

ЛІНГВІСТИЧНИЙ АНАЛІЗ ДАНИХ ІНТЕРНЕТ-МЕДІА ТА СОЦІАЛЬНИХ МЕРЕЖ У ЗАДАЧАХ ОЦІНЮВАННЯ СУСПІЛЬНИХ ПЕРЕТВОРЕНЬ¹

Анотація. Розроблено комбінований підхід до оцінювання ефективності суспільних перетворень як міри узгодженості між діями влади і очікуваннями суспільства та синергії (соціальної активності) людей, що ґрунтується на формалізованому узгодженні результатів, отриманих методом експертних оцінок, а також методами сентимент-аналізу та інтелектуального аналізу текстових повідомлень з відкритих онлайн-джерел і соціальних мереж. Ці методи реалізовано у вигляді комплексу вебсервісів та застосунків у середовищі розроблення інтегрованої онлайн-платформи Advanced Analytics Світового центру даних «Геоінформатика і сталий розвиток». Ефективність запропонованої методики продемонстровано на прикладі кількісного оцінювання ставлення населення України до дій влади, спрямованих на протидію поширенню епідемії COVID-19.

Ключові слова: вектор дій влади, вектор очікувань суспільства, вектор перетворень (реформ), лінгвістичний аналіз, контент-аналіз, аналіз даних інтернет-медіа та соціальних мереж, розвідка за відкритими джерелами.

ЗАГАЛЬНИЙ ПІДХІД ДО ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ СУСПІЛЬНИХ ПЕРЕТВОРЕНЬ

Проблема взаємодії органів влади з громадськістю є надзвичайно актуальною в будь-якій країні світу, її окремому регіоні, чи в будь-якому великому корпоративному чи суспільному утворенні. Для підготовки та обґрунтування управлінських рішень, прогнозування конфліктів і кризових ситуацій та запобігання їм існує потреба в отриманні кількісної оцінки рівня підтримки суспільством дій влади.

Будемо вважати, що суспільні перетворення (реформи), що оцінюються показником суспільних перетворень R , будуть настільки ефективними, наскільки вектор очікувань та соціальної активності суспільства \vec{S} буде за напрямом наближатися до вектора дій влади \vec{G} та наскільки високою і позитивною буде синергія суспільства (рис. 1). Під синергією суспільства будемо розуміти додаткову суспільну активність (суспільну енергію), яка утворюється в результаті взаємодії громадян країни та може бути позитивною, коли суспільство усвідомлює, що успіхи породжують нові успіхи і від'ємною, коли невдачі призводять до втрати надії людей на позитивні зміни (апатії) і, як наслідок, породжують нові невдачі.

Нехай α — це кут між векторами \vec{S} і \vec{G} , а K_S — коефіцієнт синергії суспільства. Для різних станів суспільства, які можуть змінюватися від суспільного підйому та оптимізму до втоми і апатії, коефіцієнт синергії суспільства K_S може набувати різних значень. У цьому дослідженні будемо розглядати такий ряд значень коефіцієнта синергії суспільства: $K_S = [0.25; 0.5; 0.75; 1.0]$.

Зазвичай залежність синергії суспільства від кута α є нелінійною. Граничні значення кута α є такими: $\alpha = 0^\circ$, коли напрямки вказаних двох векторів збігаються (максимальні суспільні перетворення та максимальна позитивна синергія

¹ Дослідження виконується за часткової підтримки Національного фонду досліджень України (грант 2020.01/0283) та Міністерства освіти та науки України (НДДКР0120U102060).

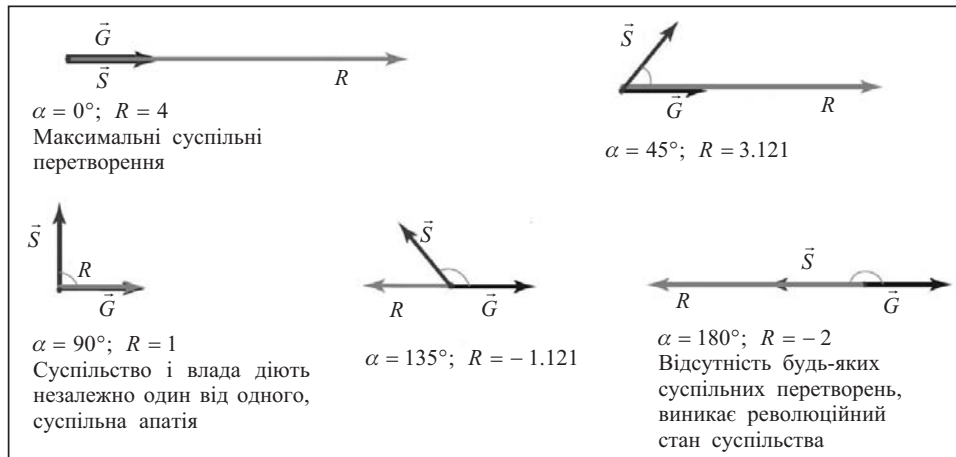


Рис. 1. Ефективність суспільних перетворень (масштаб реформ) для $K_S = 0.5$

суспільства), та $\alpha = 180^\circ$, коли їхні напрями протилежні (відсутність будь яких суспільних перетворень, утворюється максимальна негативна суспільна енергія, яка значно перевищує енергію дій влади, в результаті чого виникає революційний стан суспільства). У випадку $\alpha = 90^\circ$ вектор дій влади та вектор соціальної активності (очікувань) суспільства будуть ортогональними. Це означатиме, що для $\alpha = 90^\circ$ влада і суспільство діятимуть незалежно один від одного.

