases the accuracy of the classification of news reports. Adaptive architectures of intelligent systems are also actively being developed, capable of dynamically updating knowledge and analysis parameters depending on changes in the information environment. Such solutions provide adaptation to the emergence of new types of disinformation and create the basis for constructing explainable and context-sensitive models for detecting fake news. Earlier, in [9], the authors proposed for the first time proposed a multi-level annotation model adapted to Kazakh and Russian languages, which goes beyond the binary classification. Unlike simple models focused only on the detection of authenticity, this scheme takes into account the deep semantic structure of the text, discursive features and cultural-specific elements inherent in the Kazakh and Russian media space. The multilevel scheme allows for describing not only the fact of the presence of disinformation, but also its type, context, motive and degree of distortion of the content.

Bibliometric analysis was performed on a sample of Scopus for 2018-2026 at the request of key («fake news detection» AND semantic) in the subject area of Computer Science (with intersections with other areas). The array includes 128 documents from 66 sources, prepared by 417 authors; the average number of co-authors per publication is 4.39, the share of international co-authorship is 21.09% (see Fig. 1). The cumulative average annual growth rate is 5.2%, the average age of the document is 1.68 years, the average number of citations per publication is 14.38, indicating the relevance of the topic and active knowledge exchange.

Figure 1 – Bibliometric survey of publications (2018-2025)

There has been a steady increase from a single publication in 2018-2020 to a local high in 2024 (≈39 articles), followed by a moderate decline in 2025 (≈33) and incomplete reporting for 2026 due to the current Scopus indexing window (see Fig. 2a). The acceleration after 2022 coincides with the proliferation of large language models (LLMs) and the transition to multimodal architectures, which is further manifested in thematic clusters.

The bibliometric profile indicates the maturity and diversification of the area: the volume of publications and the range of sources are growing, international co-authorship is strengthening, and the thematic dictionary (≈329 author keywords) demonstrates a high level of fragmentation. For 2023-2025, three leading trajectories are clearly distinguished: semantic (knowledge-guided) NLP, multimodal fusion, and graph-based detection. In the following sections, we examine the evolution of methods along these lines and identify key research gaps, particularly in multilingual and low-resource contexts (including the Kazakh-Russian corpus and the HITL annotation).

A diagram created by VOSviewer (see Fig. 2b) shows the relationships between the most common terms in scientific publications on the topic of fake news detection and semantics. The size of the nodes corresponds to the frequency of term occurrence, and the distance between them reflects the strength of their co-occurrence. Colors represent thematic clusters:

red – methods of machine learning and natural language processing;

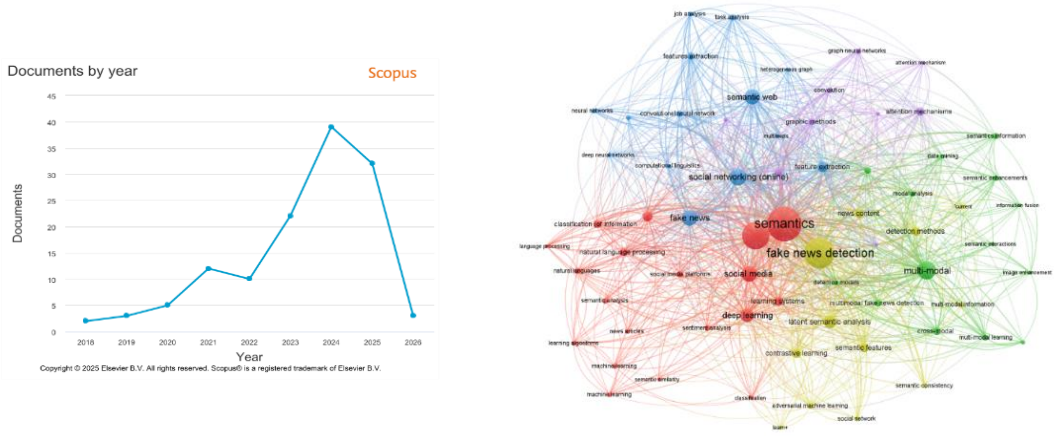
green – multimodal approaches and image analysis;

blue – social networks and information dissemination;

purple – trust, emotions and interpretation of content;

yellow – semantic and contextual aspects of data analysis, including thematic modeling and knowledge representation.

