



[DOI 10.28925/2663-4023.2023.20.142152](https://doi.org/10.28925/2663-4023.2023.20.142152)

УДК 004.942

Пучков Олександр Олександрович

кандидат філософських наук, професор, начальник

Інститут спеціального зв'язку та захисту інформації Національного технічного університету України
"Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна

ORCID ID: 0000-0002-8585-1044

iszzi@iszzi.kpi.ua

Ланде Дмитро Володимирович

доктор технічних наук, професор, завідувач кафедри

Навчально-науковий фізико-технічний інститут Національного технічного університету України
"Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна

ORCID ID: 0000-0003-3945-1178

dwlандe@gmail.com

Субач Ігор Юрійович

доктор технічних наук, доцент, завідувач кафедри

Інститут спеціального зв'язку та захисту інформації Національного технічного університету України
"Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна

ORCID ID: 0000-0002-9344-713X

igor_subach@ukr.net

Рибак Олександр Олегович

інженер кафедри

Інститут спеціального зв'язку та захисту інформації Національного технічного університету України
"Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна

ORCID ID: 0009-0004-1033-1599

rybak.oleksandr01@gmail.com

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ВИЗНАЧЕННЯ ПОЛІТИЧНОГО СПРЯМУВАННЯ ДЖЕРЕЛ ІНФОРМАЦІЇ ДЛЯ ЗАБЕЗПЕЧЕННЯ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ ДЕРЖАВИ ПІД ЧАС КРИЗОВИХ СИТУАЦІЙ

Анотація. Розроблено та запропоновано інформаційну технологію аналізу та визначення політичного спрямування джерел інформації в мережі Інтернет та соціальних мережах. На основі застосування бібліотеки машинного навчання та класифікації тексту fastText була створена модель для визначення політичного спрямування джерел відкритої інформації. Розроблено модулі сканування джерел відкритої інформації в мережі Інтернет та соціальних мережах, які забезпечують реалізацію таких функцій: формування баз даних шляхом збору інформації з визначених інформаційних ресурсів; налаштування модулів автоматичного сканування інформації з вебсайтів і соціальних мереж; повнотекстовий пошук; екстрагування сутностей із окремих документів; визначення політичного спрямування джерел інформації. Практичне значення отриманих результатів полягає в створенні діючого макету системи визначення політичного спрямування електронних джерел інформації. У системі передбачено екстрагування наступних сутностей із текстових повідомлень: ключові слова; персони; географічні локації (топоніми); організації. За допомогою високорівневого відкритого Python-фреймворку для розробки вебсистем Django розроблено вебінтерфейс системи. Також, для зручності роботи користувачів мобільних пристроїв з системою аналізу та визначення політичного спрямування джерел інформації в мережі Інтернет, розроблений Телеграм-бот. В інтерфейсі системи доступні функції пошуку інформації, за ключовими словами, періодом часу, а також окремо за кожним джерелом інформації (соціальною мережею). Система дає можливість дослідити кожне джерело окремо, перевірити наявність інформації про нього, тим самим доповнюючи базу даних джерел відкритої інформації. В системі автоматично формується дайджест з найбільш релевантними потребами користувачів



повідомленнями та концептами. Крім того, користувачі мають можливість переглянути найпопулярніші канали інформації за певним політичним спрямуванням. Реалізовано відображення найбільш актуальних топонімів на карті. Запропоновано варіанти можливого застосування системи. У теперішній час, система є актуальною і являє собою реалізацію інформаційної технології автоматизованого визначення політичного спрямування електронних джерел інформації.

Ключові слова: моніторинг, соціальні медіа, Інтернет, електронні джерела відкритої інформації, екстрагування сутностей, машинне навчання, штучний інтелект, інформаційна технологія.

ВСТУП

Постановка проблеми.

Політичне спрямування джерела інформації впливає на спосіб подання інформації в ньому. Відкриті джерела інформації, зокрема, в соціальних мережах, відіграють важливу роль у формуванні суспільної думки. Інформація з таких джерел може містити багато “сміття”, але часто така інформація, хоч і виглядає нейтральною, може містити певний підтекст, сенс, чи ідею сторони, яка хоче її комусь нав’язати.

Десь почута думка, може вплинути на свідомість людини, змінити її ставлення до сторони яку вона підтримує. Відповідно, окремі джерела відкритої інформації можуть завдавати шкоди різної величини. Виходячи з цього, урахування інформації з соціальних медіа має велике значення для забезпечення інформаційної безпеки людини, суспільства, держави.

У цій роботі запропоновано нову інформаційну технологію визначення політичного спрямування відкритих електронних джерел інформації та на її основі реалізовано систему, як один з інструментів вирішення важливих задач забезпечення інформаційної безпеки (ІБ) держави під час кризових ситуацій, зокрема, під час військової агресії російської федерації проти України. За основу було взято два політичні спрямування: ті хто підтримують територіальну цілісність України; ті, хто підтримують країну-агресора.