Для визначення показника суспільних перетворень R , значення якого буде відповідати масштабам суспільних перетворень, скористаємося моделлю, запропонованою авторами в роботі [1]:

$$R = |\vec{G}| + |\vec{S}| \cos \alpha + K_S (|\vec{G}| + |\vec{S}|)^2 \cos \alpha; \quad \alpha \in [0^\circ, 180^\circ]. \quad (1)$$

У цій моделі оцінювання кута α (з подальшим розрахунком показника суспільних перетворень R) здійснюється як з використанням методу експертних оцінок (методу Дельфі), так і шляхом автоматичного онлайн-моніторингу суспільних настроїв в інформаційному просторі, або обома способами одразу.

ОЦІНЮВАННЯ МАСШТАБУ ПОТЕНЦІЙНИХ СУСПІЛЬНИХ ПЕРЕТВОРЕНЬ З ВИКОРИСТАННЯМ МЕТОДУ ДЕЛЬФІ

Визначення експертних оцінок кута α , отриманих з використанням методу Дельфі, здійснюватиметься за шкалою Міллера (1, ..., 7). При цьому значення шкали Міллера 7 відповідає повній підтримці суспільством дій влади (кут $\alpha = 0^\circ$), а значення 1 вказує на те, що суспільство повністю не підтримує дії влади (кут $\alpha = 180^\circ$). Виберемо крок дискретизації шкали Міллера 0.5. Переведемо значення шкали Міллера у значення кута α та з використанням моделі (1) розрахуємо значення показника суспільних перетворень R для різних значень коефіцієнта синергії суспільства K_S (табл. 1).

Для узагальнення індивідуальних експертних думок та отримання колективної експертної оцінки скористаємося зваженою сумою

$$E = \sum_{i=1}^n w_i \cdot e_i, \quad (2)$$

де E — колективна експертна оцінка; $w_i, i = \overline{1, n}$, $\sum_{i=1}^n w_i = 1$, — вага індивідуальної оцінки i -го експерта; $e_i, i = \overline{1, n}$, — індивідуальна оцінка i -го експерта.

Таблиця 1. Результати оцінювання ефективності суспільних перетворень

Шкала Міллера	Кут відхилення α (°)	Показник суспільних перетворень R			
		$K_S = 0.25$	$K_S = 0.5$	$K_S = 0.75$	$K_S = 1.0$
7	0	3.000	4.000	5.000	6.000
6.5	15	2.932	3.898	4.864	5.830
6	30	2.732	3.598	4.464	5.330
5.5	45	2.414	3.121	3.828	4.536
5	60	2.000	2.500	3.000	3.500
4.5	75	1.518	1.776	2.035	2.294
4	90	1.000	1.000	1.000	1.000
3.5	105	0.482	0.224	-0.035	-0.294
3	120	0.000	-0.500	-1.000	-1.500
2.5	135	-0.414	-1.121	-1.828	-2.536
2	150	-0.732	-1.598	-2.464	-3.330
1.5	165	-0.932	-1.898	-2.864	-3.830
1	180	-1.000	-2.000	-3.000	-4.000

Ваги $w_i, i = \overline{1, n}$, повинні бути пропорційними відхиленню індивідуальної експертної оцінки від усередненого значення всіх індивідуальних експертних оцінок, тому їх обчислюють за формулами

$$d_i = |\bar{E} - e'_i|; e'_i = \frac{e_i - M_{\min}}{M_{\max} - M_{\min}}; \bar{E} = \frac{\sum_{i=1}^n e'_i}{n}; \quad (3)$$

$$w_i = \frac{1 - d_i}{\sum_{i=1}^n (1 - d_i)},$$

де $d_i, i = \overline{1, n}$, — відстань нормованої індивідуальної оцінки e'_i i -го експерта від усередненої за групою експертів оцінки \bar{E} , $M_{\min} = 1$ та $M_{\max} = 7$ — відповідно найменше та найбільше значення шкали Міллера, інші позначення такі самі, як у наведених вище формулах.

Оскільки шкала Міллера є ординальною шкалою, то для оцінювання узгодженості колективної експертної оцінки C скористаємося зваженою сумарною відстанню індивідуальних експертних оцінок від оцінки, отриманої за формулою (2):

$$C = \frac{\sum_{i=1}^n |E - e'_i|}{C_{\max}}; C_{\max} = n(n-1)(M_{\max} - M_{\min}). \quad (4)$$

Значення $C = 1$ відповідає повністю узгодженій оцінці (коли всі експерти дали одну й ту саму оцінку), значення $C = 0$ відповідає ситуації, коли множина отриманих індивідуальних оцінок містить однакову кількість мінімальних (M_{\min}) і максимальних (M_{\max}) значень за шкалою Міллера, що свідчить про відсутність консенсусу в експертній групі та неможливість отримання узгодженої колективної експертної думки.

У тому разі, коли потрібно отримати узагальнену колективну експертну думку щодо k окремих питань, як таку оцінку можна використати зважену суму

$$\tilde{E} = \sum_{i=1}^k w'_i E_i; w'_i = \frac{C_i}{\sum_{j=1}^k C_j}, \quad (5)$$

де \tilde{E} — колективна експертна оцінка для множини з k запитань, E_i — колективна експертна оцінка для i -го запитання, обчислена за формулою (2), w'_i , $\sum_{i=1}^k w'_i = 1$, — ваговий коефіцієнт оцінки E_i , пропорційний її узгодженості C_i , обчисленій за формулою (4).