The central position of the terms “fake news detection” and “semantics” underscores their key role in the structure of the research field.



(a) (b)

Figure 2 – a) Documents by year (2018-2026, Scopus). Local maximum in 2024; 2026 is incomplete; b) Visualization of a network of key terms in the field of detection of fake news

As shown in Table 1, existing approaches focus on evidence verification, multilingual pipelines, contextual retrieval, or graph neural network modeling. However, none of them provide a formally defined ontology within a unified semantic framework integrated with knowledge graph

construction and multi-level annotation. Current graph-based solutions rely primarily on structural representations without explicit semantic formalization. The proposed model differs in that it combines an explicit ontology, knowledge graph construction, transformer-based architectures, and structured multi-level annotation, while also addressing a low-resource Kazakh-Russian multilingual environment.

Materials and Methods

Figure 3 presents the full data processing pipeline, including data collection from external media sources, preprocessing, multi-level semantic annotation, knowledge graph construction, transformer-based model training, inference, and analytics. The architecture implements a closed machine learning cycle (collection → processing → annotation → training → inference → feedback) and is designed to support ontology-driven multilingual (Kazakh-Russian) misinformation detection. The figure illustrates the hypothesis that integrating semantic annotation, knowledge graphs, and transformer architectures within a unified pipeline enables structured and interpretable fake news analysis.

Table 1 – Comparison of Selected Fake News Detection Approaches

Ref	Core Method	Ontology (Formalized)	Knowledge Graph	GNN	Multilingual	Low-Resource Focus	Multilevel Semantic Annotation	Explainability Mechanism	Adaptability to New Sources	Transformer-based
[1]	Evidence verification (NN)	No	No	No	Limited	No	No	Limited (evidence scoring)	Limited	Yes
[2]	End-to-end multilingual fact-hecking	No	No	No	Yes	No	No	Limited (NLI reasoning)	Moderate	Yes
[7]	Contextualized evidence retrieval	No	Partial	No	Yes	No	No	Retrieval-based	Moderate	Yes
[8]	Heterogeneous GNN + LLM	No explicit ontology	Yes	Yes	Limited	No	No	Graph attention weights	Moderate	Yes
Proposed model	Ontology + KG + Transformers + Multilevel Annotation	Yes (explicit ontology)	Yes (explicit KG)	Optional	Kazakh-Russian	Yes	Yes (structured)	Ontology-driven interpretability	High (adaptive pipeline)	Yes

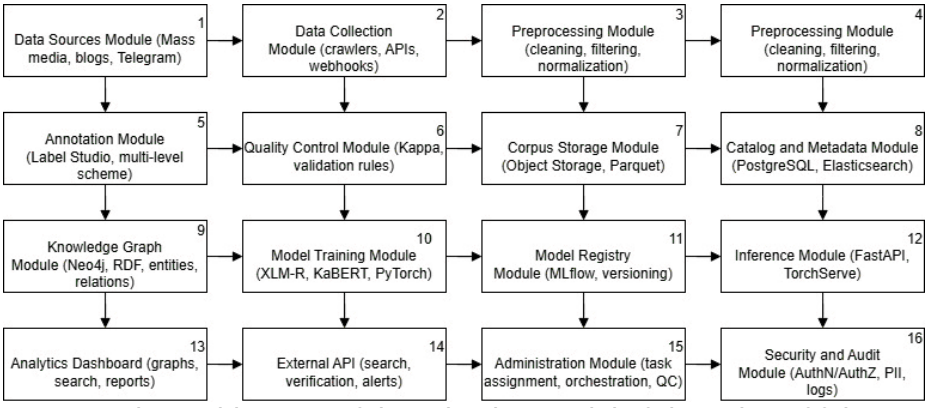


Figure 3 – The architecture of the adaptive model of detection of fake news

The technical outcome of the proposed architecture lies in enabling automatic detection and classification of fake news and subtypes of disinformation within a structured semantic framework. The integration of ontology-driven annotation, knowledge graph construction, and transformer-based modeling supports multilingual analysis (Kazakh and Russian) and provides a foundation for adaptive processing of new information sources. The system is designed to maintain consistency of semantic representation across the full processing cycle.

