Information System for the Fact-checker Support

Yurii Baryshev [1], Leonid Kupershtein [1], Vladyslav Maidanovych [1], Olesia Voitovych [1] and Serhii Prokopenko [2]

[1] *Vinnytsia National Technical University, Khmelnytske shoes str., 95, Vinnytsia, 21021, Ukraine*
[2] *Semen Kuznets Kharkiv National University of Economics, Nauky Ave., 9-A, Kharkiv, 61166, Ukraine*

Abstract

The research results of fact-checking automation are presented. The main types of fact-checking were described and the methods and techniques used by fact-checkers were analyzed. Particular attention was paid to the development of the system architecture, which consists of various modules, including a visualization module, database search, internet search, analysis and verification of news on websites and search in telegram channels. Algorithms were developed to work with these modules, allowing for efficient fact-checking. The software architecture was designed and experimental application was developed.

Keywords [1]

Fact-checking, automation, fake news, information attack, cybersecurity.

1. Introduction

In a world where inaccurate news can spread quickly and have harmful impact, the process of fact-checking is becoming vital and increasingly topical. It is a key tool in the fight against so-called fake news and disinformation. Fact-checking is not only a way to identify false statements but also a means of restoring the trust to the media. However, the process of fact-checking can be labor-intensive and time-consuming, especially when dealing with large amounts of information. Fact-checkers often have to spend hours, and sometimes days, searching for evidence to confirm or refute certain facts. In addition, the verification process itself can be challenging, as it requires not only extensive knowledge but also the ability to think critically and analyze information. In this context, the development of an automated fact-checking tool can be not only a significant contribution to the field but also a real asset for journalists, fact-checkers, and anyone working in the informational and media sectors. It is believed that these kind of tools can greatly simplify the fact-checking process, reduce the time required for it, and increase the efficiency and accuracy of the process.

The goal of our work is to increase the efficiency of the fact-checking process by automating the search, collection and analysis of information using modern information intelligent technologies.

In the modern world, inundated with vast amounts of information, it is imperative that we have confidence in the accuracy of the information we consume. In this context, this work is highly relevant, as it aims to develop effective tools to assist in verifying the authenticity of information. Future research in this field could greatly enhance capacity to identify and counteract fake news and disinformation.

2. Analysis of Fact-Checking

Fact-checking is a modern tool for journalistic investigations with a unique concept and technical features aimed at exposing populism, manipulations, and false information in the rhetoric of speakers at various levels and in media publications [1]. This procedure is crucial in the media industry, where

Information Technology and Implementation (IT&I-2023), November 20-21, 2023, Kyiv, Ukraine
EMAIL: kupershtein.lm@gmail.com (L.Kupershtein); vladmaidanovich1@gmail.com (V.Maidanovych); yuriy.baryshev@vntu.edu.ua (Y.Baryshev); voytovych.op@gmail.com (O.Voitovych); prokopenko.serhii@gmail.com (S. Prokopenko). ORCID: 0000-0003-4712-3916 (L.Kupershtein); 0009-0008-1983-479X (V.Maidanovych); 0000-0001-8964-7000 (O.Voitovych); 0000-0003-0718-4917 (S. Prokopenko); 0000-0001-8324-8869 (Y.Baryshev)

journalists, editors, and other professionals are required to effectively verify information before disclosing it. Fact-checking encompasses the verification of aspects such as factual data, quotes, sources of information, and other content.

It is important to note that fact-checking is not only about identifying false information, but also about the ability to recognize and identify subjectivity and hidden aspects, as well as to assess the quality of information in the context. Professional journalism is distinguished by the usage of reliable data and facts verified by all possible means [2].

This process also plays a significant role in cybersecurity. Disinformation and information manipulation are increasingly prevalent in the modern world, posing threats in various contexts, including wars and conflicts. Military conflicts involve not only physical combat but also the usage of information warfare, such as disinformation and manipulation, to achieve objectives and influence public opinion [3]. Fact-checking, as an investigative trend, swiftly entered Ukrainian media landscape. This was primarily in response to the wave of anti-Ukrainian propaganda that rapidly flooded information channels with fake and blatantly false reports [1].

Ukraine serves as an example of information warfare as an integral component of a military conflict. Such disinformation campaigns and information manipulation aim to sow discord, foster distrust, and create social tensions. Their consequences can be severe, impacting political stability, public safety, and the economic situation in the country. Furthermore, these information attacks can have an international impact, altering perceptions and support from allied countries and influencing political relations. To ensure cybersecurity and defend against such information threats, it is crucial to actively implement fact-checking methods and counter the disinformation. The development and utilization of tools for detecting fake news, promoting critical thinking and information literacy, and facilitating access to reliable information can mitigate the impact of disinformation campaigns and enhance cybersecurity in the context of military conflicts and other challenges in the modern world.

For instance, in the initial days of the full-scale invasion of Ukraine by the russian federation, the Gvara Media team launched the fact-checking telegram-bot *"PEREVIRKA"*. This bot assists in distinguishing fake news from real news and recognizing manipulation and propaganda. Upon receiving news submissions, the information is forwarded to project volunteers who manually verify its truthfulness and accuracy. In just the first month of the invasion, they processed approximately 30000 requests [4].