У цьому випадку узгодженість оцінки \tilde{E} можна визначити як усереднене значення:

$$\tilde{C} = \frac{\sum_{i=1}^k C_i}{k}. \quad (6)$$

Процедура проведення експертного опитування згідно із запропонованим підходом до оцінювання масштабу суспільних перетворень передбачає формування експертної групи, надання експертам опитувальника, що містить запитання стосовно конкретних дій влади протягом певного часового періоду, отримання від експертів індивідуальних оцінок кута α за шкалою Міллера, обчислення колективних експертних оцінок та їхньої узгодженості за формулами (2)–(6), а також інтерпретацію отриманих результатів.

АВТОМАТИЧНЕ ОНЛАЙН-ОЦІНЮВАННЯ МАСШТАБУ ПОТЕНЦІЙНИХ СУСПІЛЬНИХ ПЕРЕТВОРЕНЬ

Для автоматичного онлайн-оцінювання кута α (як міри неузгодженості дій влади та очікувань суспільства) та показника суспільних перетворень R скористаємося методом аналізу емоційного забарвлення інформаційних повідомлень, отриманих шляхом онлайн-моніторингу відповідних медіа-ресурсів та соціальних мереж.

За статистикою в соціальних мережах накопичується 30–40 % інформації, пов'язаної з реакцією користувачів мережі на важливі суспільні події та діяльність органів державної влади. Специфіка цієї інформації, а саме її неструктурований характер, низький рівень персистентності, наявність фейкових повідомлень, зумовлює потребу в розробленні спеціалізованих інформаційних систем, основне призначення яких — збирання, зберігання та аналітичне опрацювання неструктурованих даних, отриманих з онлайн-медіа та соціальних мереж.

Концепція одночасного застосування методів і засобів інформаційного пошуку, аналізу та агрегування даних із сучасних інформаційних потоків, покладена в основу таких систем, надає їм змогу здійснювати автоматичне сканування та первинне опрацювання інформації з вебсайтів, соціальних мереж, месенджерів тощо, створювати повнотекстові бази даних, виявляти схожі за змістом інформаційні повідомлення, екстрагувати поняття та проводити лінгвістичний контент-аналіз текстових повідомлень, визначати тональність повідомлень, виконувати дослідження динаміки тематичних інформаційних потоків. У межах завдання, розглянутого в цій роботі, особливе значення має виявлення емоційної забарвленості інформації у процесі контент-аналізу публікацій [2]. Тональність, як емоційне ставлення автора висловлювання до деякого об'єкту, виражене в тексті, в цілому можна визначити як функцію (в найпростішому випадку суму) лексичних тональностей складових його одиниць (речень) і правил їхнього поєднання [3]. У роботі запропоновано скористатися методом визначення тональності інформації шляхом застосування статистичних закономірностей, пов'язаних із присутністю певних лексем у текстах, наївного баєсівського підходу та методу нейронних мереж [4].

Запропонований підхід до кількісного оцінювання масштабу суспільних перетворень R ґрунтується на гіпотезі про адекватне відображення емоційної забарвленості повідомлень на складові вектора суспільних очікувань \vec{S} .

Розглянемо вектор суспільних очікувань \vec{S} як

$$\vec{S} = \vec{S}_- + \vec{S}_= + \vec{S}_+,$$

де \vec{S}_- — складова негативного ставлення (дій), яка відповідає протидії суспільства діям влади; $\vec{S}_=$ — нейтральна складова, яка відповідає нейтральному ставленню суспільства до дій влади; \vec{S}_+ — позитивна складова, яка відповідає схваленню (підтримці) суспільством дій влади. Відповідно до цього визначення вектори \vec{S}_- та \vec{S}_+ є колінеарними до вектора дій влади \vec{G} (напрямки векторів \vec{S}_+ та \vec{G} збігаються, напрямки векторів \vec{S}_- та \vec{G} є протилежними), а вектор $\vec{S}_=$ є ортогональним до вектора \vec{G} . Якщо вибрати ортогональний базис у такий спосіб, щоб виконувалась умова $\vec{G} = (0, 1)$, тоді справджуються такі співвідношення:

$$\begin{aligned} \vec{S}_- &= (-|\vec{S}_-|, 0), \quad \vec{S}_= = (0, |\vec{S}_=|), \quad \vec{S}_+ = (|\vec{S}_+|, 0), \\ \vec{S} &= (|\vec{S}_+| - |\vec{S}_-|, |\vec{S}_=|), \\ |\vec{S}| &= \sqrt{(|\vec{S}_+| - |\vec{S}_-|)^2 + |\vec{S}_=|^2}, \\ \cos \alpha &= \frac{|\vec{S}_+| - |\vec{S}_-|}{|\vec{S}|}, \quad \alpha \in [0^\circ, 180^\circ]. \end{aligned} \quad (7)$$

Будемо вважати, що соціальна активність суспільства відображається у потоці повідомлень різних медіа-ресурсів та соціальних мереж, а емоційна забарвленість цих повідомлень відповідає ставленню (діям) окремих індивідуумів та (або) соціальних груп до тих чи інших подій, у тому числі й тих, що є наслідком діяльності влади.