This ontological scheme defines the structural relationships between key elements of an information message in the context of fake and real news. The purpose of Figure 4 is to formalize the semantic architecture of news annotation within the proposed framework and to illustrate the claim-centered modeling approach adopted in this study. Figure 4 presents a semantic model in

which CLAIM serves as the central entity connected to message type (FAKE_NEWS / REAL_NEWS), source attribution, supporting evidence, communicative intent, target audience, and disinformation techniques. The model reflects the hypothesis that structuring news content around explicitly defined entities and relations enables consistent annotation, supports knowledge graph construction, and provides a formal basis for automated analysis. The following elements and relations are represented in the figure:

FAKE_NEWS is linked to CLAIM through the relation hasClaim, indicating that a false report contains a specific statement.

REAL_NEWS is linked to CLAIM through the relation hasClaim, showing that verified news is also structured around a defined statement.

CLAIM is linked to EVIDENCE via the relation supportedBy, representing supporting (or pseudo-supporting) information.

CLAIM is linked to SOURCE through the relation fromSource, identifying the origin of the statement.

CLAIM is linked to AUTHOR_INTENT via the relation expressesIntent, reflecting the communicative purpose behind the statement.

CLAIM is linked to TARGET_AUDIENCE through the relation targets, defining the intended recipients of the message.

CLAIM is linked to DISINFORMATION_TECHNIQUE via the relation usesTechnique, indicating the manipulation strategy applied in the message.

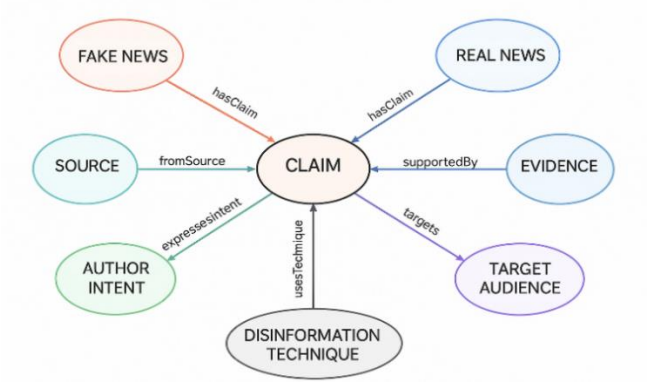


Figure 4 – Semantic model of annotation of news reports

To build a unified body of news messages, a multi-level markup scheme was developed and implemented in Label Studio. This scheme allows the identification of the key structural elements of both fake and real news, i.e., the source, the main statement, time references, and fragments containing supporting or contextualizing information, as well as the recording of manipulative techniques used in the message, the author's intent, and the target audience.

Dataset

The dataset consists of 5,000 news articles, evenly distributed between fake and real news (50% each). The corpus is balanced across two languages: 50% of the texts are in Kazakh and 50% in Russian. This balanced distribution ensures equal representation of classes and languages during training and evaluation. The average text length is approximately 230 characters, reflecting short-format news messages typical for online media and social platforms. The dataset was manually annotated using a multi-level semantic scheme implemented in Label Studio. To assess annotation consistency, inter-annotator agreement was calculated using Cohen's Kappa. The obtained values range from 0.72 to 0.81, indicating substantial agreement across annotation categories. These results confirm the internal consistency of the annotation scheme and its suitability for further model training and evaluation.

