

DOI: <http://dx.doi.org/10.30970/meu.2020.44.0.3409>

УДК [001.895+001.891.573]:[330.47+519.816+519.179.2]:[004.912+519.767.6];  
JEL C43,C83, C87, D83, O32

## ОЦІНКА АКТУАЛЬНОСТІ ПРОБЛЕМИ ЯКОСТІ СОЦІАЛЬНО-ЕКОНОМІЧНИХ ПРОГНОЗІВ ЧЕРЕЗ ПРИЗМУ ІНТЕРНЕТ-ПРОСТОРУ

Іван Твердохліб, Маркіян Тимчишин

Львівський національний університет імені Івана Франка  
79008 м. Львів, проспект Свободи, 18

e-mail: [ivan.tverdokhlib@lnu.edu.ua](mailto:ivan.tverdokhlib@lnu.edu.ua), [i\\_tverdok@i.ua](mailto:i_tverdok@i.ua), [mariktima@gmail.com](mailto:mariktima@gmail.com)

**Анотація.** В статті розглядається проблема використання систем контент-аналізу для обґрунтування доцільності тематики економічних досліджень. Поповнення інструментарію дослідника такими новітніми інформаційними технологіями є актуальним завданням з огляду на зростання ролі Інтернету в світі. З метою апробації доцільності моніторингу інформаційних джерел Інтернету в економічних дослідженнях вибрано проблему якості соціально-економічних прогнозів як об'єкт аналізу. Здійснено моніторинг Інтернет-простору системою контент-аналізу InfoStream на предмет обговорення Інтернет-спільнотою проблеми якості прогнозів. На підставі результатів моніторингу цієї науково-практичної програми у двох базах даних InfoStream (засоби масової інформації та форуми) розраховано щотижневий індекс актуальності проблеми якості прогнозів на протязі 2020 року. Для кількісного оцінювання актуальності проблеми якості прогнозів в Інтернет-просторі використано відомий репутаційний індекс Net Sentiment. Загалом отримані значення індексу Net Sentiment засвідчують стабільну позитивну тижневу актуальність проблеми якості прогнозів серед Інтернет-спільноти у 2020 році.

**Ключові слова:** інновації, прогнозування, якість прогнозів, актуальність проблеми, моніторинг Інтернет-простору, контент-аналіз, Text Mining, Infostream.

**Постановка проблеми.** У загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями. Стрімкий розвиток інформаційно-комунікаційних технологій кінця ХХ ст. – початку ХХІ ст. у поєднанні з глобалізаційними процесами призвів до все зростаючої тенденції оцифровування джерел інформації для всесвітньої мережі Інтернет. За твердженням дослідників глобальні комп'ютерні мережі містять величезну кількість інформації, за допомогою якої можна вирішувати найрізноманітніші завдання у різних сферах діяльності людини [1, 2]. Причому майже половина даних в Інтернет-просторі є безплатними у доступі [3, с. 42]. Перерахуємо деякі сфери діяльності, у яких систематично використовують джерела Інтернет-простору:

- конкурентна розвідка у теперішньому часі неможлива без використання інформаційних ресурсів Інтернет-простору [2, 3];
- дослідження структури та інформаційних потоків Інтернет-простору [1, 4, 5, 2];
- аналіз тематики і подій, що висвітлюються у ЗМІ журналістами [6, 7], у соціальних мережах [8, 9];
- відстеження репутації особи/організації, продуктів та PR-акцій [10-13];
- ідентифікація, протидія та проведення інформаційних операцій проти конкурентів [14, 2, 10];
- у IT-сфері під час тестування програмних додатків з використанням OSINT-технології [15].

Щоб забезпечити таке різноманіття завдань використання інформації з мережі Інтернет, структура інформаційного простору має бути не простою. Крім того, очевидно, потрібні спеціалізовані технології та інструменти. У контексті пошуку інформації в Інтернеті виділяють три структурні компоненти Інтернет-простору, а саме [2, с. 75-87]:

- видимий веб, тобто інформаційні джерела Інтернету, які скануються стандартними пошуковими машинами на кшталт Google;
- глибокий веб (invisible web, hidden web), що включає ще доступні джерела, які не скануються стандартними пошуковими машинами;
- темний веб (dark web) – сукупність наявних у мережі інформаційних ресурсів, які недоступні.

Причому кількісно глибокий веб набагато об'ємніший від видимого та містить якіснішу для користувача інформацію.

Власне привабливість глибокого вебу Інтернет-простору для користувачів і зумовила потребу розроблення спеціальних методів та інформаційних систем для пошуку у зазначеному сегменті інформаційного Інтернет-простору. До таких нових інструментів роботи з Інтернет відносять методи Text Mining інтелектуального аналізу даних та системи контент-аналізу для моніторингу Інтернет-простору [2, с. 26-32; 16, с. 58-75]. Однією з таких систем є InfoStream, розробником якої є українська компанія з м. Києва ТЗОВ “Інформаційний центр “Електронні вісті” [17, 18].

**Аналіз останніх досліджень і публікацій.** Поява такого нового і перспективного засобу дослідження явищ і процесів через призму Інтернет-простору привернула увагу і дослідників-економістів. В останні роки активно випробовуються можливості систем контент-аналізу для дослідження соціально-економічних процесів та напручується методологія їх використання у сфері економічної науки. Так у [19-22] оприлюднено результати розроблення прогностичних моделей з використанням моніторингу Інтернет-простору пошуковою машиною Яндекс. Ідентифікація доходів населення регіону на основі аналізу тональності їхніх пошукових запитів описана у [23]. Проблема побудови словників для аналізу контенту електронних публікацій в Інтернеті в економічних дослідженнях розглядалася у статтях [24-26, 7]. Можливості соціальних мереж у плані передбачення подій окреслено у працях [9, 8]. Методики використання моніторингу Інтернет-простору для прогнозування інфляції, демографії, економічних злочинів і кримінальної ситуації загалом описані у [20-22, 27, 28]. Ще одним напрямком застосування систем-контент-аналізу в економічних дослідженнях є оцінювання актуальності досліджуваної проблеми. Зокрема, у [29, 30] оприлюднено оцінки привабливості проблеми інформатизації Львівщини для населення на основі моніторингу ЗМІ в Інтернеті за допомогою системи InfoStream., а у [31] – відстежено динаміку кількості публікацій з проблематики інфляції у Інтернет-виданнях.