Нехай певна дія влади \vec{G} зумовила вибірку повідомлень

$$M = M_- \cup M_= \cup M_+, \quad M_i \cap M_j = \emptyset, \quad i, j \in \{-, =, +\}, \quad i \neq j,$$

де $\{-, =, +\}$ — негативне, нейтральне та позитивне емоційне забарвлення повідомлень відповідно. Тоді справджуються такі співвідношення:

$$\begin{aligned} |\vec{S}_-| \sim |M_-| \sim p_- &= \frac{|M_-|}{|M|}, \\ |\vec{S}_=| \sim |M_=| \sim p_= &= \frac{|M_=|}{|M|}, \\ |\vec{S}_+| \sim |M_+| \sim p_+ &= \frac{|M_+|}{|M|}, \end{aligned} \quad (8)$$

де $p_i, i \in \{-, =, +\}$, — частка повідомлень з відповідно негативною, нейтральною та позитивною емоційною забарвленістю.

З урахуванням (4) отримаємо оцінку

$$|\vec{S}| \sim \sqrt{(p_+ - p_-)^2 + p_=^2}. \quad (9)$$

Формули (7)–(9) будемо використовувати в моделі (1), а отриману кількісну оцінку показника суспільних перетворень R , приведену до діапазону $(1, \dots, 7)$ відповідно до значень шкали Міллера (див. табл. 1) — для перевірки узгодженості кількісних та експертних оцінок: показника суспільних перетворень R (масштабу реформ) та кута α (як міри неузгодженості дій влади та очікувань суспільства).

Розглянемо випадки, коли $p_+ = p_-$, $\cos \alpha = 0$. Якщо $p_+ = 1$, це свідчить про те, що вся множина повідомлень має нейтральне емоційне забарвлення. У випадку $p_+ = p_- = 0.5$, $p_+ = 0$ спостерігаємо в суспільстві максимальну поляризацію думок щодо дій влади, проте в обох випадках $R = 1$.

Для того, щоб розрізнити ці стани, визначимо коефіцієнт конфліктності як

$$C = -(p_+ + p_-) \frac{c_+ \log c_+ + c_- \log c_-}{\log 2},$$

де $c_+ = p_+ / (p_+ + p_-)$, $c_- = p_- / (p_+ + p_-)$.

Коефіцієнт конфліктності C є зваженою ентропією та набуває значень у діапазоні $[0, 1]$. Стану $p_+ = p_- = 0.5$, $p_+ = 0$, відповідає значення $C = 1$.

ОНЛАЙН-ПЛАТФОРМА ADVANCED ANALYTICS ДЛЯ ОЦІНЮВАННЯ СТАВЛЕННЯ СУСПІЛЬСТВА ДО ДІЙ ВЛАДИ

Запропоновані методи використано для реалізації спеціалізованого застосунку класу Social BI [5], розробленого у вигляді тематичної панелі (dashboard) за допомогою інструментів та сервісів онлайн-платформи Advanced Analytics Світового центру даних «Геоінформатика і сталий розвиток» [6].

Ця інтеграційна платформа є розподіленою інформаційною системою, до складу якої входять сервіси для розроблення та публікації односторінкових застосунків (single page application, SPA), сервіси управління даними та їхнього

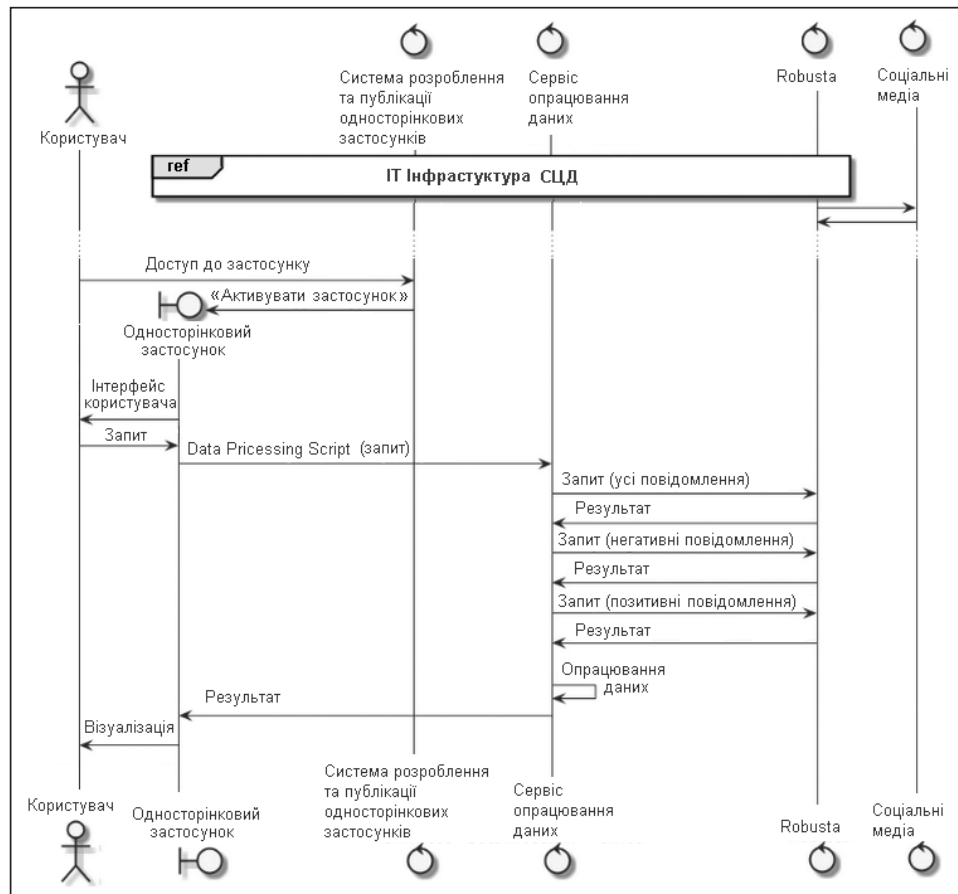


Рис. 2. Схема взаємодії компонентів онлайн-платформи Advanced Analytics Світового центру даних «Геоінформатика та сталий розвиток» під час опрацювання запитів SPA «PRO ET CONTRA»