In Ukraine, the concept of fact-checking is currently represented by several resources:

- VoxUkraine: This is a fact-checking project of the independent analytical platform Vox Ukraine. The team exposes lies, manipulations, and russian propaganda, both in Ukraine and abroad. Since 2018, they have been signatories to the Code of Ethics of the Poynter Institute's International Network of Fact-checkers, and since 2020 [5].
- Word and Deed: This resource checks the election promises of politicians [6].
- Stopfake.org: Their primary activity is aimed at combating anti-Ukrainian propaganda, statements, and facts aimed at discrediting Ukraine [7].
- Gvara Media: An independent online publication focused on social change [4].

For example of what fake news can do is to look at a recent event. Realistic images of an explosion near the Pentagon created by artificial intelligence that went viral on Twitter caused the stock market to briefly plummet. Tweets with images purportedly depicting an explosion near the Pentagon building in Arlington, Virginia, were posted by many verified Twitter accounts, including a russian state media outlet with millions of followers and a verified account claiming to be the Bloomberg news agency. Although at first glance the viral photos appeared to be genuine, they are full of hints that they were created with the help of artificial intelligence, which proves that it is all a hoax [8].

A similar event happened last year. On Nov. 10, a verified Twitter account posed as Eli Lilly wrote in a viral tweet, "We are excited to announce insulin is free now." By the next day, the drugmakers' shares plunged by about $22 billion [9].

3. Methods and Techniques of Fact-Checking

Fact-checking, as a process of verifying the truthfulness of information, includes a number of methods and techniques that help to carry out this work effectively and accurately.

These methods include:

- Checking the reliability of sources: A fact-checker determines the reliability, reputation and competence of the source of information, as well as the ability to confirm or refute the information presented.
- Comparison of data: Fact-checkers analyze information from different sources and compare it with known facts or data to detect discrepancies or inaccuracies.
- Checking the context: Fact-checkers analyze the circumstances in which the information was provided and check for distortions or deletions that could change its meaning.
- Involving experts: Fact-checkers can call on experts in the relevant field to verify specific facts or data, using their knowledge and experience.
- Objectivity of experts: Experts involved in fact-checking should be objective and free from conflicts of interest. Their experience and reputation, as well as their qualifications, should be thoroughly checked.
- Use of verified data: It is important to use reliable sources of information, such as official statistics and scientific studies, to compare and corroborate information.
- Technologies for fact-checking: Modern technologies, such as artificial intelligence, machine learning and data analytics, can automate the fact-checking process, increasing its speed and accuracy. They help to identify contradictory statements and manipulation of information.

To check the authenticity of content on the Internet special software, applications, and browser plug-ins (Trooclick, Genius, factcheck.org, and others) are used [10]. Such services are able to analyze facts and identify not only fake news but also manipulative content, identify inaccuracies or biases in the presentation of the material. They scan selected text and check it using official statistics, media reports, their own database, etc. Some of the services are able to analyze headlines, text complexity, popularity, and the structure of the site's code.

Additionally, there are also services that can aid verification process of multimedia content such as photos and videos.

Here are some of the services for photos verification:

- Google.Photo allows to upload a photo and find similar or identical images. Results can be sorted by size, upload date, and image format [11].
- Google.Lens works in a similar way to Google.Photo, but allows to search for individual parts of an image, such as buildings in a photo [12].
- TinEye - this service allows to find different versions of images, including modified versions, such as cropped, retouched or with added elements. Results can be filtered by relevance, size or differences from the original [13].
- EXIF Data/Jeffrey's Image Metadata Viewer - these tools allow to get technical information about a photo, such as the date it was taken, the location, and the device it was taken with. But, this data may not be available for some photos that have been uploaded and re-uploaded multiple times [14].
- FotoForensics - this service analyses images using the ELA (error level analysis) method. It compares the image with its compressed version. It helps to identify which part of the image has been altered, whether it is darker or brighter compared to a similar area that has remained unchanged [15].

These tools help fact-checkers detect photographic manipulation and verify the authenticity of images.

Fact-checking technologies are powerful tools, but they cannot replace the role of a fact-checker. The manual element remains an integral part of the process, as context, cultural differences and subjectivity of information need to be taken into consideration.

The efficient usage of technology, combined with expert analysis, can lead to more objective and reliable results in fact-checking.

4. Problem Statement

Fact-checking is a complex and intensive process that requires a significant amount of time and effort to verify the accuracy of facts, news, and information disseminated in the media and online sources. Given the vast volume of information available, manually verifying all data is highly resource-intensive in terms of time and resources.

That's why the automation of fact-checking processes is becoming increasingly important. The usage of automated algorithms and software solutions allows for the efficient analysis of large amounts of data, rapid verification of facts, and the identification of potential attempts at disinformation and information manipulation. This, in turn, enhances the speed and accuracy of the fact-checking process, enabling a quicker response to the spread of false information and preventing its adverse consequences.

The primary objective of our work is to develop an information system based on software tool to support the work of fact-checkers, with the goal of providing a convenient and effective means of detecting fake news.