Аналіз останніх досліджень та публікацій. Аналіз літератури, пов’язаної з темою дослідження, дозволяє виділити кілька важливих напрямів, пов’язаних з питаннями роботи з відкритими даними для вирішення проблем ІБ. До першого напрямку можна віднести: збір інформації з відкритих джерел [1], аналіз інформаційних потоків [2, 3], навчання машинних бібліотек. Другий напрям присвячений: питанням автоматичного збору даних, прикладних програмних інтерфейсів (API) та інструментів й алгоритмів машинного навчання, а також застосуванню геоінформаційних методів [4]. У статті [5] порівнюється ефективність трьох різних алгоритмів векторизації слів – Word2Vec, GloVe та fastText для застосування у задачі класифікації текстів з використанням згорткових нейронних мереж (CNN); у [6] описано проблеми, пов’язані з ефективною і точною класифікацією тексту, а також запропоновано методи для покращення продуктивності та точності класифікації. У роботі [7] описується модель fastText для створення векторних представлень слів та текстів, яка дозволяє покращити ефективність і точність класифікації, та, відповідно, дозволяє враховувати інформацію про частини слів при створенні їх векторних представлень.

Метою роботи є створення інструментальних засобів для збору (постійного моніторингу) інформації, аналізу (екстрагування сутностей із текстових повідомлень) та

виявлення політичного спрямування джерел відкритої інформації для вирішення задач інформаційної безпеки держави.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Виклад основного матеріалу дослідження. Розроблена технологія включає наступні функціональні компоненти.

1. У першій - здійснюється збір інформації (сканування) з відкритих джерел (месенджерів, соціальних мереж та ін.). Підсистема сканування реалізована за допомогою засобів мови програмування Python та додаткових модулів, таких як selenium (інструмент для автоматизації роботи в веббраузері) та beautifulsoup4 (бібліотека, яка дозволяє легко отримувати інформацію з вебсторінок). Збір інформації здійснюється періодично та/або за потреби користувача. Зібрана інформація зберігається для подальшої роботи з нею.

2. Друга компонента реалізує функції навчання моделі. Для автоматичного визначення політичного спрямування каналу соціальної мережі, за основу було взято бібліотеку fastText [6, 7]. fastText — це одна з бібліотек машинного навчання для обробки текстової інформації, створена Лабораторією досліджень штучного інтелекту Facebook (*Facebook's AI Research lab, FAIR*) [8-11]. Дана бібліотека дозволяє реалізовувати моделі та алгоритми машинного навчання для отримання векторних представлень слів. Слід зазначити, що Facebook пропонує доступні попередньо натреновані моделі для 294 мов [12]. Для представлення слів в бібліотеці fastText застосовується нейронна мережа. Проте, робота з бібліотекою fastText пов'язана із вирішенням певних проблем. У наслідок того, що в досліджуваних соціальних медіа публікації ведуться найчастіше двома мовами (українською та російською), то бібліотека не завжди дає точні результати. Тому, для більш точного визначення політичного спрямування джерел, було здійснено окреме навчання модулів українською та російською мовами. Відповідно, всі повідомлення було розділено на дві частини, російськомовні та українськомовні (див. рис. 1).

Навчання здійснювалось наступним чином:

– створено чотири словники: “проукраїнські” (відповідають повідомленням проукраїнської тематики), “антиукраїнські” (відповідають повідомленням антиукраїнської тематики), і відповідно, «проросійські» та «антиросійські». За допомогою розробленого скрипту мовою Python, було перевірено кожне повідомлення на наявність у ньому кількості “позитивних” та “негативних” слів, окремо українською та окремо російською мовами. Якщо при перевірці, повідомлення містить більше “позитивних” ніж “негативних” слів (для точного визначення було взято поріг у 90%), то повідомлення маркується як проукраїнське/проросійське. Коли ж повідомлення не має явної переваги “позитивних” чи “негативних” слів, наприклад, 60% або 40%, то воно маркується як нейтральне.

Таким чином здійснювались наступні маркування: label-pos (“позитивне”); label-neg (“негативне”); label-neu (“нейтральне”).

Завантажено повідомлення з даними маркерами до бібліотеки fastText. Спочатку - всі повідомлення українською мовою, з маркуваннями “позитивні”/“негативні”/“нейтральні”, та проведено тренування першої моделі, що працюватиме з джерелами, які публікуються українською мовою. Так само проведено

навчання другої моделі, яка визначатиме політичне спрямування повідомлень написаних російською мовою.

У результаті проведених заходів, отримано дві моделі (рис. 1), кожна з яких може працювати з певною мовою, визначати політичне спрямування кожного повідомлення та маркувати його автоматично.