опрацювання, сервіси автоматичного збору та інтелектуального аналізу великих даних. Особливістю цієї системи є використання такого інструментарію інтеграції, розробленого у Світовому центрі даних «Геоінформатика і сталий розвиток» (СЦД-Україна), який надає змогу об'єднувати різноманітні гетерогенні ресурси та сервіси в межах єдиної IT-інфраструктури [7, 8].

Схему взаємодії компонентів онлайн-платформи наведено на рис. 2.

У процесі реалізації застосунку для оцінювання ставлення суспільства до дій влади на основі аналізу даних інтернет-медіа та соціальних мереж (SPA «PRO ET CONTRA») було застосовано такі компоненти онлайн-платформи: систему розроблення та публікації односторінкових застосунків (SPA Host); сервіс опрацювання даних (Data Processing Service); систему моніторингу соціальних медіа (Media Sources) та лінгвістичного аналізу великих обсягів даних, отриманих із соціальних медіа (Robusta).

ПРАКТИЧНЕ ЗАСТОСУВАННЯ ОНЛАЙН-ПЛАТФОРМИ ADVANCED ANALYTICS НА ПРИКЛАДІ ПОШИРЕННЯ COVID-19

Нижче наведено результати аналізу ставлення суспільства до дій влади, спрямованих на посилення карантинних заходів у зв'язку з поширенням пандемії COVID-19. Ці результати отримано за допомогою розробленого застосунку «PRO ET CONTRA» на тестовому наборі даних системи Robusta, накопичених протягом 200 днів з найбільш популярних 11 соціальних медіа та 100 новинних українських вебсайтів (обсяг вибірки становить понад 2 млн. повідомлень).

Було сформовано такий запит мовою системи Robusta:

(\"ужесточен карантин\"|\\"посилен карантин\") (україн|украин),

який здійснює вибірку повідомлень, що містять у своєму тексті фрази «ужесточение карантина» або «посилення карантину» в різних відмінках і множині, а також слова, які починаються з «україн» або «украин» (регістр літер не має значення).

На рис. 3 наведено тематичну панель загальних результатів аналізу, з якої видно, що протягом проміжку часу з 12.03.2020 по 08.08.2020 вибрано 913 повідомлень. З них 579 — негативні, 294 — нейтральні, 40 — позитивні (це становить у частках 0.60, 0.32 та 0.04 відповідно). Співвідношення між цими повідомленнями відображено на векторній діаграмі (рис. 3). На основі цих даних розраховано оцінки норми вектора очікувань суспільства $|\vec{S}| = 0.672$ та $\alpha = 151^\circ$. На векторній діаграмі в центрі тематичної панелі (рис. 3) зображено співвідношення між векторами дій влади \vec{G} , соціальної активності (очікувань) суспільства \vec{S} та показником суспільних перетворень R .

Результати аналізу свідчать про те, що ставлення суспільства до дій влади, спрямованих на посилення карантинних заходів, є достатньо консолідованим (коефіцієнт конфліктності $C = 0.23$) та суттєво негативним (оцінка за шкалою Міллера — 2 бали). Це зумовлює спротив діям влади ($R = -0.82$) у формі ігнорування карантинних обмежень.

Як видно з рис. 3, маємо два піки погравлення соціальної активності суспільства. Перший пік (01.04.2020–10.04.2020) пов'язаний з уведенням надзвичайної ситуації та обмежувальних карантинних заходів. 25 березня було запроваджено надзвичайну ситуацію по всій території України; карантин продовжено до 24 квітня, закрито заклади торгівлі, крім продуктових магазинів, аптек, автозаправок і банків. Зупинено роботу метрополітенів у Києві, Харкові та Дніпрі, припинено міжміське та міжобласне автомобільне, залізничне та авіаційне сполучення.