Functionality Requirements:

- Identification of fake news: The tool should be capable of analyzing news content to determine its credibility. This may involve analyzing the text, source, author, and other relevant attributes of the news.
- Search for similar news: The tool should have the capability to search for similar news stories within a database or Internet. This feature will assist fact-checkers in identifying instances of news republishing or the spread of fake news.
- Parsing comments: The tool should be able to collect comments on news articles from various sources, social media, and other platforms. This will enable fact-checkers to analyze public reactions and opinions regarding a specific news item.
- Analysis of publication dates and authorship: The tool should be able to analyze the publication date and author information. This functionality will help fact-checkers assess the relevance of the news and evaluate the credibility of the author.

Interface Requirements:

- User-Friendly Interface: The tool should feature an intuitive and user-friendly interface that enables fact-checkers to use it effortlessly and access the necessary information.
- Visualization of Results: The tool should be capable of visually presenting the results of news analysis. This visual representation will allow fact-checkers to quickly assess the credibility and dissemination of fake news.

Performance Requirements**:**

- Processing Speed: The tool should have the capability to process a substantial volume of news rapidly, enabling fact-checkers to promptly identify and respond to fake news.
- Efficient Resource Utilization: The tool should optimize the utilization of resources, including computing power and memory, to ensure the efficiency of the fact-checkers work.

The tool should include features such as searching for similar news within the database, searching for news on Telegram channels, parsing comments, identifying fake news, analyzing publication dates and authorship. Additionally, the tool should offer a user-friendly interface, the ability to visualize results, and efficient performance.

So, we have discussed the concept of fact-checking and its significance in the modern information environment. We have also described the primary types of fact-checking. Following an analysis of the methods and techniques employed by fact-checkers, we have identified the key steps in the fact-checking process, including fact verification, source and citation checks, context analysis, and the utilization of specialized tools and data.

Comprehending these methods and techniques is essential for the effective execution of fact-checking tasks. The study has pinpointed the principal challenges associated with fact-checking, such as politicization and the impact of social media.

The development of such a tool will help to increase the efficiency and accuracy of information verification, as well as help to provide reliable and accurate information to a wide audience by automating the fact-checking process.

5. Information system architecture

The information system architecture includes several modules. It is shown in Figure 1. It consists of six modules: visualization module, database-search module, telegram-search module, network-search module, analysis and verification module and data-save module. System has its own database also.

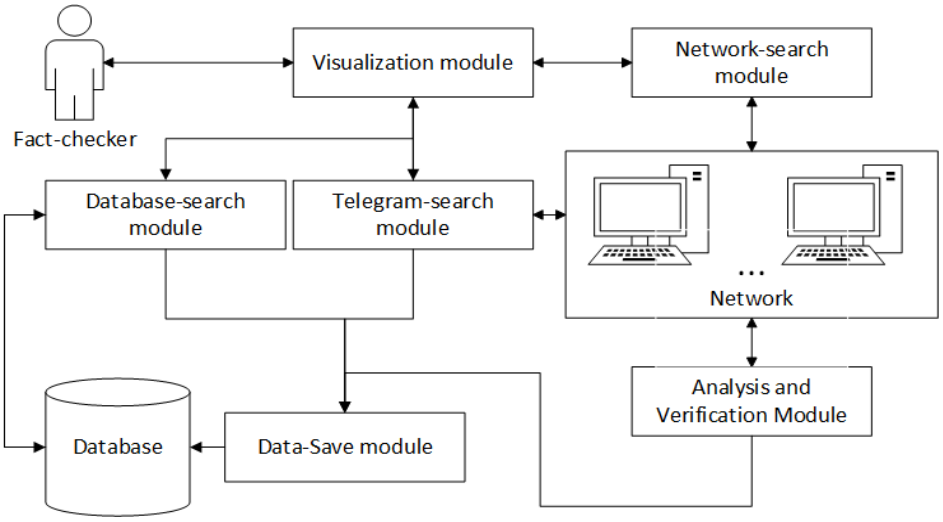

Figure 1: Information system architecture

Let's consider each module separately. The visualization module is responsible for launching the system and presenting the graphical interface to an user. Upon launching the application, the module initializes graphical components such as windows, icons, images, and control buttons for user interaction with the tool. The module handles the display of windows in response to user actions. It also establishes the necessary connections and displays the system's results. After data entry, the module passes this information to the subsequent processing stages.

The database search module searches and compares the entered data with the existing database of fake and true news. It utilizes algorithms and methods to identify matches and determine the credibility of the news. If the news is found in the database, the component returns the corresponding verification result. However, frequently incoming news may not be unique, as different information channels, in various formats, cover certain events even though they pertain to the same event. Consequently, the same facts can be processed multiple times. Machine-based comparison of the received message with ones stored in the database enables the automatic processing of similar requests. The task of comparing and determining text similarity is relatively straightforward. Currently, there are many tools available for implementing this.

One of the most popular methods of converting text into a numerical vector is to use vector embedding. A vector embedding is a numerical representation of a word or phrase that preserves the semantic relationships between words [16]. For example, words with similar meanings will have similar vector representations.

One of the simplest approaches is the bag-of-words approach. In this approach, the text is considered as a set of individual words, and a vector is created that shows which words are present in the text and how often they occur. Each word is considered a separate feature, and its number in the text is the value of this feature [17].

Another approach is TF-IDF (Term Frequency-Inverse Document Frequency), which is a statistical method for assessing the importance of a word in a document by taking into account its frequency in the document and weighing the frequency of the word in other documents. The TF-IDF is calculated for each word in a corpus of documents and gives a numerical value that reflects its relevance to each individual document.