Актуальність прогнозів для управління економікою обґрунтував відомий російський економіст М.Д. Кондратьєв ще майже 100 років тому, аналізуючи методологію розробки планів у СРСР. Він пов'язав якість планування з передбаченням ходу протікання соціально-економічного явища та оцінкою ефекту, що очікується від запланованих рішень, вперше звернув увагу на складність проблеми якості прогнозів [32]. Починаючи з 2-ої половини ХХ ст. фахівці періодично обговорювали проблему якості передбачень у сфері економіки і підприємництва з метою виявлення причин недостовірності соціально-економічних прогнозів та удосконалення методології їх розроблення (див. аналіз у [33]). І в наш час залежність управлінських рішень від прогнозів є визначальною. Приміром здійснений у [34] аналіз впливу прогнозування на обґрунтування бюджетних показників урядами США засвідчив їх добру узгодженість з відхиленням до 1%. Натомість неякісний макроекономічний прогноз призвів до значного недонадходження державного бюджету України у 2019 році [35]. З метою збору та консолідації оцінок точності здійснених соціально-економічних прогнозів у [36, 37] обґрунтовано доцільність спеціалізованого інформаційного web-ресурсу підтримки рішень у сфері прогнозування.

**Постановка завдання** (формулювання цілей статті). Неодмінною передумовою будь-якого наукового дослідження є усвідомлення його актуальності. Зазичай така характеристика тематики дослідження обґрунтовується на основі аналізу стану її вирішення через призму фахових публікацій. Можливості сучасних комп'ютерних мереж уможливають залучення до обговорення проблем і інших верств населення, що дає змогу оцінювати суспільну значущість тематики наукового дослідження. Отже, з урахуванням зазначених вище новітніх тенденцій інформатизації наукової сфери **метою дослідження** була апробація методології застосування систем контент-аналізу для визначення значущості проблеми якості прогнозів для соціуму. Власне проблема якості соціально-економічних прогнозів, яка наразі ще немає задовільного вирішення, була вибрана як об'єкт інтелектуального аналізу методами Text Mining існуючих систем моніторингу Інтернет-простору. З цією метою була оцінена динаміка зміни актуальності проблеми якості прогнозування серед соціуму через призму її обговорення в Інтернет-просторі.

**Виклад основного матеріалу дослідження** з обґрунтуванням отриманих наукових результатів. Моніторинг Інтернет-простору здійснювався за допомогою професійної системи контент-аналізу InfoStream [17, 18], яка використовується у навчальному процесі ЛНУ ім. Івана Франка як Інтернет-сервіс у режимі online згідно місячного пакета послуг BizOnline [17]. Уточнена методика аналізу згадувань проблеми в Інтернет-джерелах за допомогою InfoStream описана у статті [29, с. 596-598]. Саме її і було використано для оцінювання актуальності проблеми якості прогнозування через призму Інтернет-простору.

Оскільки згідно пакета послуг BizOnline ретроспектива пошуку у БД системи InfoStream обмежена 6-ма місяцями від дати моніторингу, то нами було здійснено два точкові моніторинги:

- перший у липні 2020 р., який охоплював проміжок часу від січня 2020 р. по липень 2020 р. Основний пошуковий запит мовою InfoRes [18] був сформульований як фраза "якість прогнозування" тільки українською мовою. Додаткові опції запиту – пошук в головній БД (ЗМІ), усунення дублів, врахування морфології тексту запиту;
- другий у кінці листопада 2020 р., який охоплював період травень 2020 р. – листопад 2020 р. Пошуковий запит 1-го моніторингу був видозмінений і став таким -

“(якість прогнозування) | (якість прогнозування)”. Додаткові опції залишилися такими самими.

Обидва запити отримали непорожні множини релевантних документів. Приміром, система для 2-го запиту знайшла у своїй БД ЗМІ 4680 документів, які за змістом релевантні тексту запити.

Обидва набори релевантних документів аналізувалися за допомогою операцій консолідації та візуалізації їх характеристик у контексті проблеми якості прогнозів. Зокрема, була відстежена динаміка появи в Інтернет документів релевантних наборів в розрізі днів, тижнів і місяців з урахуванням їх тональності щодо проблеми якості прогнозів, тематика сюжетів (інформаційних тематичних потоків). Наприклад, динаміка згадувань проблеми якості прогнозів по днях для 2-го запити як копія екрану InfoStream зображена на рис. 1. Як бачимо, цій проблематиці переважно присвячено по декілька десятків публікацій на день, що уже засвідчує певну актуальність проблеми для соціуму.

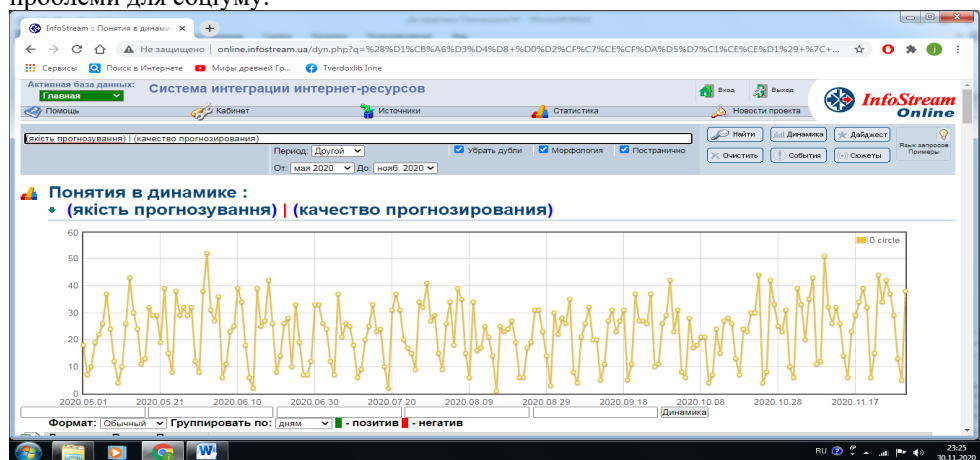


Рис. 1. Динаміка появи документів у ЗМІ з проблематики якості прогнозів на проміжку травень 2020 р. – листопад 2020 р. у розрізі днів (розробка авторів)

Щоб порівняти динаміку появи у Інтернет згадувань проблеми якості прогнозування для обох запитів, консолідуємо денні оцінки обсягів релевантних до кожного запити документів і візуалізуємо на відповідних часових відтинках за допомогою системи тижневі агреговані дані щодо кількості документів для запитів (див. рис. 2, 3). Просте порівняння графіків на цих малюнках засвідчує переважання у декілька разів кількості релевантних документів у тиждень для 2-го запити, аніж для 1-го. Причина криється у тому, що текст 2-го запити включає концепт пошуку ще російською мовою. Особливістю системи InfoStream є орієнтація її БД на пострадянський простір. Російська мова ще є переважаною для спілкування в ньому. Проблема якості прогнозування як інтелектуальна є привабливою для російськомовних ЗМІ.