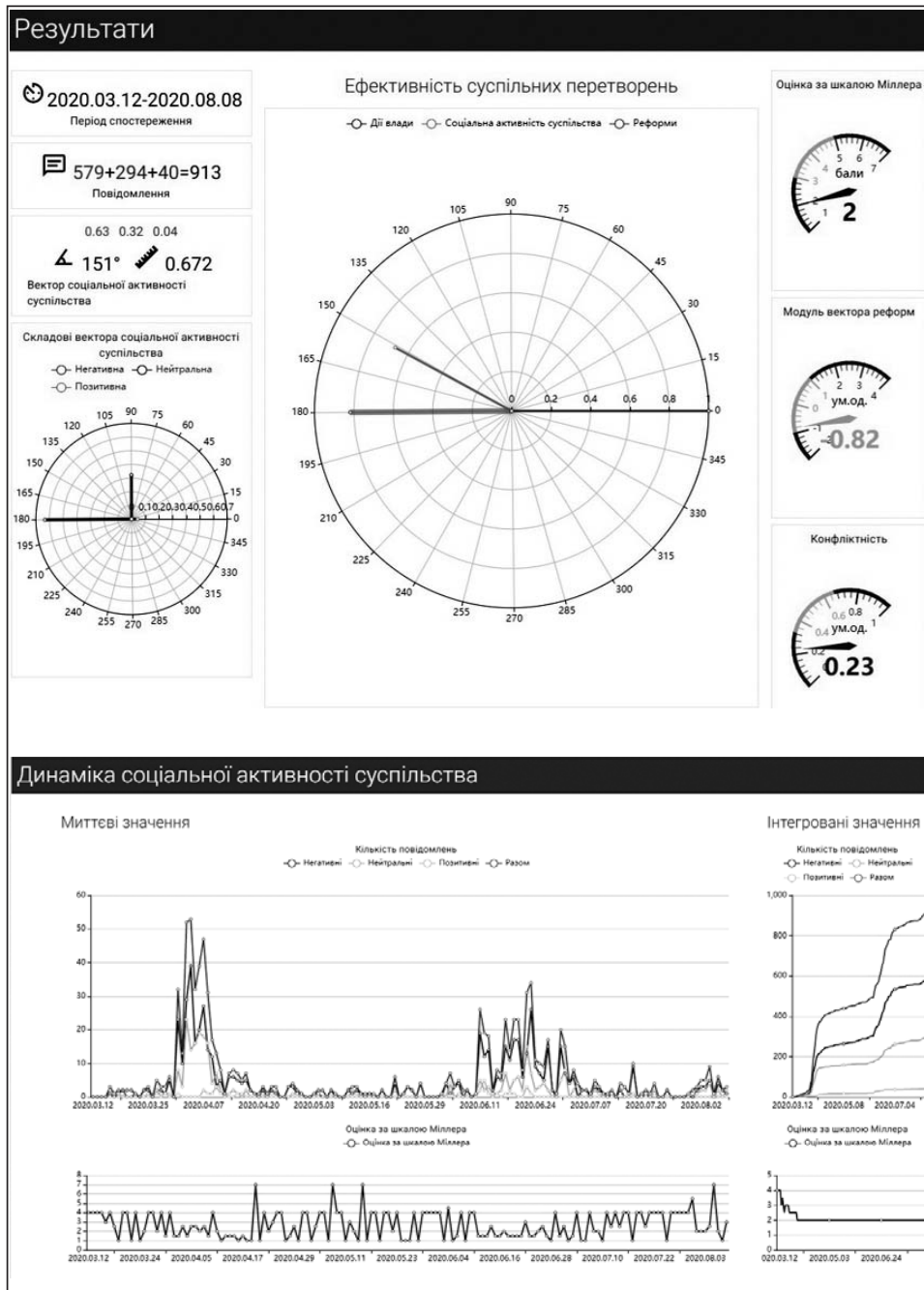


Рис. 3. Тематичні панелі застосунку «PRO ET CONTRA»

2 квітня уряд ухвалив постанову про посилення карантинних обмежень. 6 квітня в Україні було введено нові обмежувальні карантинні заходи: заборона на відвідування парків і зон відпочинку, у громадських місцях — обов’язкове носіння масок, особам, старшим 60 років, рекомендовано залишатися вдома.

Другий пік (11.06.2020–30.06.2020) пов’язаний з істотним погіршенням епідемічної ситуації в Україні та продовженням адаптивного карантину. 11 червня через істотне погіршення епідемічної ситуації в Україні Міністерство охорони здоров’я (МОЗ) вдається до посилення карантину. Зокрема, посилюються обме-

Експертне опитування щодо подій, які пов'язані з діями влади в період 2020.06.11 - 2020.06.22

Шановний експерте! Дайте оцінку приведеним нижче подіям за шкалою Міллера:

- 1 - критичний стан соціально-економічних характеристик суспільства, виникла революційний стан в країні;
- 2,3 - економічна стагнація, догнотривалий спад соціально-економічних характеристик суспільства;
- 4 - суспільство і влада діють незалежно один від одного, стан суспільної апатії;
- 5,6 - високі суспільні перетворення;
- 7 - максимальні суспільні перетворення.

11 червня через істотне погіршення епідемічної ситуації в Україні міністерство охорони здоров'я вдається до посилення карантину. Зокрема, посилюються обмеження на в'їзд в Україну іноземцям (необхідність мати страховий поліс).

ВАША ВІДПОВІДЬ

★ ★ ★ ☆ ☆ ☆ ☆ 2 (економічна стагнація, догнотривалий спад соціально-економічних характеристик суспільства)

12 червня з'явилася інформація про те що Уряд планує скасувати основне зовнішнє незалежне тестування (ЗНО) для вступників до закладів вищої освіти 2020 року. Пізніше МОЗ оприлюднило три сценарії проведення ЗНО у 2020 році: проведення у попередньо встановлені терміни з 25 червня до 17 липня, зміна дат проведення та повне його скасування.