Рис. 1. Схеми навчання моделі fastText

На схемі, що наведено на рис. 1, відображено порядок навчання моделі. Спочатку відбувається завантаження інформаційного потоку (повідомлення із соціальних медіа) та за допомогою бібліотеки машинного навчання та класифікації тексту fastText розділення його на дві групи повідомлень: повідомлення українською мовою; повідомлення російською мовою. Далі, розділивши ці повідомлення, за допомогою оригінальних скриптів, написаних мовою програмування Python, а також словників “позитивних” і “негативних” слів, здійснюється їх підрахунок у кожному повідомленні. У результаті чого, відповідно до кількості “позитивних” та “негативних” слів, здійснюється маркування повідомлень: позитивних (`__label-pos`); негативних (`__label-neg`); нейтральних (`__label-neu`). Далі, на основі отриманих та завантажених до бібліотеки fastText повідомлень, здійснюється тренування моделі. Таким чином, у результаті відпрацьованих дій, отримано дві моделі для визначення політичного спрямування джерел відкритої інформації.

3. Визначення політичного спрямування джерела інформації здійснюється шляхом підрахунку маркованих повідомлень у ньому. При цьому, задається поріг спрацьовування класифікаційного правила. Наприклад, якщо кількість негативних повідомлень перевищує 70%, то джерело – негативне, тобто в нашому випадку підтримує країну-агресора. Відповідно, джерела, спрямування яких визначити важко, маркуються як нейтральні. Після визначення найбільшого відсотку вмісту повідомлень певного спрямування, здійснюється розрахунок рейтингу джерел, наприклад, від найбільш негативних, тобто тих, які явно підтримують агресора та війну, до найменш негативних. За допомогою постійного моніторингу джерел та маркування повідомлень в них, розрахунок рейтингу можна здійснювати з більшою точністю.

Для ефективного використання системи розроблено вебінтерфейс для інтеграції її з програмними застосунками, написаними мовою програмування JavaScript, а також з високорівневим відкритим Python-фреймворком Django для розробки вебсистем.

Початкова сторінка створеного сайту (рис. 2) показує користувачу топ каналів за двома політичними спрямуваннями - проукраїнськими та проросійськими. На ній подається загальний контекст зі всіх джерел, наявних в базі: ключові слова, персони, локації та організації, що, на даний час, обговорюються в джерелах відкритого доступу. Наприклад, складна ситуація у Бахмуті, боротьба церков УПЦ та ПЦУ, передача Україні американської військової техніки (танки Abrams, тощо.), взяття під варту Дональда Трампа та ін. Саме ці події були найбільш обговорюваними під час роботи над цією статтею. Тому, результат виводу контексту, зображений на рис. 2, відповідає інформації на даний проміжок часу.

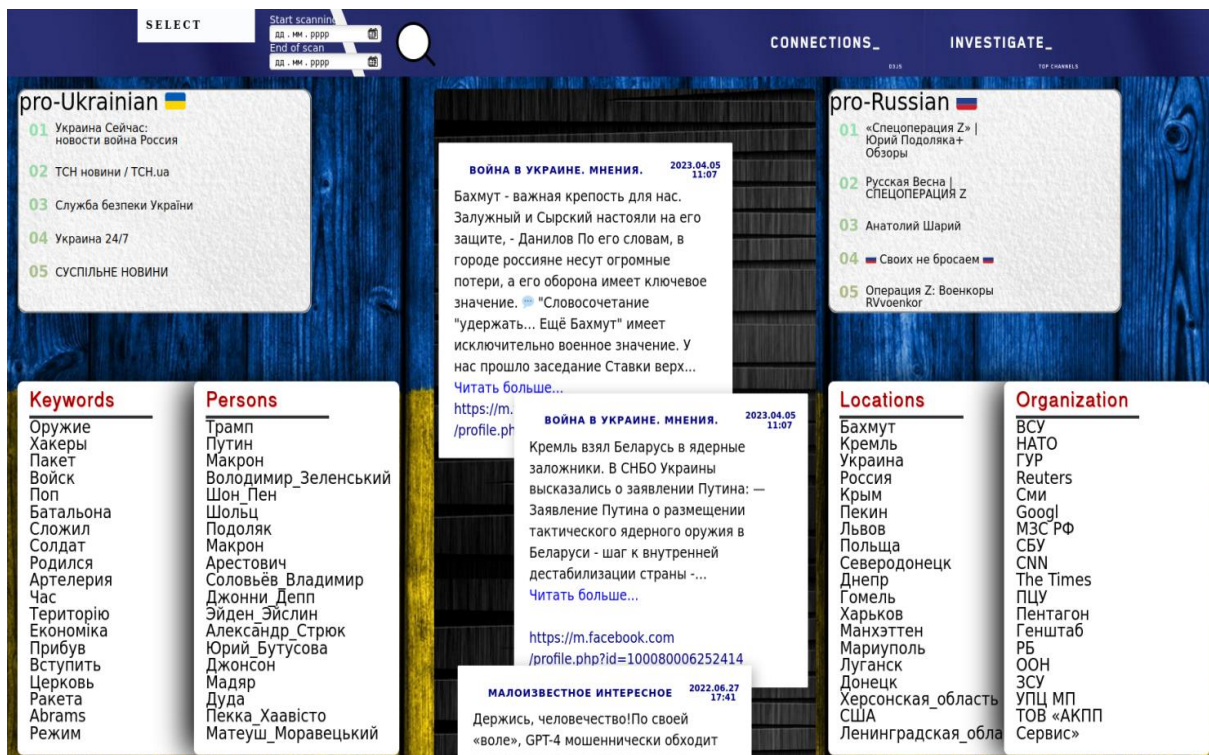


Рис.2. Вебінтерфейс системи

За потреби, користувач має можливість переглянути інформацію та контекст за певний період часу, вказавши період у вікні дати в форматі, наприклад: 01.04.2023 по 12.04.2023.