The TF-IDF is often used to represent documents in a collection as numerical vectors those reflect the importance of each word in a set of words (the number of words in the set determines the dimension of the vector) in each document. Such model is called a vector model and makes it possible to compare texts by comparing their respective vectors using given metric (Euclidean distance, cosine measure, Manhattan distance, Chebyshev distance, etc.), i.e. to perform a cluster analysis [18].

Next is the Word2Vec model, which learns to predict words based on their context. It creates vector representations of words in which similar words have similar vectors [19].

Another one is the GloVe model (Global Vectors for Word Representation), which combines statistics of word occurrences in texts with matrix decomposition [19].

The FastText model uses the Word2Vec approach, but treats words as a set of character substrings. It allows to obtain vector representations even for words that were not included in the training corpus [19]. There is also the Universal Sentence Encoder Multilingual model from Google. It is a powerful tool for converting text sentences into vector representations (embeddings). The main goal of the model is to preserve the semantic relations between sentences in the vector space [20].

For our work, we have chosen the Universal Sentence Encoder Multilingual model, which is based on transformers. It provides a convenient and powerful way to convert text into numerical vectors. Its multilingual support allows to work with text in different languages, which is important for analyzing news and content from different sources. In addition, the model has good performance and high accuracy, which makes it an ideal choice for our fact-checking and news content analysis tasks.

The primary steps for implementing machine fuzzy text matching are:

- Cleaning of stop words: prepositions, conjunctions, pronouns.
- Cleaning of invalid data: links, etc.
- Conversion to a numeric vector format: at this stage, a trained neural network model processes the text and transforms it into a numeric vector array, for example:

““Окупанти застосували хімічну зброю на Херсонщині” ->
[[0.00342214 -0.00027666 -0.02173961 ... -0.01806361 0.02199496 0.04954417]
[-0.06651231 -0.08472251 0.0423876 ... -0.02535929 -0.01699802 0.05733533]
[-0.04076068 -0.093428 -0.02202856 ... -0.0024216 -0.00762561 0.06122636]
...
[0.04959109 -0.0801333 -0.06416482 ... -0.02055199 0.0410036 -0.04278645]
[-0.03331999 -0.09735778 0.00543897 ... 0.03228471 0.01209306 -0.04241692]
[-0.0637934 -0.08842427 -0.02345617 ... 0.04337208 -0.03944801 0.04867269]]

Our array has 512 elements. Each element is represented by a floating point number.

Our program utilizes a database that is sourced from a dump of the "ПЕРЕВІРКА" Telegram bot. This database serves as the foundation for data retrieval and operations within our application.

The resulting vector arrays can be easily compared using the cosine similarity measure and the percentage of similarity between two texts can be determined.

The database structure contains the News table with the following fields:

- image: A field containing a link to the image associated with the news item. This allows you to store one or more images related to a particular news item.
- video: A field containing a link to a video associated with the news item. It allows you to store one or more videos related to a particular news item.
- fakeStatus: A Boolean field that indicates whether the news is fake. A value of "true" indicates a true news item, and a value of "false" indicates a fake news item.
- createdAt: A field containing the date and time the news was created. This allows you to track when a particular news item was added to the database.
- text: A text field that contains the text of the news item. It stores the main content of the news item, which will be used for further analysis and verification.
- id: A unique identifier for the news item. This field allows you to uniquely identify and access a specific news item in the database.

This database structure allows to store information about news, including images, videos, fake news status, date of creation and the news text itself. A unique identifier helps to unify and uniquely identify each news item in the system.

The Telegram search module allows to automatically find news in the channels and groups of this platform that have similar content to the entered news text. This module provides an opportunity to effectively analyze and monitor news content, receive up-to-date information from a large number of sources and identify news that have a similar context or topic. Usage of the Telegram search module allows to quickly and conveniently review news from different sources in one place, as well as analyze their similarities and identify related topics or events. This will help users to track up-to-date information, conduct research, and draw informed conclusions based on a wide range of sources.

The Internet search module allows users to search for news that was not found in the database. It queries a variety of information sources, including news websites, social media and other sources, to find related information about a news item. The module displays a list of URLs of web pages where the

news is mentioned and the date of its publication. It also provides special buttons next to each result, which are needed for further analysis and verification of the news.

The analysis and verification module uses special algorithms and methods to assess the reliability of the news. It performs text processing, detection of suspicious patterns, context analysis and other techniques to determine whether a news item is fake or not. The verification result is presented in an appropriate way, allowing the user to understand how reliable the news is.

The general algorithm for checking news for authenticity can be summarized as follows:

Step 1. Input data: Receiving the news for analysis and verification.

Step 2. Pre-processing:

- Parsing data from the web page
- Identification of the author.
- Parsing the publication date of the news.
- If supported by the website, parsing of comments.

Step 3. Context analysis:

- Analyzing the context of the news, including keywords, topics and other parameters.
- Identify suspicious patterns or incorrect information.
- Check whether links to other sources are present. Each news item should be linked to a source.

Step 4. Language and style analysis:

- Determining the language of the news text and checking compliance with language rules.
- Take into account the style and phrases used that may indicate fake news.