ВАША ВІДПОВІДЬ

★ ★ ★ ☆ ☆ ☆ ☆ 3 (економічна стагнація, догнотривалий спад соціально-економічних характеристик суспільства)

Рис. 4. Фрагмент анкети експертного опитування

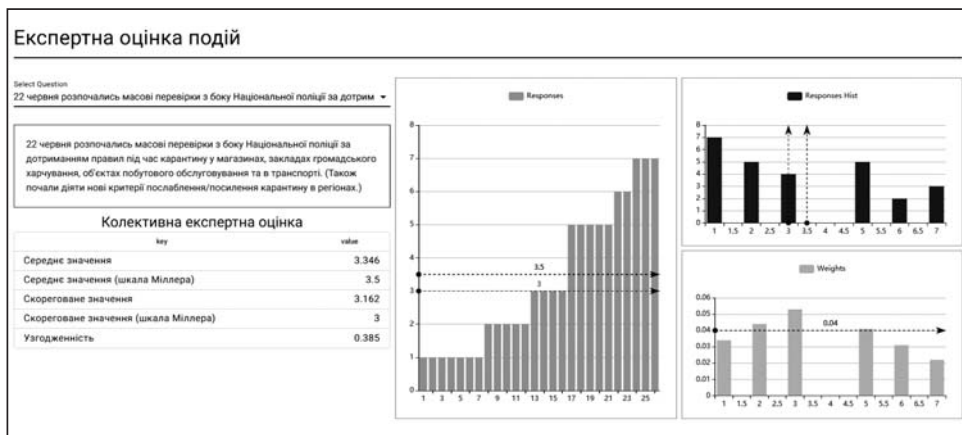


Рис. 5. Тематична панель для аналізу колективної експертної оцінки

ження на в'їзд в Україну іноземцям (обов'язкова наявність страхового полісу). 12 червня з'явилася інформація про те, що уряд планує скасувати основне зовнішнє незалежне тестування (ЗНО) для вступників до закладів вищої освіти 2020 року. Пізніше МОЗ оприлюднило три сценарії проведення ЗНО у 2020 році: у попередньо встановлені терміни з 25 червня до 17 липня, зміна дат проведення та повне його скасування. 17 червня уряд ухвалив рішення про продовження адаптивного карантину до 31 липня з оновленими правилами його дотримання. 22 червня розпочалися масові перевірки Національною поліцією України дотримання цих правил під час карантину в магазинах, закладах громадського харчування, об'єктах побутового обслуговування та у транспорті. Також почали діяти нові критерії послаблення/посилення карантину в регіонах.

У результаті аналізу дій влади протягом періоду другого піку активності було сформовано анкету з чотирьох запитань, фрагмент якої зображено на рис. 4.

Цю анкету було запропоновано експертній групі у складі 30 осіб, які дали індивідуальні експертні оцінки ставлення суспільства до дій влади. На основі цих оцінок було сформовано та проаналізовано колективну експертну оцінку за допомогою відповідної тематичної панелі (рис. 5)

Таблиця 2. Результати експертного опитування

Характеристика дій влади	Результат експертного оцінювання ставлення суспільства до дій влади (за шкалою Міллера)	Узгодженість експертних оцінок
11 червня через істотне погіршення епідемічної ситуації в Україні Міністерство охорони здоров'я вдається до посилення карантину. Зокрема, посилено обмеження на в'їзд в Україну іноземцям (обов'язкова наявність страхового полісу).	4	0.35
12 червня з'явилася інформація про те, що уряд планує скасувати основне зовнішнє незалежне тестування (ЗНО) для вступників до закладів вищої освіти 2020 року. Пізніше МОЗ оприлюднило три сценарії проведення ЗНО в 2020 році: у попередньо встановлені терміни з 25 червня до 17 липня, зміна дат проведення та повне його скасування.	3	0.49
17 червня уряд ухвалив рішення про продовження адаптивного карантину до 31 липня з оновленими правилами його дотримання.	3.5	0.37
22 червня Національна поліція розпочала масові перевірки дотримання правил під час карантину у магазинах, закладах громадського харчування, об'єктах побутового обслуговування та у транспорті. Також почали діяти нові критерії послаблення/посилення карантину в регіонах.	3	0.39
Анкета в цілому	3.34	0.40

Отримані результати експертного опитування зведено в табл.2.

Для того самого періоду кількісна оцінка за шкалою Міллера — 2, узгодженість індивідуальних експертних оцінок — 0.65.

Отже, експертна група порівняно з результатами аналізу повідомлень дала більш оптимістичні, проте менш узгоджені оцінки.

ВИСНОВКИ

1. Розроблено методику оцінювання ефективності суспільних перетворень, в основу якої покладено визначення міри неузгодженості дій влади та очікувань суспільства та синергії (соціальної активності) людей, залежно від зазначеної міри неузгодженості.

2. Запропоновано методи експертного та автоматичного онлайн-оцінювання масштабу суспільних перетворень на основі результатів експертних опитувань відповідно до методу Дельфі та результатів моніторингу відкритих онлайн-публікацій і соціальних мереж з використанням лінгвістичного аналізу великої кількості повідомлень з різним емоційним забарвленням.

3. На основі інтеграції різноманітних гетерогенних ресурсів та сервісів у межах єдиної ІТ інфраструктури Світового центру даних «Геоінформатика і сталий розвиток» розроблено спеціалізований застосунок «PRO ET CONTRA», що входить до складу комплексу програмних засобів Інформаційно-аналітичного ситуаційного центру «СІЦД-Україна» та призначений для експертного та автоматичного онлайн-оцінювання масштабу суспільних перетворень.

4. На прикладі кількісного оцінювання ставлення населення України до дій влади, спрямованих на посилення карантинних заходів у зв'язку з поширенням пандемії COVID-19, продемонстровано ефективність запропонованої методики та розроблених онлайн-інструментів.