При натиску на карточку Locations відкриється карта, яка відмічає дані топоніми (рис. 3).

У лівому верхньому куті є кнопка 'INVESTIGATE_' для переходу на наступну сторінку, що дає можливість дослідити будь яке джерело окремо (див. рис. 4).

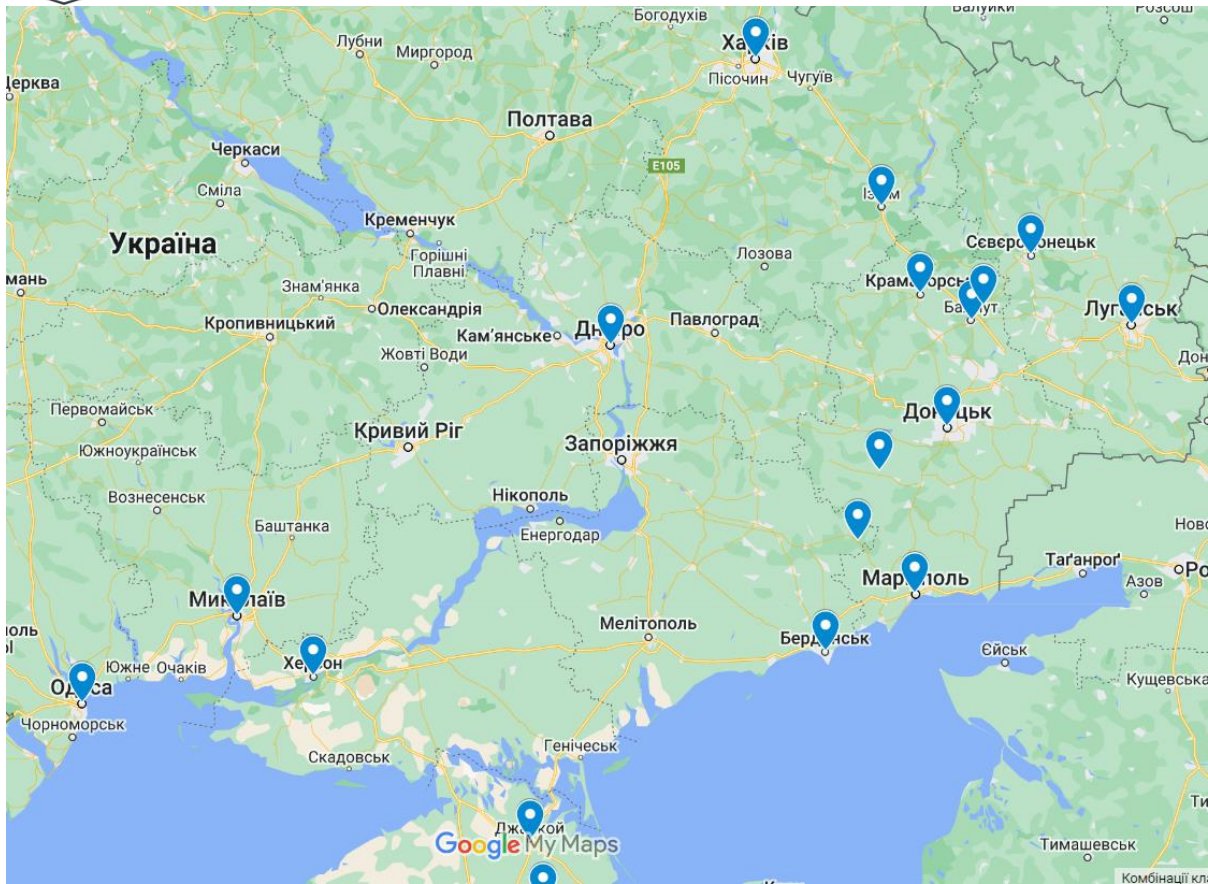
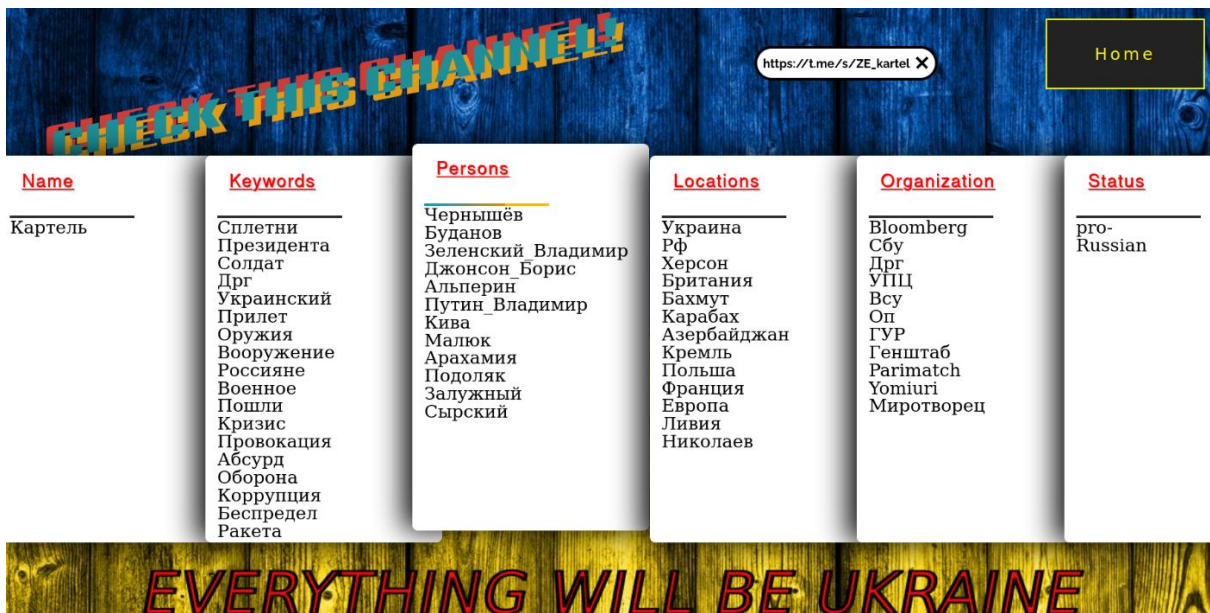