Для кількісного оцінювання рівня актуальності проблеми якості прогнозування у Інтернет-просторі потрібно ще аналізувати тональність релевантних запитам документів. Зазвичай системи контент-аналізу за допомогою методів Text Mining класифікують зміст документів у контексті запити як позитивний, або негативний, або нейтральний.

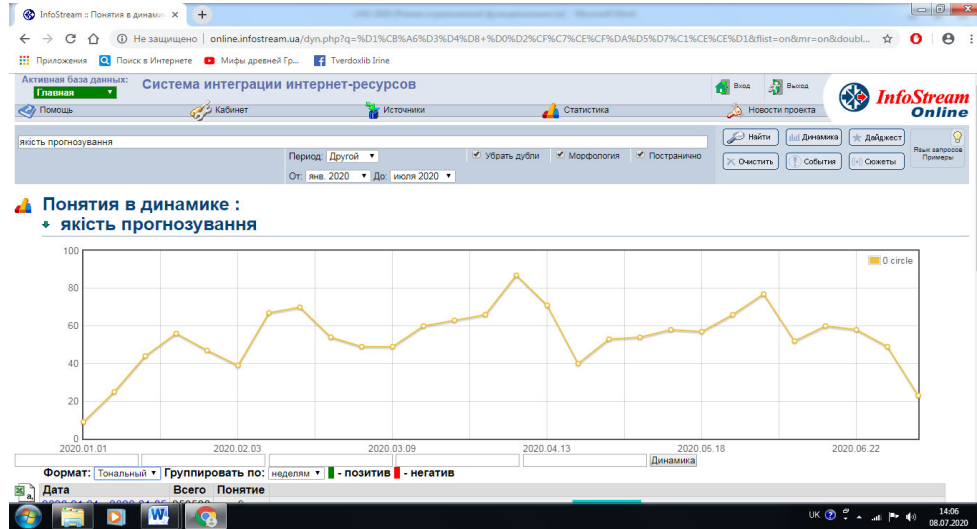


Рис. 2. Динаміка згадувань проблеми якості прогнозування у ЗМІ на проміжку часу січень 2020 р. – липень 2020 р. у розрізі тижнів (розробка авторів)

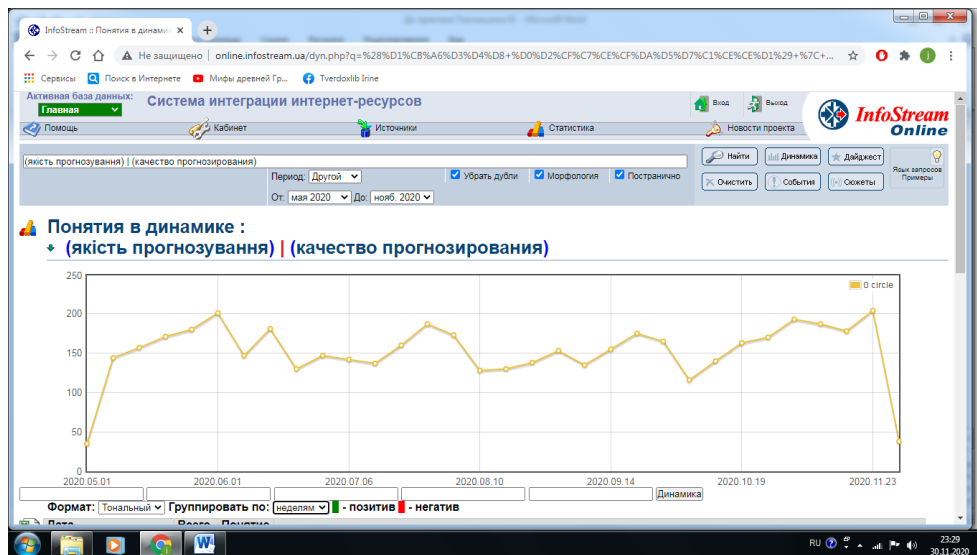


Рис. 3. Динаміка згадувань проблеми якості прогнозування у ЗМІ на проміжку часу травень 2020 р. – листопад 2020 р. у розрізі тижнів (розробка авторів)

Маючи оцінки кількості документів у вказаних групах за період часу можна розрахувати величину актуальності відповідної запити проблеми. На практиці використовують різні міри для такого оцінювання. Нами використано відомий мережевий репутаційний індекс Net Sentiment [38]. Хоча він використовується найчастіше для оцінювання репутації організацій чи осіб, але відомі приклади його застосування і для оцінки актуальності проблем [29, 30]. Значення індексу Net

Sentiment обчислюється за формулою [38]  $I(t) = [k_1(t) + k_2(t) - k_3(t)] / K(t)$ , де  $I(t)$  є кількісною оцінкою репутації у період  $t$ ,  $K(t)$  позначає кількість релевантних запиту (проблеми) документів у період  $t$ , а змінні  $k_1(t), k_2(t), k_3(t)$  задають кількості документів у період  $t$  відповідно позитивних, нейтральних і негативних. Очевидно, що  $I(t) \in [-1; +1]$ . Тобто  $I(t) = -1$  за умови негативної тональності усіх віднайдених документів,  $I(t) = 1$ , якщо документи негативного характеру відсутні, а  $I(t) = 0$  коли кількості позитивних та негативних документів однакові.

Отже, вибравши опції тональності та часовий період для операції Динаміка системи InfoStream, отримали агреговані дані щодо кількості релевантних документів у розрізі їх тональностей. На основі цих даних і було розраховано індекси актуальності проблеми якості прогнозування у розрізі тижнів та місяців. Оцінки актуальності зазначеної проблеми на проміжку часу січень – липень 2020 р. у розрізі місяців наведено у табл. 1. У 2-й колонці цієї таблиці зазначено кількість документів у БД ЗМІ системи контент-аналізу InfoStream, серед яких здійснювався пошук у відповідний місяць. Третя графа показує обсяг вибірки для 1-го запиту у відповідний період часу, а в наступних стовпцях зазначено розподіл вибірки за тональністю змісту документів. Характерним є переважаюча кількість документів нейтрального змісту щодо проблеми якості прогнозування та наявність негативних публікацій для кожного місяця 2020 р. В останній графі наведено розраховані значення індексу Net Sentiment, які і будуть оцінками актуальності проблеми якості прогнозування в Інтернет-просторі. Оскільки на часовому проміжку січень – липень 2020 р. величини  $I(t)$  знаходяться у межах інтервалу  $[0,8095238; 0,9437229]$ , то можна стверджувати досить значний позитивний рівень актуальності проблеми якості прогнозування серед україномовних публікацій у ЗМІ.