СПИСОК ЛІТЕРАТУРИ

1. Zgurovsky M., Boldak A., Yefremov K. et al. Foresight and construction of the strategies of socio-economic development of Ukraine on mid-term (up to 2020) and long-term (up to 2030) time horizons. Scientific advisor of the project acad. of NAS of Ukraine M. Zgurovsky. International Council for Science etc. 2nd ed. Kyiv: NTUU «Igor Sikorsky KPI», Publ. house «Polytechnica», 2016. 184 p. URL: <http://wdc.org.ua/en/node/186284>.
2. Pozzi F.A., Fersini E., Messina E., Liu B. Sentiment analysis in social networks. Morgan Kaufmann, 2017. 284 p.
3. Додонов А.Г., Ландэ Д.В., Путятин В.Г. Компьютерные сети и аналитические исследования. Киев: ИПРИ НАН Украины, 2014. 486 с.
4. Ландэ Д.В. Основы интеграции информационных потоков. Киев: «Инжиниринг», 2006. 240 с.
5. Застосунок «PRO ET CONTRA». URL: <https://jace-dev.herokuapp.com/design/proetcontra#/>.
6. Онлайн-платформа Advanced Analytics СЦД-Україна. URL: <http://open.wdc.org.ua/#/>.
7. Korenko D., Boldak A. Approach to organization of client-server interaction for implementation of model-view-controller pattern in distributed systems. *Proc. International Conference on Security, Fault Tolerance, Intelligence (ICSFTI2019)* (14–15 May 2019, Kyiv, Ukraine). Kyiv: Igor Sikorsky Kyiv Polytechnic Institute, Publ. house «Polytechnica», 2019. P. 182–187.
8. Boldak A., Yefremov K. WDC-Ukraine's distributed platform for supporting research data life cycle. *Proc. 17th International Conference on System Fnalysis and Information Technologies (SAIT 2015)* (22–25 June 2015, Kyiv, Ukraine). Kyiv, 2015. P. 210–211.

Надійшла до редакції 02.11.2020

М.З. Згуровский, Д.В. Ланде, А.А. Болдак, К.В. Ефремов, М.Н. Перестюк **ЛИНГВИСТИЧЕСКИЙ АНАЛИЗ ДАННЫХ ИНТЕРНЕТ-МЕДИА И СОЦИАЛЬНЫХ СЕТЕЙ** **В ЗАДАЧАХ ОЦЕНИВАНИЯ ОБЩЕСТВЕННЫХ ПРЕОБРАЗОВАНИЙ**

Аннотация. Разработан комбинированный подход к оценке эффективности общественных преобразований как меры несогласованности между действиями власти и ожиданиями общества и синергии (социальной активности) людей, основанный на формализованном согласовании результатов, полученных методом экспертных оценок и методами сентимент-анализа и интеллектуального анализа текстовых сообщений из открытых онлайн-источников и социальных сетей. Эти методы реализованы в виде комплекса веб-сервисов и приложений в среде разработки интегрированной онлайн-платформы Advanced Analytics Мирового центра данных «Геоинформатика и устойчивое развитие». Эффективность предложенного подхода продемонстрирована на примере количественного оценивания отношения населения Украины к действиям власти, направленным на противодействие распространения эпидемии COVID-19.

Ключевые слова: вектор действий власти, вектор ожиданий общества, вектор преобразований (реформ), лингвистический анализ, контент-анализ, анализ данных Интернет медиа и социальных сетей, разведка на основе открытых источников.

M. Zgurovsky, D. Lande, A. Boldak, K. Yefremov, M. Perestyuk

**LINGUISTIC ANALYSIS OF INTERNET MEDIA AND SOCIAL NETWORKS DATA
IN PROBLEMS ON ASSESSMENT OF SOCIAL TRANSFORMATIONS**

Abstract. A combined approach has been developed to assess the effectiveness of social transformations as a measure of inconsistency between the actions of the authorities and the expectations of society and the synergy (social activity) of people, based on formalized coordination of the results obtained by the method of expert assessments and methods of sentiment analysis and intelligent analysis of text messages from open online sources and social networks. These methods are implemented as a set of web services and applications in the development environment of the Advanced Analytics integrated online platform of the World Data Center “Geoinformatics and Sustainable Development”. The effectiveness of the proposed approach is demonstrated by the example of a quantitative assessment of the attitude of the population of Ukraine to the actions of the authorities aimed at countering the spread of the COVID-19 epidemic.

Keywords: vector of government actions, vector of society’s expectations, vector of transformations (reforms), linguistic analysis, content analysis, linguistic sentiment analysis of Internet media data and social networks, open source intelligence.

Згуровський Михайло Захарович,

академік НАН України, доктор техн. наук, професор, ректор Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», e-mail: zgurovsm@hotmail.com.

Ланде Дмитро Володимирович,

доктор техн. наук, професор, завідувач відділу Інституту проблем реєстрації інформації НАН України, Київ, e-mail: dwlande@gmail.com.

Болдак Андрій Олександрович,

кандидат техн. наук, доцент, доцент кафедри Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», e-mail: boldak.andrey@gmail.com.

Єфремов Костянтин Вікторович,

директор Навчально-наукового комплексу «Світовий центр даних «Геоінформатика та сталий розвиток» Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», e-mail: k.yefremov@gmail.com.

Перестюк Марія Миколаївна,

кандидат фіз.-мат. наук, професор, директор Інституту моніторингу якості освіти Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», e-mail: maria@perestyuk.com.