Results

The news item (see Fig. 5) about the alleged free public transport in Astana on behalf of City Transportation Systems (CTS) was annotated in Label Studio as FAKE_NEWS with the subtypes FABRICATED_NEWS and MISLEADING_CONTENT; additionally, the disinformation techniques CLICKBAIT and EXAGGERATION were identified, and the author's intent was classified as ATTRACTION_ATTENTION and FINANCIAL_GAIN. The text highlights the entities of SOURCE ("City Transportation Systems (CTS)"), CLAIM (a single fragment that includes the title and the main

promise of free travel for 6 and 12 months), TIME ("on the occasion of the 15th anniversary"), and two fragments of EVIDENCE (a description of the "Arnaya card" and a proposal with a hyperlink). There are semantic relationships between the entities: HAS_CLAIM from SOURCE to CLAIM (source → claim), HAS_TIMESPAN from CLAIM to TIME (claim → TIME), and two SUPPORTED_BY relations reflecting that the main statement relies on pseudo-evidence (claim → EVIDENCE).

This example illustrates how the proposed annotation scheme captures the internal structure of promotional disinformation by linking the main claim with temporal framing and supporting pseudo-evidence. The figure demonstrates the applicability of the ontology-driven model to Kazakh-language content and shows how communicative intent and manipulation techniques are formalized within a unified semantic framework.

Figure 5 – An example of a manual annotation in the Kazakh language indicating all the entities of the scheme

The markup carried out in Label Studio allowed the formalization of the structure of fake news messages, highlighting their key components, such as source, statement, time binding**,** and elements of pseudo-proof, as well as establishing connections between them through semantic relationships. This approach not only systematizes the internal logic of disinformation dissemination but also creates a uniform annotated corpus necessary for the subsequent training and validation of automatic fake news detection models. The use of multi-level markup also enables the analysis of typical manipulation strategies, author intent, and target groups, which may improve the quality of model predictions and contribute to the development of more robust NLP systems against the spread of misinformation.

Annotation Quality Assessment (Cohen's Kappa)

To assess the consistency and reproducibility of the annotation scheme, percentage distributions of annotated entities were compared between two independent annotators, and Cohen's Kappa was calculated. Higher Kappa values for SOURCE (0.81) and CLAIM (0.78) indicate clearer formal definitions and stable identification in news texts. Lower values for categories such as AUTHOR_INTENT and FAKE_NEWS reflect their greater contextual and interpretative complexity (Table 2) [9].

Table 2 – Distribution of annotated entities and inter-annotator agreement (Cohen's Kappa) across main categories

Category	Annotator 1 (%)	Annotator 2 (%)	Cohen's Kappa
FAKE_NEWS	6.5%	6.3%	0.72
REAL_NEWS	7.0%	6.9%	0.74
SOURCE	25.3%	25.2%	0.81
CLAIM	28.5%	28.4%	0.78
EVIDENCE	8.0%	8.1%	0.76
DISINFORMATION_TECHNIQUE	6.5%	6.6%	0.74
AUTHOR_INTENT	6.2%	6.1%	0.73
TARGET_AUDIENCE	6.0%	6.1%	0.72
TIMESTAMP	5.0%	5.2%	0.75

As a result of the annotation quality assessment using Cohen's Kappa coefficient, the proposed annotation scheme demonstrated a substantial level of inter-annotator agreement. The annotation was performed according to predefined guidelines within the Label Studio environment, and agreement was evaluated across the main annotation categories (FAKE_NEWS, REAL_NEWS, SOURCE, CLAIM, EVIDENCE, DISINFORMATION_TECHNIQUE, AUTHOR_INTENT, TARGET_AUDIENCE, TIMESTAMP). The obtained Cohen's Kappa values range from 0.72 to 0.81, corresponding to the level of substantial agreement, while the SOURCE category (0.81) approaches almost perfect agreement. The highest agreement was observed for SOURCE (0.81) and CLAIM (0.78), indicating clearer and more stable interpretation of these categories during annotation. In contrast, slightly lower values for FAKE_NEWS (0.72), AUTHOR_INTENT (0.73), and TARGET_AUDIENCE (0.72) reflect their greater semantic complexity and contextual dependence.

Discussion

The study presents an ontology-based framework for fake news detection that combines semantic markup, knowledge graph construction, and transformer-based modeling. The proposed annotation scheme formalizes key elements of news texts, including claim, source, evidence, time reference, author intent, and target audience. Manual annotation of 5,000 news articles in Kazakh and Russian showed consistent application of the scheme. The Cohen's Kappa values (0.72-0.81) indicate stable inter-annotator agreement across categories. More structurally defined entities, such as SOURCE and CLAIM, demonstrated higher agreement, while context-dependent categories showed slightly lower consistency. The results confirm that the ontology and multi-level markup can be applied in a multilingual setting and can serve as a structured basis for further model training and analysis.