Рис. 3. Приклад відображення топонімів на картах



| Name | Keywords | Persons | Locations | Organization | Status |
|---------|---|--|---|---|-------------|
| Картель | Сплетни Президента Солдат Дрг Украинский Прилет Оружия Вооружение Россияне Военное Пошли Кризис Провокация Абсурд Оборона Коррупция Беспредел Ракета | Чернышёв Буданов Зеленский Владимир Джонсон Борис Альперин Путин Владимир Кива Малюк Арахамия Подольак Залужный Сырский | Украина Рф Херсон Британия Бахмут Карабах Азербайджан Кремль Польша Франция Европа Ливия Николаев | Bloomberg Сбу Дрг УПЦ Всу Оп ГУР Генштаб Parimatch Yomiuri Миротворец | pro-Russian |

Рис. 4. Сторінка дослідження джерела інформації

Для того щоб дослідити джерело, потрібно вставити посилання нього (Telegram-канал, Facebook-сторінка, тощо) в поле пошуку та активізувати його, в результаті чого запуститься розроблений скрипт, який збере інформацію з джерела та завантажить її в натреновану модель бібліотеки fastText, на основі якої відбувається визначення його

політичного спрямування. На виході отримуємо контекст цього джерела, відображений у карточках та його статус, в даному випадку pro-russian.

Схему роботи модуля виводу інформації про джерело наведено на рис. 5.