Таблиця 1

Динаміка зміни індексу актуальності проблеми якості прогнозування у розрізі місяців 2020 року

Місяць року, $t$	Обсяг БД в документах	Обсяг вибірки, документів, $K(t)$	з них			Індекс проблеми, $I(t)$
			позитив, $k_1(t)$	нейтральні, $k_2(t)$	негатив, $k_3(t)$	
2020.01	2579065	178	37	130	11	0,9382022
2020.02	2653973	231	46	172	13	0,9437229
2020.03	3014939	242	31	166	45	0,8140496
2020.04	2988778	286	44	207	35	0,8776224
2020.05	2854640	247	50	179	18	0,9271255
2020.06	2856867	257	43	188	26	0,8988327
2020.07	739398	63	13	38	12	0,8095238

Джерело: розраховано і сформовано авторами на основі звітів системи InfoStream

З метою прискіпливішого відстеження зміни актуальності досліджуваної проблеми в Інтернет-просторі здійснили консолідацію документів релевантних наборів для обох запитів у розрізі тональності їх змісту та тижнів року. Динаміка зміни

індексу актуальності проблеми якості прогнозування у розрізі тижнів року показана на рис. 4 (для 1-го запиту) та на рис. 5 для другого.

Як випливає з рис. 4, на часовому проміжку січень – липень 2020 р. величина тижневого індексу актуальності проблеми в україномовних ЗМІ коливається у межах від 0,6 до 1,0. Крім того, спостерігається значний спад зацікавленості цією проблемою у 3-й тиждень березня 2020 р. (така собі воронка відгуків за аналогією з маркетингу [39]).

На часовому інтервалі травень – листопад 2020 р. тижневий індекс актуальності проблеми більш стабільний і незначно коливається у межах від 0,8 до 0,91304348. Основна причина – другий запит охоплює і російськомовні джерела Інтернету, що збільшує обсяги вибірок релевантних проблемі документів.

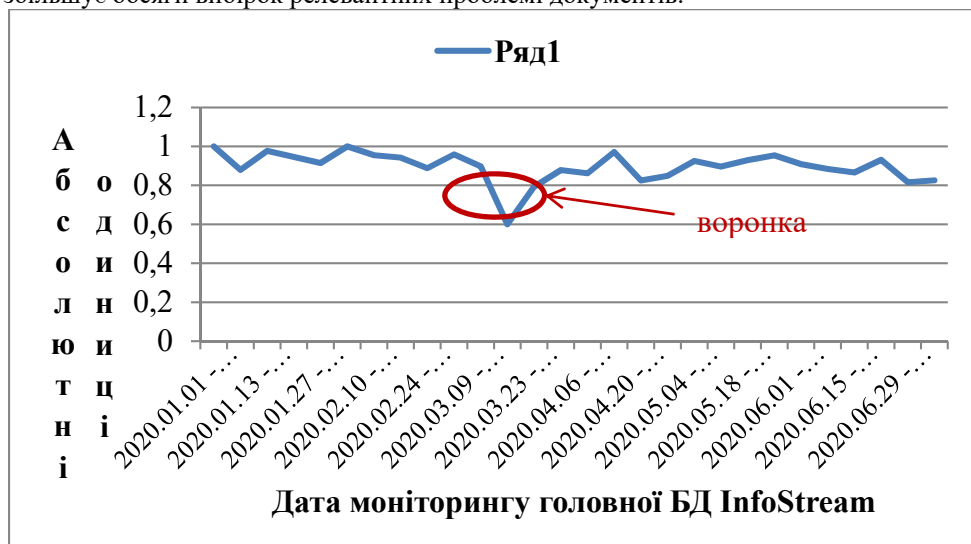


Рис. 4. Динаміка зміни актуальності проблеми якості прогнозування в україномовних ЗМІ у розрізі тижнів на основі результатів виконання InfoStream першого запиту (розробка авторів)

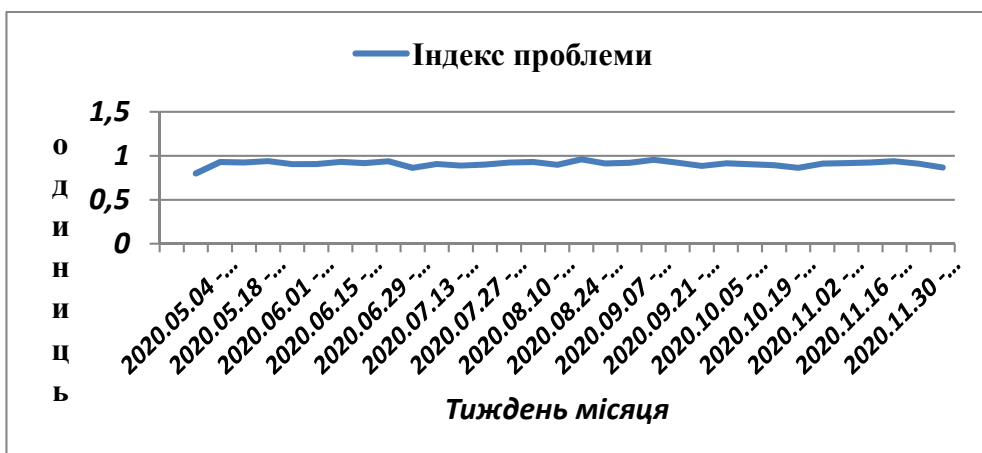


Рис. 5. Динаміка зміни актуальності проблеми якості прогнозування в ЗМІ у розрізі тижнів на основі результатів виконання InfoStream другого запиту (розробка авторів)

**Висновки та перспективи подальших досліджень.** Проведене дослідження актуальності проблеми якості прогнозування в Інтернет-просторі за допомогою системи контент-аналізу InfoStream засвідчило наявність певного постійного рівня зацікавленості цією проблемою Інтернет-спільноти. Це певний сигнал для науковців про доцільність дослідження аспектів ефективності прогнозів. Щодо доцільності використання систем контент-аналізу у економічних дослідженнях, то з одного боку проаналізовані зразки їх застосувань засвідчують деякі переваги під час вивчення об'єктів і процесів, а з другого – все залежить від проблеми дослідження. Ефективність прогнозів є складною науковою проблемою, результати її досліджень як правило оприлюднюються у спеціалізованих інформаційних джерелах, які у більшості не входять у БД системи InfoStream. Тому вибірки релевантних документів часто не містять спеціалізованих документів з такої проблематики і їх сюжети не повністю окреслюють сучасний стан вивчення проблеми. Проте затребуваність проблем для соціуму такі системи досить адекватно відстежують.