Limitations

The study focuses on the design of the ontology and annotation framework. A comparative experimental evaluation with baseline machine learning or transformer-only models was not conducted. The dataset is limited to 5,000 news texts in Kazakh and Russian, which restricts generalization to other languages and media contexts. The annotation process relies on manual labeling, which requires time and expert involvement. The current implementation considers only textual content and does not include multimodal signals such as images or video. The system has also not been tested in real-time deployment conditions.

Conclusion

The paper proposes an integrated approach to building an intelligent fake news detection system that combines ontological modeling, semantic markup, and modern NLP methods. A bibliometric analysis of Scopus publications (2018-2026) shows a steady rise in interest in fake news detection and semantics, a shift from classical ML methods to semantic, graph-based, and multimodal models, and the strengthening of transformer architectures in news analysis. The developed adaptive system architecture supports the full data processing cycle: collection, preprocessing, annotation, storage, model training, and visual analytics. A key element is the ontological model and multi-level markup scheme implemented in Label Studio, formalizing the internal structure of news messages. The annotation scheme captures key entities (source, claim, EVIDENCE, TIME) and classification categories (DISINFORMATION_TECHNIQUE, REAL_NEWS / FAKE_NEWS, FAKE_SUBTYPE, AUTHOR_INTENT, TARGET_AUDIENCE), as well as their semantic relationships. This ensures high interpretability and forms a corpus suitable for machine learning. Manual annotation results confirmed that the scheme consistently describes both fake and real news. The identification of entities and communicative characteristics provides a holistic representation of news logic. The scheme forms a foundation for developing robust NLP models in multilingual environments. The proposed architecture and ontology enable next-generation intelligent systems for detecting fake news and analyzing disinformation processes. Future work includes experimental validation, markup automation, corpus expansion, multimodal models, and explainable AI to strengthen trust and transparency.

The proposed ontology-driven framework can be applied in media monitoring systems and digital content moderation platforms. The structured representation of claims, sources, and manipulation techniques enables more transparent analysis of news content and may support faster identification of potentially misleading information. The integration of semantic annotation and knowledge graph construction can reduce manual verification workload by structuring information

for automated processing. The approach may also contribute to improving trust in digital platforms by providing interpretable explanations of detection results.

References

1. Credible, unreliable or leaked? Evidence verification for enhanced automated fact-checking / Z. Chrysidis et al // Proceedings of the 3rd ACM International Workshop on Multimedia AI against Disinformation. – 2024. – P. 73-81. https://doi.org/10.1145/3643491.366027.
2. Setty V. Factcheck editor: Multilingual text editor with end-to-end fact-checking / V. Setty // Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval. – 2024. – P. 2744-2748. https://doi.org/10.1145/3626772.3657663.
3. Toward automated factchecking: Developing an annotation schema and benchmark for consistent automated claim detection / L. Konstantinovskiy et al // Digital threats: research and practice. – 2021. – Т. 2, № 2. – P. 1-16. https://doi.org/10.1145/3412869.
4. Indonesian fake news detection using various machine learning technique / L. Triyono et al // International Journal on Informatics Visualization. – 2023. – Т. 7, № 3. – P. 726-732. https://doi.org/10.30630/joiv.7.3.1243.
5. Abumansour A.S. Check-worthy claim detection across topics for automated fact-checking / A.S. Abumansour, A. Zubiaga // PeerJ Computer Science. – 2023. – Т. 9. – P. e1365. https://doi.org/10.7717/peerj-cs.1365.
6. Alghamdi J. A comparative study of machine learning and deep learning techniques for fake news detection / J. Alghamdi, Y. Lin, S. Luo // Information. – 2022. – Т. 13, № 12. – 576. https://doi.org/10.3390/info13120576.
7. Nasir A. Credia: Contextualized retrieval of evidence for open-domain fact verification / A. Nasir, M. Wasim, S. Nasir // Knowledge and Information Systems. – 2025. – P. 1-31. https://doi.org/10.1007/s10115-025-02400-x.
8. Multiknowledge and LLM-inspired heterogeneous graph neural network for fake news detection / B. Xie et al // IEEE Transactions on Computational Social Systems. – 2024. https://doi.org/10.1109/TCSS.2024.3488191.
9. A multi-level annotation model for fake news detection: Implementing Kazakh-Russian corpus via Label Studio / M. Sambetbayeva et al // Big Data and Cognitive Computing. – 2025. – Т. 9, № 8. – H. 215. https://doi.org/10.3390/bdcc9080215.
10. Wang X. Monolingual and Multilingual Misinformation Detection for Low-Resource Languages: A Comprehensive Survey / X. Wang, W. Zhang, S. Rajtmajer // Inf. Process. Manag. – 2025. – № 62. – P. 102789.