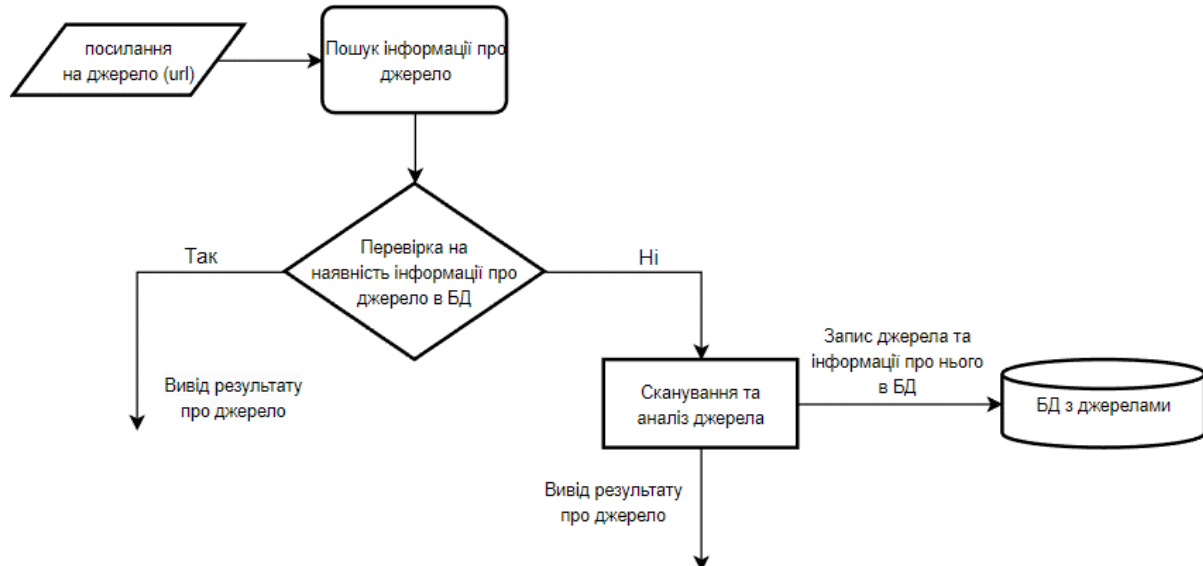


Рис 5. Схema роботи модуля виводу інформації про джерело

З рис. 5 видно, що коли користувач вводить в поле пошуку посилання на джерело, система починає його пошук у базі даних. У разі наявності інформації про джерело, тобто його політичного спрямування та контексту, система надає інформацію про нього. У результаті того, що система періодично сканує джерела відкритої інформації, ці дані будуть актуальними. У разі відсутності інформації про джерело, система запускає модулі сканування, збирає інформацію з цього джерела за останні декілька днів, аналізує її, визначає її політичне спрямування і контекст та надає результат пошуку користувачеві системи. Посилання введене користувачем додається до бази даних джерел відкритої інформації для подальшого сканування та збору інформації з цього джерела.

Сканування джерел, що записані у БД, здійснюється періодично, за розписом наприклад, кожні дві години. Таким чином інформація, що міститься у системі, завжди є актуальною. Також, інформація про джерело, його спрямування та контекст завжди обновлюється та також є актуальною.

Телеграм-бот. Для того, щоб користувачам мобільних пристроїв було зручно використовувати систему аналізу та визначення політичного спрямування джерел інформації в мережі Інтернет, за допомогою мови програмування Python був розроблений Телеграм-бот (рис. 6). Він містить такі ж самі можливості, що і розглянута система, окрім пошуку повідомлень за датою публікації. Можливість легко використовувати систему з Телеграм-боту може привернути більше користувачів та спростити проблему пошуку джерел відкритої інформації.