Ще однією перевагою застосування систем контент-аналізу у економічних дослідженнях є менший обсяг релевантних запитів документів Інтернет-простору та краща якість реферативного опису змісту документа. Приміром, Google Chrome на пошуковий запит “якість прогнозування” зазвичай видасть приблизно 49 900 000 посилань на релевантні документи з дуже нечіткими описами їх змісту у контексті запиту. Зрозуміло, що такий обсяг вибірки повністю дослідником не буде опрацьовано, а відбір потрібних вимагатиме затрат часу. Власне зменшення затрат часу фахівців на пошук актуальних джерел інформації з проблематики дослідження в Інтернет-просторі стає тим вагомим чинником, який і окреслює доцільність використання такого інструментарію у економічній науці.

Аналіз напрямків застосування систем контент-аналізу у економічних дослідженнях засвідчує можливість врахування з їх допомогою нечітких людиноорієнтованих параметрів проблем, які вкрай важко оцінювати у моделях. Саме така їх властивість і визначає основний вектор зацікавленості серед дослідників, особливо молодих.

#### Список використаних джерел

1. Ланде Д. Аналіз інформаційних потоків у глобальних комп'ютерних мережах (за матеріалами наукової доповіді на засіданні Президії НАН України 25 січня 2017 р.) *Вісник НАН України*. 2017. № 3. С. 46–54.
2. Додонов А.Г., Ландэ Д.В., Прищепа В.В., Путятин В.Г. Конкурентная разведка в компьютерных сетях. Киев : ИПРИ НАН Украины, 2013. 250 с.
3. Ландэ Д., Прищепа В. Школа веб-разведки. *Телеком*. 2007. №6. С. 40-45.
4. Ландэ Д. Динамика информационных потоков и информационные операции. *Информационные технологии для менеджмента*. 2010. № 10. С. 22-26.
5. Ландэ Д., Снарский А. Попытки объять необъятное, или World Wide Web под прицелом. *Сети и бизнес*. 2007. №4 (35). С. 18-24.
6. Городниченко Ю., Талавера О., Сюй Х. Говорить президент: як відрізняються виступи Порошенка в Україні та за кордоном. *Європейська правда*. 16.08.2016. URL: <http://www.eurointegration.com.ua/articles/2016/08/16/7053438/>
7. Павлишенко Б.М. Використання методів машинного навчання та семантичних ознак в інтелектуальному аналізі текстових даних. *Електроніка та інформаційні технології*. 2020. Вип. 13. С. 3–18.
8. Bohdan M. Pavlyshenko. Forecasting of events by tweets data mining. *Electronics and information technologies*. 2018. Issue 10. P. 71–85.



9. Bohdan M. Pavlyshenko. Can twitter predict royal baby's name? *Electronics and information technologies*. 2019. Issue 11. P. 52–60.
10. Рябий М., Хатян О., Багацький С. Модель виявлення PR-впливу через публікації в Інтернет ЗМІ. *Ukrainian Scientific Journal of Information Security*. 2015. Vol. 21. Issue 2. p. 131-139.
11. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Модели репутации и информационного управления в социальных сетях. *Математическая теория игр и ее приложения*. 2009. Т. 1. Вып. 2. С. 14-37.
12. Wobot. URL: <http://www.wobot.ru>
13. Мегафон: будущее зависит от тебя. Анализ информационного поля компании и конкурентов в СМИ. URL: <https://www.mlg.ru/products/media>
14. Горбулін В.П., Додонов О.Г., Ланде Д.В. Інформаційні операції та безпека суспільства: загрози, протидія, моделювання: монографія. Київ : Інтертехнологія, 2009. 164 с.
15. Розвідка на основі відкритих джерел. *Вікіпедія: вільна енциклопедія*. URL: [https://en.wikipedia.org/wiki/Open-source\\_intelligence](https://en.wikipedia.org/wiki/Open-source_intelligence)
16. Ландэ Д.В., Снарский А.А., Безсуднов Й.В. Интернетика: Навигация в сложных сетях: модели и алгоритмы. Москва : Книжный дом «ЛИБРОКОМ», 2009. 264 с.
17. InfoStream® - моніторинг ЗМІ та соцмереж. URL: <http://infostream.ua/UKR/>
18. InfoStream. Мониторинг новостей из Интернет: технология, система, сервис: научно-методическое пособие. /А.Н. Григорьев та ін. - Киев, ООО “Старт 98”, 2007. 40 с.
19. Болдырева А. Построение прогнозных моделей экономической и социальной конъюнктуры по интенсивности запросов в поисковой сети Интернет. *Современная экономика: теория, политика, инновации, сборник студенческих научных работ*, Москва, РАНХиГС. 2016. С. 36–61.
20. Boldyreva A., Koshulko O. GMDH Helps to Build Models Based on Queries to Yandex for Forecast of Economic Crimes. Proc. of the 7th International Workshop on Inductive Modeling, NASU, Praga Tech. Univ. 2015. pp. 9–11.
21. Boldyreva, A. Demographic Forecasts Based on Queries to Yandex Search Machine. Proc. of 7th International Workshop on Inductive Modeling, NASU, Praga Tech. Univ. 2015. pp. 7–8.
22. Boldyreva A., Alexandrov M., Koshulko O., Sobolevskiy O. Queries to Internet as a tool for analysis of regional police work and forecast of crimes in regions. Proc. of 15th Mexican Intern. Conf. on Artificial Intelligence, Springer, LNAI. vol. 10061. 2016. pp. 291-302.
23. Болдырева А., Александров М., Суркова Д. Негативно окрашенные слова в поисковых запросах Интернета как индикатор уровня среднедушевых доходов населения в федеральных округах РФ. *Сб. трудов НАН Украины «Индуктивное моделирование сложных систем»*. 2015. Вып. 7. С. 77–92.
24. Boldyreva A., Sobolevskiy O., Alexandrov M., Danilova V. Creating collections of descriptors based on Internet queries. Proc. of 15th Mexican Intern. Conf. on Artificial Intelligence, Springer, LNAI. Vol. 10061. 2016. pp. 303-314.
25. Береснева Д., Александров М. Динамические словари для изучения Интернет новостей. *Сб. трудов “Математическое моделирование социальных процессов”*. Москва : Соц. Фак. МГУ, 2015. Вып.17. С. 27-33.
26. O. Kushnir, V. Buryi, S. Grydzhan, L. Ivanitskyi, S. Rykhlyuk. Zipf’s and heaps’ laws for the natural and some related random texts. *Electronics and information technologies*. 2018. Issue 9. P. 94–105.
27. Ейнес Ю. Рекомендаційні техніки для аналізу криміналістичних даних. *Електроніка та інформаційні технології*. 2018. Вип. 10. С. 64–70.
28. Болдырева А. Интегральный метод исследования настроений пользователей Интернета с помощью поисковых запросов. *Математическое моделирование и*