Funding: *This research has been funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan Grant AP26195591.*

Ж.Б. Ламашева[1,2], М.А. Самбетбаева[1,2,*], А.Н. Некесова[1,2], Б.Х. Абдығалым[1,2,3], Н. Тасболатұлы[1,3]

[1]«Астана» халықаралық ғылыми кешені,
Қазақстан Республикасы, Астана қ., Қабанбай батыр даңғылы, 8
[2]Л.Н. Гумилев атындағы Еуразия ұлттық университеті,
010008, Қазақстан Республикасы, Астана қ., Сәтпаев к-сі, 2
[3]Астана Халықаралық университеті,
010000, Қазақстан Республикасы, Астана қ., Қабанбай батыр даңғылы, 8
*e-mail: sambetbayeva_ma_1@enu.kz

ОНТОЛОГИЯЛЫҚ МОДЕЛЬ МЕН МЕДИА МӘТІНДЕРІН СЕМАНТИКАЛЫҚ БЕЛГІЛЕУ НЕГІЗІНДЕ ҰЙЫМДАСТЫРЫЛҒАН ЗИЯТКЕРЛІК ЖАСАНДЫ ЖАҢАЛЫҚТАРДЫ АНЫҚТАУ ЖҮЙЕСІ

Бұл мақалада медиа мәтіндерді онтологиялық модельдеу мен семантикалық аннотациялау негізінде жалған жаңалықтарды анықтау тәсілі ұсынылады. Зерттеу аясында 2018-2026 жылдар аралығындағы Scopus дерекқорындағы жарияланымдарға библиометриялық талдау жүргізіліп, жалған жаңалықтарды анықтау және семантикалық тәсілдер саласындағы негізгі үрдістер айқындалды. Нәтижелер дәстүрлі машиналық оқыту әдістерінен трансформерлік, графтық және семантикалық модельдерге көшу үрдісін көрсетеді. Деректерді жинау, алдын ала өңдеу, Label Studio жүйесінде көпдеңгейлі аннотациялау, білім графын құру, модельдерді оқыту, инференс және аналитика кезеңдерін қамтитын бейімделетін жүйе архитектурасы ұсынылды. Жаңалық мәтіндерінің негізгі элементтерін құрылымдайтын формалды онтологиялық модель әзірленді:

тұжырым (claim), дереккөз (source), дәлел (evidence), автор ниеті (author intent), мақсатты аудитория (target audience) және дезинформация әдісі (disinformation technique). Жүйе қазақ және орыс тілдеріндегі көптілді өңдеуді қолдайды. Деректер жиыны 5 000 жаңалық мәтінінен тұрады, олардың жартысы жалған, жартысы шынайы және екі тіл арасында теңгерімді бөлінген. Аннотация сапасы Cohen's Kappa коэффициенті арқылы бағаланды (0,72–0,81), бұл аннотацияның келісімді орындалғанын көрсетеді. Ұсынылған тәсіл көптілді ортада жалған жаңалықтарды автоматты анықтау жүйелерін әрі қарай дамыту мен бағалау үшін құрылымдық негіз ұсынады.