Step 5. Checking the source:

- Verifying the news from other trusted sources.
- Checking the credibility and reliability of the news source.
- Checking whether the source is satirical and the news itself is a joke.

Step 6. Using of external sources:

- Querying external sources or databases for more information about the news.
- Compare the results and confirm or deny the facts of the news.

Step 7. Yielding a result:

- Based on the collected information and analysis, determining whether the news is fake.
- Displaying the result, which indicates whether the news is authentic or fake.

Step 8. Output data: Receiving the result of the fake news checking.

This algorithm describes the main steps and procedures performed in the News Analysis and Verification module. The implementation of specific algorithms and methods may vary depending on the needs. In our software, the algorithm of the analysis and verification module is based on rules and weights. First of all, we parse the elements of the web page where the news is placed, such as the title, author, date of publication, and author. Next, we use rules to check the authenticity of the news. For example, we check whether the source of the news is unreliable: if true, it is set to 1, otherwise to 0. We also check whether the news text is typed: if there are typos, it is set to 1, otherwise to 0. Additionally, we check for manipulative words in the news text: if they are found, it is set to 1, otherwise to 0. Also, if there is no author of the news, it is set to 1, otherwise to 0. We use weights for each rule. In our case, each rule has a weight of 0.25. Finally, we calculate the newsworthiness score by calculating the weighted sum of the rule values.

The value is calculated using the formula:

$$fake\_news\_score = (source\ weight \cdot source\ rule + text\ weight \cdot text\ rule + manipulation\ weight \cdot \cdot manipulation\ rule + author\ weight \cdot author\ rule).$$

The data storage module is responsible for saving news and verification results in the database. It provides efficient storage and management of the collected data, ensuring its security and availability. This component may use a database or other storage mechanisms to ensure efficient access to data and to ensure data integrity and confidentiality. Each of these modules interacts with the others, transmitting and receiving the necessary data for the system to operate efficiently. This interaction ensures the integration of functionality and improves the system's ability to verify the accuracy of news. We have utilized the capabilities of the Python programming language and the JetBrains PyCharm development

environment to construct our system. While the program's user interface was created using the customtkinter library, we have also leveraged a host of other powerful tools to extend its functionality.

For the transformation of embeddings, we have turned to the renowned TensorFlow library, a stalwart in the field of machine learning and deep neural networks. Its extensive capabilities are instrumental in converting textual data into numerical vectors, thereby enabling effective comparison and similarity measurement. Additionally, for calculating similarity metrics, we have incorporated the sklearn library. This dynamic library equips our system with the ability to assess the similarity between text vectors, employing various metrics like cosine similarity, Euclidean distance, and more.

In the realm of communication with Telegram, we have utilized Telethon, a Python library that facilitates interaction with the Telegram API. This integration enables our system to access and search Telegram channels, expanding the horizons of information retrieval. When parsing news from Internet content, our system use the Google API Client library. This tool enhances the capacity of our system to conduct web searches and retrieve relevant data. These integrated tools and libraries fortify the core of our system, allowing it to function as a comprehensive fact-checking and news analysis platform.

6. Experiments

To launch the program, we need to download the software and run it. The user opens the program window on the "Check by database" tab. It is shown in Figure 2.

Figure 2: Database check window

On the side of the program is a window with controls:

- "Database check" screen;
- "Telegram check" screen;
- "Checking the network";
- element for changing the application theme;
- element for changing the interface scale.

Let’s test the application's intaraction with databases. We already have a ready-made database of fake news. Let's create an array of NumPy vectors for it. To do this, one is to press the "Select File" button and then "Convert". After that, the application asks where to save the file. Once the file is successfully saved, one can select it using the Select File button to continue working. Select database and an array of vectors. Enter the news to be checked in the text box. Enter the number of most similar results those user wants to display. Press the "Search" button. The result is shown in Figure 3.

Let’s test the developed application with Telegram. We have a ready-made database of links to Telegram channels. Since the algorithm works using the Telegram API, to get started, one needs to get the API ID and API HASH values and enter them into the special fields of application. This procedure needs to be done only once, and these values will be saved in the future.

After entering the API ID and API HASH values, one needs to enter the news we need, select file with the channel database, and enter the date for which application should search for the news. The result is shown in Figure 4. After pressing the "News" button, user will see all similar posts from all channels. As a result, program outputs the link to the channel, news text, link to the post and date of publication. The result is shown in Figure 5. Let's test the parsing of comments. Since we have searched for news by channel, and the results contain links to posts, let's take one of them.

To start parsing, one needs to paste the link into a special field, and if necessary, select black or white list filtering, and click the "Comments" button. The result of parsing can be seen in Figure 6. Let's try to perform the same action, but with the whitelist filter enabled. The result is shown in Figure 7.