- інформатика социальных процессов, сборник трудов МГУ, социологический факультет. 2016. Вып. 18. С. 26–34.
29. Твердохліб І.П., Блонський Н.А., Костюк Д.В. Емпірична оцінка доцільності моніторингу інформаційного простору Інтернет в економічних дослідженнях. *Економіка та суспільство: Електронне наукове фахове видання*. Вип. 11. Мукачів, 2017. С. 593–602. URL: <http://www.economyandsociety.in>.
  30. Твердохліб І.П., Костюк Д.В. Моніторинг стану інформатизації Львівщини в Інтернет-просторі за допомогою систем контент-аналізу. *"Formation of Knowledge Economy as the Basis for Information Society"* [Thesis of the 7th International Scientific Seminar] / by I.Y.Shtuler, D.V. Solokha, A.A. Oleshko, Z.B.Zhyvko etc. Kyiv-Venice-Verona: IAI-S-NAM, 2018. P. 42-47.
  31. Береснева Д. Моделирование динамики числа публикаций по экономической тематике в интернет-изданиях. *Сб. трудов студентов РАНХиГС «Современная экономика: теория, политика, инновации»*. Москва: РАНХиГС, 2016. С. 174-181.
  32. Кондрагьев Н.Д. План и предвидение (к вопросу о методах составления перспективных планов развития народного хозяйства и сельского хозяйства в частности). *Проблемы экономической динамики*. Москва: Экономика, 1989. С.91-134.
  33. Твердохліб І., Петрик Г. Ефективність економіко-математичного моделювання процесів трансформаційної економіки: прагматичний аспект. *Вісник Львівського університету. Серія економічна*. 2005. Вип. 34. С. 227-241.
  34. Гурвич Е. Насколько точны макроэкономические и бюджетные прогнозы? *Вопросы экономики*. 2006. №9. С. 4-20.
  35. Пацкан В. Через низьку якість макроекономічного прогнозування недо-надходження державного бюджету у 2019 році сягнули 43 млрд. грн. 2019. URL: <https://rp.gov.ua/PressCenter/News/?id=860> (дата звертання: 26.12.2020)
  36. Твердохліб І. Концепція інформаційного Web-ресурсу підтримки рішень у сфері прогнозування. *Інформація, комунікація, суспільство 2014: матеріали 3-ї міжнар. наук. конф. ICS-2014*. Львів: Видавництво Львівської політехніки, 2014. С. 28-29.
  37. Твердохліб І.П. Сервіс-орієнтована архітектура інформаційного web-ресурсу підтримки рішень у сфері прогнозування. *«Сучасні проблеми інформатики в управлінні, економіці та освіті»* : [матеріали XIII міжнар. наук. семінару] / за наук. ред. д.е.н., проф. М. М. Єрмошенка. Київ : Національна академія управління, 2014. С. 65-70.
  38. Мониторинг СМИ от Медиалогии. Мониторинг прессы. URL: <http://www.mlg.ru/solutions/pr/monitoring/>
  39. Воронка продажів: що це таке і як її побудувати? *TurboSMS*. URL: <https://turbosms.ua/ua/articles/voronka-prodazh-hto-eto-takoe.html>

#### References

1. Lande D. (2017). Analiz informatsiinykh potokiv u hlobalnykh kompiuternykh mrezhakh (za materialamy naukovoi dopovidi na zasidanni Prezydii NAN Ukrainy 25 sichnia 2017 r.) [Analysis of information flows in global computer networks (based on the scientific report at the meeting of the Presidium of the NAS of Ukraine on January 25, 2017)]. *Visnyk of the National Academy of Sciences of Ukraine*, no. 3, pp. 46–54. (in Ukrainian)
2. Dodonov A.G., Lande D.V., Prishchepa V.V., Putyatin V.G. (2013) Konkurentnaya razvedka v komp'yuternykh setyakh [Competitive intelligence in computer networks]. Kiev : IPRI NAN Ukrainy. (in Russian)
3. Lande D., Prishchepa V. (2007). Shkola veb-razvedki [School of Web Intelligence]. *Telekom*, no. 6, pp. 40–45. (in Russian)
4. Lande D. (2010). Dinamika informatsionnykh potokov i informatsionnye operatsii [Dynamics of information flows and information operations]. *Information Technology for Management*, no. 10, pp. 22-26. (in Russian)