***Түйін сөздер:** фейк жаңалықтар, фейк жаңалықтарды анықтау, онтологиялық модель, семантикалық белгілеу, семантикалық талдау.*

Ж.Б. Ламашева[1,2], М.А. Самбетбаева[1,2,*], А.Н. Некесова[1,2], Б.Х. Абдығалым[1,2,3], Н. Тасболатұлы[1,3]
[1]Международный научный комплекс «Астана»,
Республика Казахстан, г. Астана, проспект Кабанбай Батыра, 8
[2]Евразийский национальный университет имени Л.Н. Гумилёва,
010008, Республика Казахстан, г. Астана, ул. Сатпаева, 2.
[3]Международный университет Астана,
010000, Республика Казахстан, г. Астана, проспект Кабанбай Батыра, 8
*e-mail: sambetbayeva_ma_1@enu.kz

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ВЫЯВЛЕНИЯ ФЕЙКОВЫХ НОВОСТЕЙ НА ОСНОВЕ ОНТОЛОГИЧЕСКОЙ МОДЕЛИ И СЕМАНТИЧЕСКОЙ РАЗМЕТКИ МЕДИА-ТЕКСТОВ

В данной статье представлен подход к выявлению фейковых новостей на основе онтологического моделирования и семантической аннотации медиатекстов. В исследовании проведён библиометрический анализ публикаций базы данных Scopus за 2018-2026 годы с целью выявления тенденций в области обнаружения фейковых новостей и семантических методов анализа. Результаты показывают переход от традиционных методов машинного обучения к трансформерным, графовым и семантическим моделям. Предложена адаптивная архитектура системы, включающая сбор данных, предварительную обработку, многоуровневую разметку в Label Studio, построение графа знаний, обучение моделей, инференс и аналитический модуль. Разработана формальная онтологическая модель, структурирующая ключевые элементы новостных текстов: утверждение (claim), источник (source), доказательство (evidence), намерение автора (author intent), целевая аудитория (target audience) и техника дезинформации (disinformation technique). Система поддерживает многоязычную обработку на казахском и русском языках. Датасет включает 5 000 новостных сообщений, равномерно распределённых между фейковыми и реальными новостями и сбалансированных по языкам. Качество разметки оценивалось с использованием коэффициента Cohen's Kappa (0,72-0,81), что свидетельствует о согласованности аннотации. Предложенный подход создаёт структурную основу для дальнейшей разработки и оценки автоматизированных систем выявления фейковых новостей в многоязычной среде.

***Ключевые слова:** фейковые новости, обнаружение фейков, онтологическая модель, семантическая разметка, семантический анализ.*

Сведения об авторах

Жанар Бейбутовна Ламашева – Phd, старший преподаватель кафедры Информационных систем, Евразийский национальный университет имени Л.Н. Гумилева, старший научный сотрудник Международного научного комплекса «Астана», Астана, Республика Казахстан; e-mail: lamasheva_zhb@enu.kz. ORCID: https://orcid.org/0000-0002-9535-2636.

Мадина Аралбаевна Самбетбаева* – Phd, ассоциированный профессор кафедры информационных систем, Евразийский национальный университет имени Л.Н. Гумилева, ведущий научный сотрудник Международного научного комплекса «Астана», Республика Казахстан; e-mail: sambetbayeva_ma_1@enu.kz. ORCID: https://orcid.org/0000-0001-9358-1614.

Анаргуль Аймуратовна Некесова – магистр технических наук, докторант кафедры информационных систем, Евразийский национальный университет имени Л.Н. Гумилева, младший научный сотрудник Международного научного комплекса «Астана», Республика Казахстан; e-mail: aimurat_anara@mail.ru. ORCID: https://orcid.org/0009-0001-7642-0648.

Баянғали Хайерберліұлы Абдығалым – магистр технических наук, докторант кафедры информационных систем, Евразийский национальный университет имени Л.Н. Гумилева, Инженер программист Международного научного комплекса «Астана», Республика Казахстан; e-mail: bayangali.abd@gmail.com. ORCID: https://orcid.org/0009-0001-8872-7428.