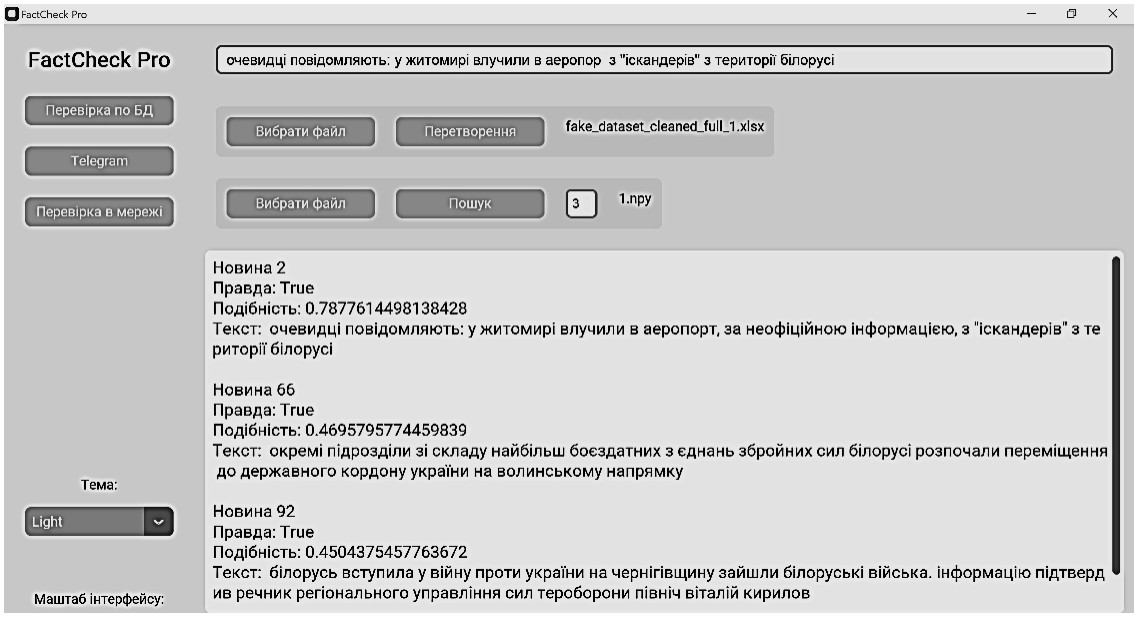

Figure 3: Database search result

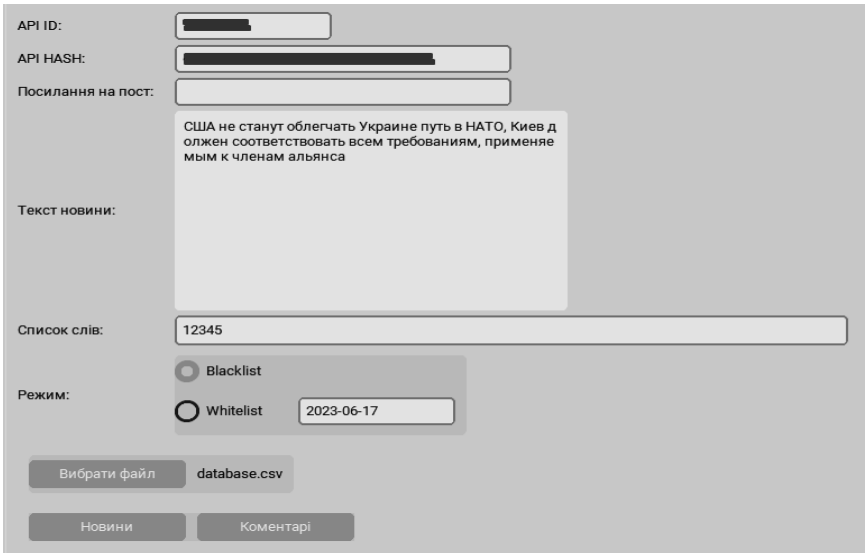

Figure 4: Search interface for Telegram channels

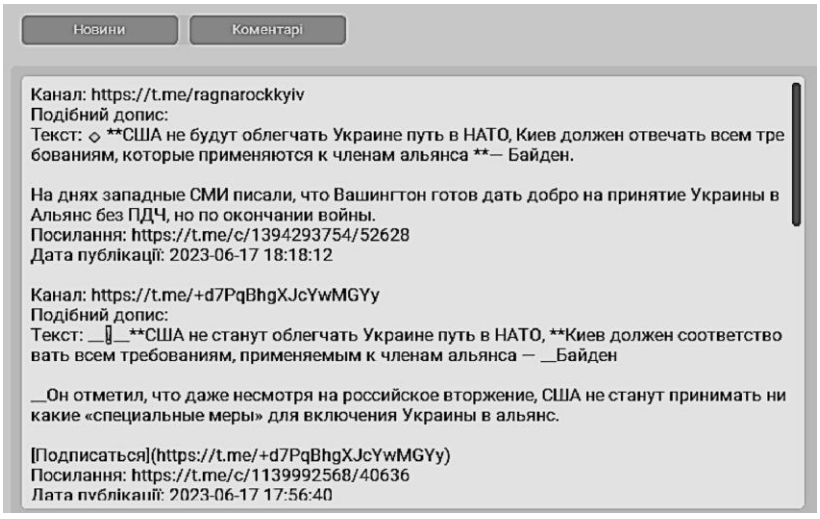

Figure 5: Search result in Telegram channels

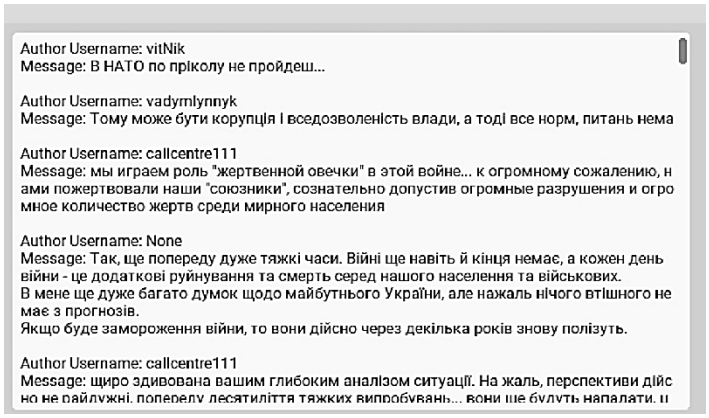

Figure 6: The result of parsing comments without a filter

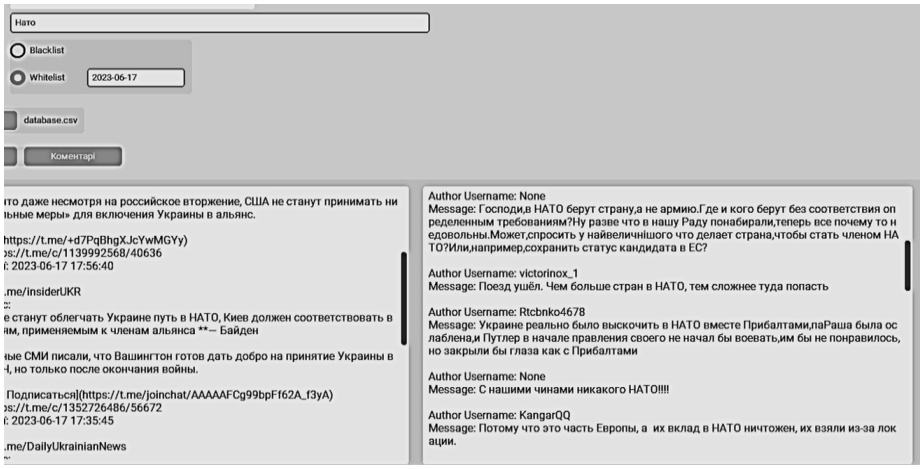

Figure 7: Parsing result with whitelist filter

Figure 8: News search result in the network

Let's test the search for news on the web. To do this, one is to enter the news in the search field and press the "Search" button. The search results will be displayed in the text box. The result is shown in Figure 8. As it can be seen from fig. 8 the results include the headline of the news and the link to it. Next, we can test the news verification on the web page. Since the search returned certain results, we

can take the link from there. Let’s paste the selected link into a special field and press the "Check. The verification process yield the result shown in Figure 9. As we can see, the Fake News Score is 0.0, which means the news is most likely true. Let's try checking out https://panorama.pub, a site that publishes only joke and fake news The result is shown in Figure 10.

As we can see, Fake News Score: 1. This means that the news is fake.