5. Lande D., Snarskiy A. (2007). Popytki ob"yat' neob"yatnoe, ili World Wide Web pod pritselom [Attempts to grasp the immensity, or the World Wide Web at gunpoint]. *Networks and business*, no. 4 (35), pp. 18-24.
6. Horodnychenko Yu., Talavera O., Siui Kh. (2016). Hovoryt prezident: yak vidrizniaiutsia vystupy Poroshenka v Ukraini ta za kordonom [The president says: how different are Poroshenko's performances in Ukraine and abroad]. *Yevropeiska pravda*. 16.08.2016. URL: <http://www.eurointegration.com.ua/articles/2016/08/16/7053438/>
7. Pavlyshenko B.M. (2020). Vykorystannia metodiv mashynnoho navchannia ta semantychnykh oznak v in-telektualnomu analizi tekstovykh danykh [The use of machine learning methods and semantic features in the intellectual analysis of text data]. *Electronics and information technologies*, issue 13, pp. 3–18. (in Ukrainian)
8. Bohdan M. Pavlyshenko. (2018). Forecasting of events by tweets data mining. *Electronics and information technologies*, issue 10, pp. 71–85.
9. Bohdan M. Pavlyshenko. (2019). Can twitter predict royal baby's name? *Electronics and information technologies*, issue 11, pp. 52–60.
10. Riabiy M., Khatian O., Bahatskiy S. (2015). Model vyavleniia PR-vplyvu cherez publikatsii v Internet ZMI [Model of PR-influence detection through publications in Internet mass media]. *Ukrainian Scientific Journal of Information Security*, vol. 21, issue 2, pp. 131-139. (in Ukrainian)
11. Gubanov D.A., Novikov D.A., Chkhartishvili A.G. (2009). Modeli reputatsii i informatsionnogo upravleniia v sotsial'nykh setyakh [Reputation and information management models in social networks]. *Mathematical game theory and its applications*, vol. 1, issue 2, pp. 14-37. (in Russian)
12. Wobot. URL: <http://www.wobot.ru>
13. Megafon: budushchee zavisit ot tebya. Analiz informatsionnogo polya kompanii i konkurentov v SMI [Megaphone: the future depends on you. Analysis of the information field of the company and competitors in the media]. URL: <https://www.mlg.ru/products/media/>
14. Horbulin V.P., Dodonov O.H., Lande D.V. (2009). Informatsiini operatsii ta bezpeka suspilstva: zahrozy, protydiia, modeliuvannia: monohrafiia [Information operations and security of society: threats, counteraction, modeling: monograph]. Kyiv : Intertekhnolohiia. (in Ukrainian)
15. Rozvidka na osnovi vidkrytykh dzherel [Exploration based on open sources]. *Wikipedia: free encyclopedia*. URL: [https://en.wikipedia.org/wiki/Open-source\\_intelligence](https://en.wikipedia.org/wiki/Open-source_intelligence)
16. Lande D.V., Snarskiy A.A., Bezsudnov I.V. (2009). Internetika: Navigatsiia v slozhnykh setyakh: modeli i algoritmy [Internetics: Navigating Complex Networks: Models and Algorithms]. Moscow : Knizhnyy dom «LIBROKOM». . (in Russian)
17. InfoStream® - monitorynh ZMI ta sotsmerezh [InfoStream® - monitoring of mass media and social networks]. URL: <http://infostream.ua/UKR/>
18. Grigor'ev A.N., Lande D.V., Borodnikov S.A., Mazurkevich R.V., Pats'ora V.N. (2007). InfoStream. Monitoring novostey iz Internet: tekhnologiya, sistema, servis: nauchno-metodicheskoe posobie [Monitoring news from the Internet: technology, system, service: scientific and methodological manual]. Kiev, OOO "Start 98". (in Russian)
19. Boldyreva A. (2016). Postroenie prognoznykh modeley ekonomicheskoy i sotsial'noy kon"yunktury po intensivnosti zaprosov v poiskovoy seti Internet [Construction of predictive models of economic and social conditions based on the intensity of queries in the Internet search network]. *Modern economics: theory, politics, innovation, collection of student research papers*, Moskva, RANKhiGS,. pp. 36–61. (in Russian)
20. Boldyreva A., Koshulko O. (2015). GMDH Helps to Build Models Based on Queries to Yandex for Forecast of Economic Crimes. Proc. of the 7th International Workshop on Inductive Modeling, NASU, Praga Tech. Univ., pp. 9–11.

21. Boldyreva, A. (2015). Demographic Forecasts Based on Queries to Yandex Search Machine. Proc. of 7th International Workshop on Inductive Modeling, NASU, Praga Tech. Univ., pp. 7–8.
22. Boldyreva A., Alexandrov M., Koshulko O., Sobolevskiy O. (2016). Queries to Internet as a tool for analysis of regional police work and forecast of crimes in regions. Proc. of 15th Mexican Intern. Conf. on Artificial Intelligence, Springer, LNAI. vol. 10061, pp. 291-302.
23. Boldyreva A., Aleksandrov M., Surkova D. (2015). Negativno okrashennye slova v poiskovykh zaprosakh Interneta kak indikator urovnya srednedushevnykh dokhodov naseleniya v federal'nykh okrugakh RF [Negatively colored words in Internet search queries as an indicator of the level of per capita income of the population in the federal districts of the Russian Federation]. *Sat. Proceedings of the National Academy of Sciences of Ukraine "Inductive Modeling of Complex Systems"*, issue 7, pp. 77–92. (in Russian)
24. Boldyreva A., Sobolevskiy O., Alexandrov M., Danilova V. (2016). Creating collections of descriptors based on Internet queries. Proc. of 15th Mexican Intern. Conf. on Artificial Intelligence, Springer, LNAI. Vol. 10061, pp. 303-314.
25. Beresneva D., Aleksandrov M. (2015). Dinamicheskie slovari dlya izucheniya Internet novostey [Dynamic dictionaries for learning Internet news]. *Sat. works «Mathematical modeling of social processes* Moscow : Sots. Fak. MGU, issue 17, pp. 27-33. (in Russian)
26. O. Kushnir, V. Buryi, S. Grydzhan, L. Ivanitskyi, S. Rykhlyuk (2018). Zipf's and heaps' laws for the natural and some related random texts. *Electronics and information technologies*, issue 9, pp. 94–105.
27. Eines Yu. (2018). Rekomendatsiini tekhniki dlia analizu kryminalistychnykh danykh. *Electronics and information technologies*, issue 10, pp. 64–70. (in Ukrainian)
28. Boldyreva A. (2016). Integral'nyy metod issledovaniya nastroeniy pol'zovateley Interneta s pomoshch'yu poiskovykh zaprosov [An integral method for researching the mood of Internet users using search queries]. *Mathematical modeling and informatics of social processes, collection of works of Moscow State University, Faculty of Sociology*, issue 18, pp. 26–34. (in Russian)
29. Tverdokhlib I.P., Blonskyi N.A., Kostyuk D.V. (2017). Empyrychna otsinka dotsilnosti monitorynhu informatsiinoho prostoru Internet v ekonomichnykh doslidzhenniakh [EMPIRICAL RATING OF EXPEDIENCY OF MONITORING OF INFORMATION SPACE THE INTERNET IN ECONOMIC RESEARCHES]. *Economy and society: Electronic scientific professional publication*, issue 11, Mukachiv, pp. 593-602. URL: <http://www.economyandsociety.in.ua> (in Ukrainian)
30. Tverdokhlib I.P., Kostyuk D.V. (2018). Monitorynh stanu informatyzatsii Lvivshchyny v Internet-prostori za dopomohoiu system kontent-analizu [Monitoring the state of informatization of Lviv region in the Internet space using content analysis systems]. *"Formation of Knowledge Economy as the Basis for Information Society"* [Thesis of the 7th International Scientific Seminar] / by I.Y.Shtuler, D.V. Solokha, A.A. Oleshko, Z.B.Zhyvko etc. Kyiv-Venice-Verona: IAIS-NAM, 2018, pp. 42-47. (in Ukrainian)
31. Beresneva D. (2016). Modelirovanie dinamiki chisla publikatsiy po ekonomicheskoy tematike v internet-izdaniyakh [Modeling the dynamics of the number of publications on economic topics in Internet publications]. *Sat. works of RANEPА students «Modern economy: theory, politics, innovation»*. Moscow: RANEPА, pp. 174-181. (in Russian)
32. Kondrat'ev N.D. (1989). Plan i predvidenie (k voprosu o metodakh sostavlениya perspektivnykh planov razvitiya narodnogo khozyaystva i sel'skogo khozyaystva v chastnosti) [Plan and foresight (on the question of methods of drawing up long-term plans for the development of the national economy and agriculture in particular)]. *Problems of economic dynamics*. Moscow: Ekonomika, pp. 91-134. (in Russian)
33. Tverdokhlib I, Petryk Gh. (2005) Efektyvnistj ekonomiko-matematychnogho modeljuvannja procesiv transformacijnoji ekonomiky: prahmatychnyj aspect [The