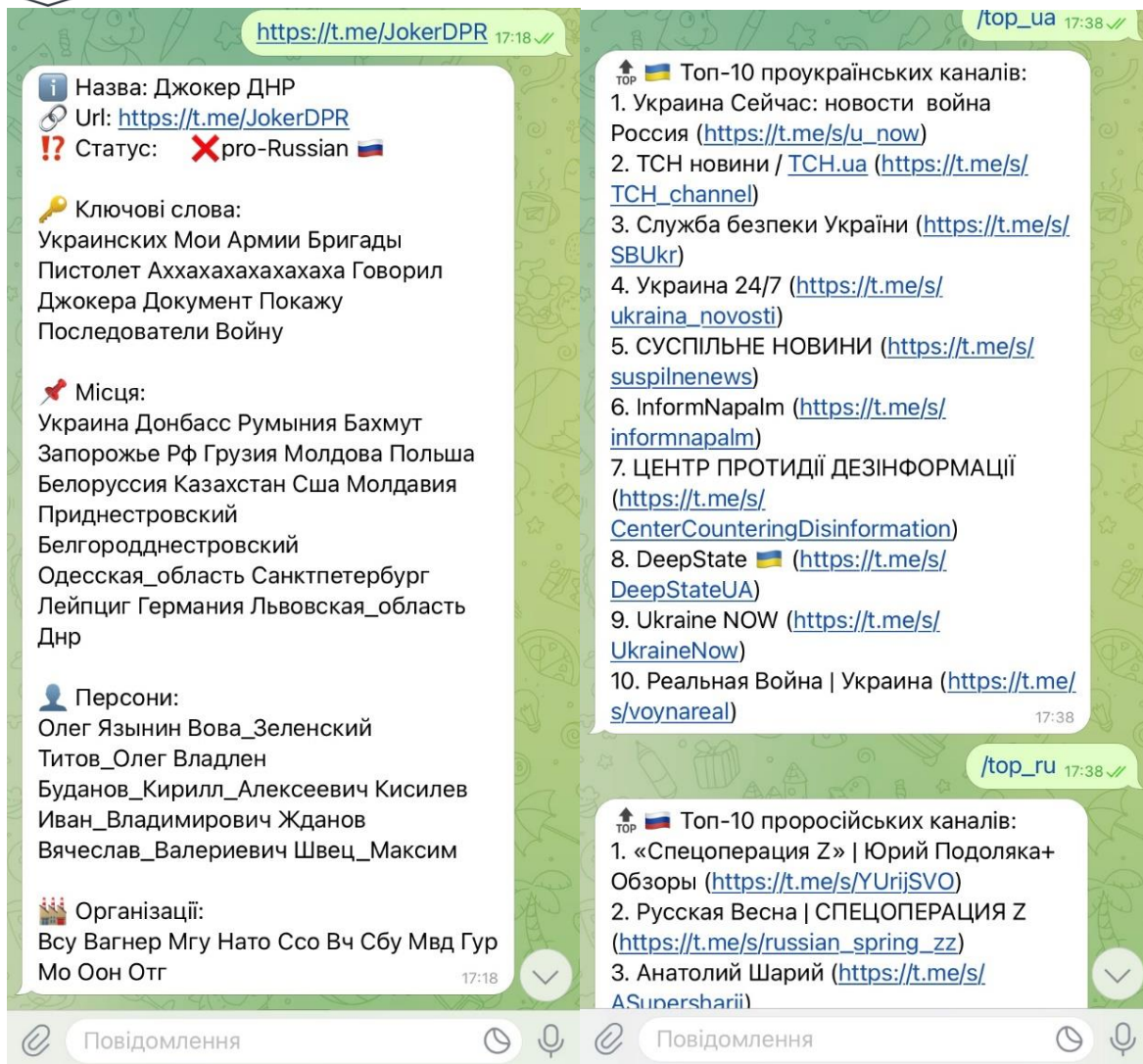


Рис. 6. Приклади роботи із системою у Телеграм-боті

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

До цього часу у відкритому доступі не існувало засобів автоматизованого визначення політичного спрямування інформаційних ресурсів. Ця задача вирішувалась епізодично в окремих інформаційних центрах із залученням експертів. У цій роботі було розглянуто актуальність даного питання в контексті умов сьогодення та запропоновано систему автоматизованого визначення політичного спрямування відкритого електронного джерела інформації.

Розроблено інформаційну технологію та інструментальні засоби її реалізації – комплекс програмних модулів для навчання моделей за допомогою бібліотеки fastText, які дають можливість визначати політичне спрямування відкритого електронного джерела інформації на основі застосування технологій нейронних мереж.

За допомогою мови програмування Python та її вебфреймворків було розроблено вебінтерфейс та Телеграм-бот для використання системи з мобільних пристроїв. Ця можливість надає користувачу можливість з легкістю визначити політичне спрямування



та контекст джерела, звідки він збирається та отримує інформацію. Ще однією перевагою запропонованої інформаційної технології є те, що користувачі мають можливість самостійного нарощування бази даних відкритими джерелами інформації для забезпечення зворотного зв'язку та краудсорсингу.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- 1 Ланде, Д., Пучков, О., Субач, І. (2020). Система аналізу великих обсягів даних з питань кібербезпеки із соціальних медіа. *Collection Information Technology and Security*, 8(1), 4–18. <https://doi.org/10.20535/2411-1031.2020.8.1.217993>
- 2 Lande, D., Subach, I., Puchkov, O., Soboliev, A. (2021). A Clustering Method for Information Summarization and Modelling a Subject Domain. *Information & Security: An International Journal* 50, 1, 79-86. <https://doi.org/10.11610/isij.5013>.
- 3 Lande, D., Snarskii, A., Dmytrenko, O., Subach, I. (2020). Relaxation time in complex network. У *ARES 2020: The 15th International Conference on Availability, Reliability and Security*. ACM. <https://doi.org/10.1145/3407023.3409231>
- 4 Layton, R., Watters, P. A. (2016) Automating open source intelligence: algorithms for OSINT: Elsevier, Syngress. <https://doi.org/10.1016/C2014-0-02170-3>.
- 5 Dharma, E., Gaol, F., Warnars, H., Soewito, B. (2022). The accuracy comparison among word2vec, glove, and fasttext towards convolution neural network (cnn) text classification. Jakarta.
- 6 Bojanowski, P., Grave, E., Joulin, A., Mikolov, T. (2016). Bag of Tricks for Efficient Text Classification. <https://arxiv.org/abs/1607.01759>
- 7 Bojanowski, P., Grave, E., Joulin, A., Mikolov, T. (2017). Enriching Word Vectors with Subword Information. <https://arxiv.org/abs/1607.04606>
- 8 Ryan, K. (2016). Facebook's New Open Source Software Can Learn 1 Billion Words in 10 Minutes. <https://www.inc.com/kevin-j-ryan/facebook-open-source-fasttext-learns-1-billion-words-in-10-minutes.html>
- 9 Low, C. (2016). Facebook is open-sourcing its AI bot-building research. <https://www.engadget.com/2016-08-18-facebook-open-sourcing-fasttext.html>
- 10 Mannes, J. (2016). Facebook's Artificial Intelligence Research lab releases open source fastText on GitHub. <https://techcrunch.com/2016/08/18/facebooks-artificial-intelligence-research-lab-releases-open-source-fasttext-on-github>
- 11 Mannes, J. (2017). Facebook's fastText library is now optimized for mobile. <https://techcrunch.com/2017/05/02/facebooks-fasttext-library-is-now-optimized-for-mobile>
- 12 Sabin, D. (2017). Facebook Makes A.I. Program Available in 294 Languages. <https://www.inverse.com/article/31075-facebook-machine-learning-language-fasttext>



Oleksandr Puchkov

candidate of philosophy science, professor, head
Institute of special communications and information security National technical university of Ukraine Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine
ORCID ID: 0000-0002-8585-1044
iszzi@iszzi.kpi.ua

Dmytro Lande

doctor of technical science, professor, head of department
Educational and Scientific Physical and Technical Institute of the National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
ORCID ID: 0000-0003-3945-1178
dwlande@gmail.com