Нурболат Тасболатулы – PhD, ассоциированный профессор Высшей школы информационных технологий и инженерии Международного университета Астана, ведущий научный сотрудник

Международного научного комплекса «Астана», Республика Казахстан; e-mail: nurbolat.tasbolatuly@aiu.edu.kz. ORCID: https://orcid.org/0000-0002-0511-7000.

Авторлар туралы мәліметтер

Жанар Бейбутовна Ламашева – Phd, Л.Н. Гумилев атындағы Еуразия ұлттық университетінің «Ақпараттық жүйелер» кафедрасының аға оқытушысы, «Астана» Халықаралық ғылыми кешенінің аға ғылыми қызметкері, Астана, Қазақстан Республикасы; e-mail: lamasheva_zhb@enu.kz. ORCID: https://orcid.org/0000-0002-9535-2636.

Мадина Аралбайқызы Самбетбаева* – PhD, Ақпараттық жүйелер кафедрасының қауымдастырылған профессоры, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, «Астана» халықаралық ғылыми кешенінің жетекші ғылыми қызметкері, Қазақстан Республикасы; e-mail: sambetbayeva_ma_1@enu.kz. ORCID: https://orcid.org/0000-0001-9358-1614.

Анаргуль Аймуратовна Некесова – техника ғылымдарының магистрі, ақпараттық жүйелер кафедрасының докторанты, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, «Астана» Халықаралық ғылыми кешенінің кіші ғылыми қызметкері, Қазақстан Республикасы; e-mail: aimurat_anara@mail.ru. ORCID: https://orcid.org/0009-0001-7642-0648.

Баянғали Хайерберліұлы Абдығалым – техникалық ғылымдар магистрі, ақпараттық жүйелер кафедрасының докторанты, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, «Астана» халықаралық ғылыми кешенінің бағдарламашы инженері, Қазақстан Республикасы; e-mail: bayangali.abd@gmail.com. ORCID: https://orcid.org/0009-0001-8872-7428.

Нұрболат Тасболатұлы – PhD, Аста халықаралық университетінің Ақпараттық технологиялар және инженерия жоғары мектебінің қауымдастырылған профессоры, «Астана» халықаралық ғылыми кешенінің жетекші ғылыми қызметкері, Қазақстан Республикасы; e-mail: nurbolat.tasbolatuly@aiu.edu.kz. ORCID: https://orcid.org/0000-0002-0511-7000.

Information about the authors

Zhanar Beibutovna Lamasheva – PhD, Senior lecturer at the department of Information Systems, L.N. Gumilyov Eurasian National University, senior Researcher at the International Scientific Complex «Astana» Astana, Kazakhstan; e-mail: lamasheva_zhb@enu.kz. ORCID: https://orcid.org/0000-0002-9535-2636.

Madina Aralbayevna Sambetbaeva* – PhD, associate professor of the Department of Information Systems, L.N. Gumilyov Eurasian National University, leading researcher at International Science Complex Astana, Republic of Kazakhstan; e-mail: sambetbayeva_ma_1@enu.kz. ORCID: https://orcid.org/0000-0001-9358-1614.

Anargul Nekessova – Master of Technical Sciences, Phd student of the Department of Information Systems at L.N. Gumilyov Eurasian National University, Junior Researcher at the International Scientific Complex «Astana», Republic of Kazakhstan; e-mail: aimurat_anara@mail.ru. ORCID: https://orcid.org/0009-0001-7642-0648.

Bayangali Khayerberliuly Abdygalym – master of technical sciences, Phd student of the Department of Information Systems, L.N. Gumilyov Eurasian National University, Software engineer at International Science Complex Astana, Republic of Kazakhstan; e-mail: bayangali.abd@gmail.com. ORCID: https://orcid.org/0009-0001-8872-7428.

Nurbolat Tasbolatuly – PhD, Associate Professor at the School of Information Technology and Engineering, Astana International University, Leading Researcher at the Astana International Scientific Complex, Republic of Kazakhstan; e-mail: nurbolat.tasbolatuly@aiu.edu.kz. ORCID: https://orcid.org/0000-0002-0511-7000.

Received 05.01.2026
Revised 23.02.2026
Accepted 11.03.2026