- effectiveness of economic and mathematical modeling of transformational economics: a pragmatic aspect]. *Visnyk of the Lviv University. Series Economics*, issue 34, pp. 227-241 (in Ukrainian)
34. Gurvich E. (2006). Naskol'ko tochny makroekonomicheskie i byudzhetye prognozy? [How accurate are macroeconomic and fiscal forecasts?]. *Voprosy Ekonomiki*, no. 9, pp. 4-20. (in Russian)
  35. Patskan V. (2019). Cherez nyzku yakist makroekonomichnoho prohnozuvannia nedonadkhozhenia der-zhavnoho biudzhetu u 2019 rotsi siahnuly 43 mlrd. hrn. 2019 [Due to the low quality of macroeconomic forecasting, the state budget shortfall in 2019 reached UAH 43 billion]. URL: <https://rp.gov.ua/PressCenter/News/?id=860> (in Ukrainian)
  36. Tverdokhlib I. (2014) Koncepcija informacijnogho Web-resursu pidtrymky rishenj u sferi prohnozuvannja [The concept of information Web-resource to support solutions in the field of forecasting]. Proceedings of the 3<sup>rd</sup> *Internacional Academic Conference «Information, Communication, Society 2014» (Ukraine, Lviv - Slavske, May 21-24, 2014)*, Ljviv : Lviv Polytechnic National University, pp. 28-29 (in Ukrainian)
  37. Tverdokhlib I.P. (2014) Servis-orientovana arkhitektura informatsiinoho web-resursu pidtrymky rishen u sferi prohnozuvannia [Service-oriented architecture of information web-resource to support solutions in the scope of forecasting]. Proceedings of the *XIII International Scientific Seminar «Modern problems of computer science in management, economics and education» (Ukraine, Kyiv - Svityaz, June 30 - July 4, 2014)* (eds. prof. M. M. Yermoshenko), Kyiv : National Academy of Management, pp. 65-70 (in Ukrainian)
  38. Monitoring SMI ot Medialogii. Monitoring pressy [Media monitoring from Medialogia. Press monitoring]. URL: <http://www.mlg.ru/solutions/pr/monitoring/> (in Russian)
  39. Voronka prodazhiv: shcho tse take i yak yii pobuduvaty? [Sales funnel: what is it and how to build it?]. *TurboSMS*. URL: <https://turbosms.ua/ua/articles/voronka-prodazhiv-cto-eto-takoe.html> (in Ukrainian)

#### ACTUALITY ASSESSMENT OF THE QUALITY PROBLEM OF SOCIO-ECONOMIC FORECASTS THROUGH THE PRISM OF INTERNET SPACE

**Ivan Tverdokhlib, Markijan Tymchyshyn**

*Ivan Franko National University of Lviv, Prospekt Svobody 18, UA – 79008, Ukraine  
E-mail: [ivan.tverdokhlib@lnu.edu](mailto:ivan.tverdokhlib@lnu.edu), [i\\_tverdok@i.ua](mailto:i_tverdok@i.ua), [mariktima@gmail.com](mailto:mariktima@gmail.com)*

**Annotation.** The article considers the problem of using content analysis systems to substantiate the expediency of economic researches subjects. Replenishing the researcher's tools with such the latest information technologies is an actual task given the growing role of the Internet in the world. Monitoring of the Internet space as a promising direction in modern economic research is noted. In order to test the expediency of monitoring the information sources of the Internet in economic research, the problem of socio-economic forecasts quality was chosen as the object of analysis. The Internet space was monitored by the InfoStream content analysis system to discuss the problem of forecast quality by the Internet community. Based on the monitoring results of this scientific and practical program in two databases InfoStream (media and forums) a weekly index of the actuality of quality forecast problem during year 2020 was calculated. The well-known reputation index Net Sentiment was used to quantify the actuality of the forecast quality problem in the Internet space. As a specific feature of the research, we note the need to formalize the concepts of search queries in two

languages (Ukrainian and Russian), as InfoStream databases are focused on the post-Soviet information space. In general, the obtained values of the Net Sentiment index testify the stable weekly actuality of the forecast quality problem among the Internet community in the year 2020. This can be interpreted as a signal to the researcher about the feasibility of analyzing such issues. We also note the significant predominance the number of Russian-language sources on the quality of forecasts over Ukrainian-language ones, which indirectly testifies to the predominant focus the discussion of this problem in the post-Soviet space on the still Russian language. Regarding the feasibility of using content analysis systems in scientific economic research, we note the dependence of the possible positive effect on the completeness of the database coverage such systems information sources of the Internet, and the adequacy of formalizing the researcher's requirements in their search languages.

**Key words:** innovations, forecasting, quality of forecasts, actuality of the problem, monitoring of Internet space, content analysis, Text Mining, InfoStream.

*Стаття надійшла до редколегії 04.10.2020*  
*Прийнята до друку 29.12.2020*