Ihor Subach

doctor of technical science, associate professor, head of department
Institute of special communications and information security National technical university of Ukraine Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine
ORCID ID: 0000-0002-9344-713X
igor_subach@ukr.net

Oleksandr Rybak

department engineer
Institute of special communications and information security National technical university of Ukraine Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine
ORCID ID: 0009-0004-1033-1599
rybak.oleksandr01@gmail.com

INFORMATION TECHNOLOGY FOR DETERMINING THE POLITICAL DIRECTION OF INFORMATION SOURCES TO ENSURE THE INFORMATION SECURITY OF THE STATE DURING CRISIS SITUATIONS

Abstract. The information technology of analysis and determination of the political orientation of information sources on the Internet and social networks has been developed and proposed. Based on the application of the machine learning library and fastText text classification, a model was created to determine the political orientation of open information sources. Modules for scanning sources of open information on the Internet and social networks have been developed, which ensure the implementation of the following functions: the formation of databases by collecting information from specified information resources; setting up modules for automatic scanning of information from websites and social networks; full-text search; extracting entities from individual documents; determination of the political orientation of information sources. The practical significance of the obtained results lies in the creation of a working layout of the system for determining the political orientation of electronic sources of information. The system provides extraction of the following entities from text messages: keywords; persons; geographical locations (toponyms); organizations. The web interface of the system was developed with the help of the high-level open Python framework for the development of Django web systems. Also, for the convenience of mobile device users with the system of analysis and determination of the political direction of information sources on the Internet, the Telegram bot was developed. Information search functions are available in the system interface by keywords, time period, and separately by each information source (social network). The system makes it possible to examine each source separately, to check the availability of information about it, thereby supplementing the database of open information sources. The system automatically generates a digest with messages and concepts most relevant to users' needs. In addition, users have the opportunity to view the most popular channels of information according to a certain political orientation. The display of the most relevant toponyms on the map has been implemented. Variants of possible application of the system are proposed. Currently, the system is relevant and represents the implementation of information technology for the automated determination of the political direction of electronic sources of information.



Keywords: monitoring, social media, Internet, electronic sources of open information, essence extraction, machine learning, artificial intelligence, information technology.

REFERENCES (TRANSLATED AND TRANSLITERATED)

- 1 Lande, D., Puchkov, O., Subach, I. (2020). Systema analizu velykykh obsiahiv danykh z pytan kiberbezpeky iz sotsialnykh media. Collection Information Technology and Security, 8(1), 4–18. <https://doi.org/10.20535/2411-1031.2020.8.1.217993>
- 2 Lande, D., Subach, I., Puchkov, O., Soboliev, A. (2021). A Clustering Method for Information Summarization and Modelling a Subject Domain. Information & Security: An International Journal 50, 1, 79-86. <https://doi.org/10.11610/isi.5013>.
- 3 Lande, D., Snarskii, A., Dmytrenko, O., Subach, I. (2020). Relaxation time in complex network. U ARES 2020: The 15th International Conference on Availability, Reliability and Security. ACM. <https://doi.org/10.1145/3407023.3409231>
- 4 Layton, R., Watters, P. A. (2016) Automating open source intelligence: algorithms for OSINT: Elsevier, Syngress. <https://doi.org/10.1016/C2014-0-02170-3>.
- 5 Dharma, E., Gaol, F., Warnars, H., Soewito, B. (2022). The accuracy comparison among word2vec, glove, and fasttext towards convolution neural network (cnn) text classification. Jakarta.
- 6 Bojanowski, P., Grave, E., Joulin, A., Mikolov, T. (2016). Bag of Tricks for Efficient Text Classification. <https://arxiv.org/abs/1607.01759>
- 7 Bojanowski, P., Grave, E., Joulin, A., Mikolov, T. (2017). Enriching Word Vectors with Subword Information. <https://arxiv.org/abs/1607.04606>
- 8 Ryan, K. (2016). Facebooks New Open Source Software Can Learn 1 Billion Words in 10 Minutes. <https://www.inc.com/kevin-j-ryan/facebook-open-source-fasttext-learns-1-billion-words-in-10-minutes.html>
- 9 Low, C. (2016). Facebook is open-sourcing its AI bot-building research. <https://www.engadget.com/2016-08-18-facebook-open-sourcing-fasttext.html>
- 10 Mannes, J. (2016). Facebooks Artificial Intelligence Research lab releases open source fastText on GitHub. <https://techcrunch.com/2016/08/18/facebook-artificial-intelligence-research-lab-releases-open-source-fasttext-on-github>
- 11 Mannes, J. (2017). Facebooks fastText library is now optimized for mobile. <https://techcrunch.com/2017/05/02/facebook-fasttext-library-is-now-optimized-for-mobile>
- 12 Sabin, D. (2017). Facebook Makes A.I. Program Available in 294 Languages. <https://www.inverse.com/article/31075-facebook-machine-learning-language-fasttext>