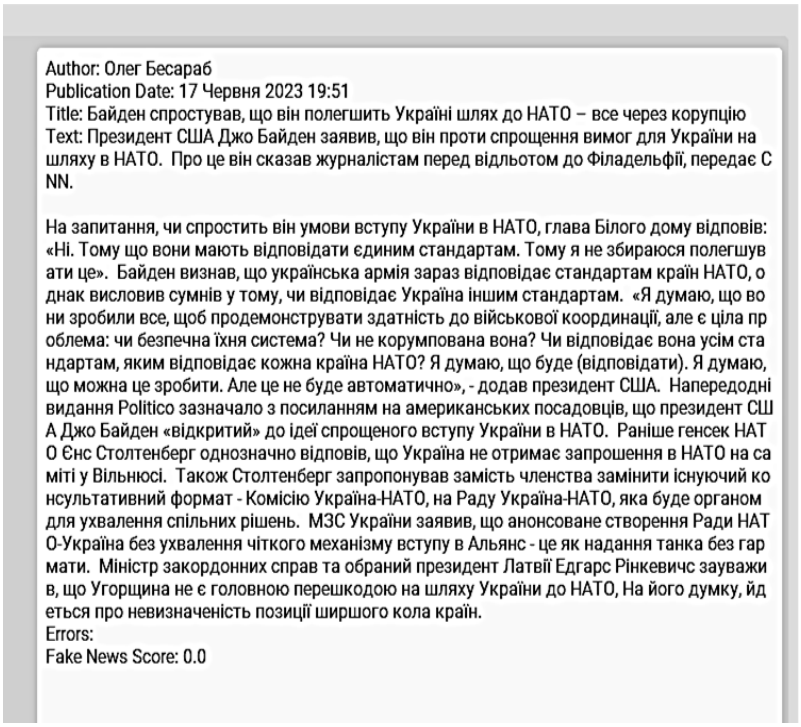

Figure 9: The result of checking the news on the web page

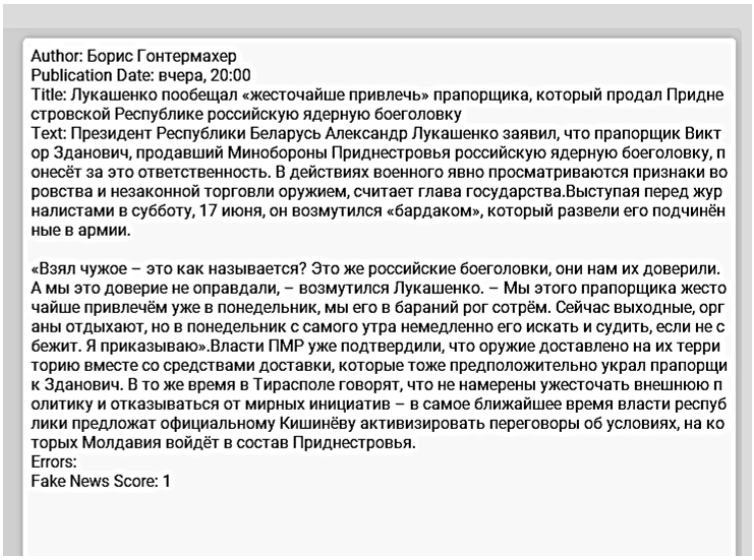

Figure 10: The result of checking a joke site

7. Conclusion and future works

In this article, we have presented a comprehensive system for fact-checking and analyzing news content, which utilizes the power of machine learning, natural language processing, and data integration. Our system is designed to streamline the fact-checking process and enhance its efficiency, providing users with a valuable tool for verifying the credibility of news items. We have integrated various modules, each serving a distinct purpose, to create a unified platform for news analysis and

verification. The effectiveness of our system is based on its ability to automate and expedite tasks that typically demand hours of manual fact-checking. A human fact-checker may spend considerable time searching both social media and Internet publications, trusted information sources, analyzing context, and scrutinizing language and style for signs of misinformation or manipulation. With our system, these tasks are performed in seconds, offering a rapid response and flexibility in customization. Users can customize the system to their specific needs and preferences. One of the key strengths of our system is its integration with machine learning techniques, including state-of-the-art NLP transformers. We are currently exploring the development of a module for automatic text classification using these advanced technologies. This addition would further enhance our system's ability to discern the authenticity of news by leveraging the latest advancements in machine learning.

8. References

[1] Gorokhovsky O. M. Factchek as a trend of investigations: opportunities and prospects: a practical guide. Dnipro: LIRA, 2017. 133 p.

[2] Challenges of the digital space: social networks, verification and fact-checking. URL: https://www.aup.com.ua/vikliki-cifrovogo-prostoru-socialn.

[3] Zhadko V. O., Kharytonenko O. I., Poltavets Y. S. Hybrid war and journalism. Problems of information security. Kyiv: Drahomanov National Pedagogical University, 2018. 356 p.

[4] Krainikova, T., & Prokopenko, S. (2023). Waves of disinformation in the hybrid Russian-Ukrainian war. Current Issues of Mass Communication, 33, 12–25. https://doi.org/10.17721/CIMC.2023.33.12-25

[5] VoxCheck. URL : https://voxcheck.voxukraine.org.

[6] Word and deed. URL : https://www.slovoidilo.ua.

[7] Stopfake.org. URL : https://www.stopfake.org/ru/glavnaya-2.

[8] Pentagon explosion hoax goes viral. URL: https://www.bleepingcomputer.com/news/security/pentagon-explosion-hoax-goes-viral-after-verified-twitter-accounts-push.

[9] Eli Lilly lost billions after fake Twitter account promotes free insulin. URL: https://www.beckershospitalreview.com/pharmacy/eli-lilly-lost-billions-after-fake-twitter-account-promotes-free-insulin.html.

[10] Fake control: how to recognize fakes and not get caught in a lie. website. URL: https://vsim.ua/IT-novyny/feyk-kontrol-yak-rozpiznati-feyki-i-ne-potrapiti-na-gachok-brehni-novi-11170324.html.

[11] Google Photo. URL: https://images.google.com.

[12] Google.Lens. URL: https://lens.google/intl/uk.

[13] TinEye. URL: https://tineye.com.

[14] Exif Data Viever. URL: https://exifdata.com.

[15] Fotoforensics. URL: https://fotoforensics.com.

[16] Reisinger, Joseph; Mooney, Raymond J. (2010). Multi-Prototype Vector-Space Models of Word Meaning. Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Los Angeles, California: Association for Computational Linguistics. pp. 109–117.

[17] E. Lee, F. Rustam, I. Ashraf, P. B. Washington, M. Narra and R. Shafique, "Inquest of Current Situation in Afghanistan Under Taliban Rule Using Sentiment Analysis and Volume Analysis," in IEEE Access, vol. 10, pp. 10333-10348, 2022, doi: 10.1109/ACCESS.2022.3144659.

[18] Che-Wen Chen, Shih-Pang Tseng, Ta-Wen Kuan, Jhing-Fa Wang “Outpatient Text Classification Using Attention-Based Bidirectional LSTM for Robot-Assisted Servicing in Hospital” vol. 19, 2020. URL: https://www.researchgate.net/publication/339343963_Outpatient_Text_Classification_Using_Attention-Based_Bidirectional_LSTM_for_Robot-Assisted_Servicing_in_Hospital.

[19] Yugesh Verma, “Word2Vec vs GloVe – A Comparative Guide to Word Embedding Techniques” URL: https://analyticsindiamag.com/word2vec-vs-glove-a-comparative-guide-to-wordembedding-techniques.

[20] Multilingual Universal Sentence Encoder for Semantic Retrieval. URL: https://ai.googleblog.com/2019/07/multilingual-universal-sentence-encoder.html.