

УДК 330:[004.5+004.7+004.8+004.9](002.8)

Емпірична оцінка доцільності моніторингу інформаційного простору Інтернет в економічних дослідженнях

Твердохліб І.П.

кандидат економічних наук, доцент,
Львівський національний університет імені Івана Франка

Блонський Н.А., Костюк Д.В.

студенти
Львівського національного університету імені Івана Франка

У статті розглянуто проблему автоматизації пошуку у мережі Інтернет документів, релевантних меті аналітичного дослідження, за допомогою систем контент-аналізу. Проаналізовано причини використання систем моніторингу Інтернет-простору в управлінні організаціями. Обґрунтування управлінських рішень на засадах інформаційного управління економічними об'єктами відзначено як основний чинник упровадження систем контент-аналізу в їхній контур управління. Обґрунтовано технологічний процес застосування систем контент-аналізу в аналітичних дослідженнях. На прикладі системи InfoStream моніторингу Інтернет-простору оцінено доцільність використання контент-аналізу в бізнес-аналітиці.

Ключові слова: економіка, бізнес-аналітика, інформаційне управління, моніторинг Інтернет-простору, text mining, контент-аналіз, InfoStream.

Твердохліб І.П., Блонський Н.А., Костюк Д.В. ЭМПИРИЧЕСКАЯ ОЦЕНКА ЦЕЛЕСООБРАЗНОСТИ МОНИТОРИНГА ИНФОРМАЦИОННОГО ПРОСТРАНСТВА ИНТЕРНЕТ В ЭКОНОМИЧЕСКИХ ИССЛЕДОВАНИЯХ

В статье рассмотрена проблема автоматизации поиска в сети Интернет документов, релевантных цели аналитического исследования, при помощи систем контент-анализа. Проанализированы причины использования систем мониторинга Интернет-пространства в управлении организациями. Обоснование управленческих решений на принципах информационного управления экономическими объектами отмечено как основной фактор внедрения систем контент-анализа в их контур управления. Обосновано технологический процесс применения систем контент-анализа в аналитических исследованиях. На примере системы InfoStream мониторинга Интернет-пространства оценена целесообразность применения контент-анализа в бизнес-аналитике.

Ключевые слова: экономика, бизнес-аналитика, информационное управление, мониторинг Интернет-пространства, text mining, контент-анализ, InfoStream.

Tverdokhlib I.P., Blonskij N.A., Kostyuk D.V. EMPIRICAL RATING OF EXPEDIENCY OF MONITORING OF INFORMATION SPACE THE INTERNET IN ECONOMIC RESEARCHES.

The article considers search automation of relevant to the analytical research goal documents in the Internet with the help of content analysis system. The reasons of using monitoring systems of Internet-space in management of the organizations was analyzed. Substantiation of management decisions on the basis of the information management of economic objects marked as the main factor of implementation of content analysis systems in their contour management. Technological process of content analysis systems application in analytical researches was substantiated. By the example of InfoStream system of monitoring Internet space evaluated the feasibility of using content analysis in business intelligence.

Keywords: economy, business-analytics, information management, monitoring of Internet-space, text mining, content-analysis, InfoStream.

Постановка проблеми у загальному вигляді. Значний приріст обсягів даних в організаціях приводить до зростання затрат на їх опрацювання. Приміром, на підставі обстежень фірм аналітики стверджують, що офісні працівники витрачають до половини свого робочого часу на рутинне опрацювання контенту даних, і оцінюють усереднено втрати

підприємств у перерахунку на одного працівника у \$2500-3500 [1].

Останнім часом значні зусилля фахівців спрямовані на розроблення методів і систем Text Mining [2], оскільки більшість знань зберігається у текстовій формі [3, с. 13; 4]. З огляду на зростаючу роль мережі Інтернет як інформаційного середовища зберігання і поширення

документів, такі методи і системи найчастіше використовуються для пошуку потрібних фахівцям даних. Зазвичай агентно-орієнтовані програмні системи, базовані на методах Text Mining та призначені для пошуку інформації за запитами користувачів, відомі як *системи контент-аналізу або системи моніторингу Інтернет-простору* [5, 6]. Вони уможливають реалізацію контекстно-орієнтованого управління знаннями в організації [3, с. 224].

Сьогодні ринок систем контент-аналізу динамічно розвивається [7, 8]. Практика їх застосувань окреслила такі типові сфери використання, як *збирання інформації щодо об'єктів спостережень у конкурентній розвідці* [9, с. 16-71], *відстеження репутації фірми* [9, с. 117-148; 10], *аналіз соціальних мереж* [5; 9, с. 98-116; 11], *протидія інформаційним загрозам організації* [12, с. 125-140; 9, с. 149-189], *формування аналітичних оглядів і зведень у журналістиці* [6, 13], *реферування текстів* [6, 14] тощо.

Отже, можна констатувати дещо обмежений характер застосування систем контент-аналізу в економіці, підприємстві й управлінні та в економічних дослідженнях. З огляду на значний приріст інформації, швидкість змін у бізнесі, можливості комунікаційних технологій системи такого класу здатні покращити ефективність процесу обґрунтування управлінських рішень в організації за рахунок урахування контексту управлінської проблеми та пришвидшення опрацювання великих обсягів інформації з мережі Інтернет.

Аналіз останніх досліджень і публікацій. Методи Text Mining та технології опрацювання неструктурованих даних розглядаються багатьма фахівцями, зокрема, і у працях [2, с. 211-241; 15, с. 203-263; 5, с. 177-335; 9, с. 26-30]. Причини впровадження систем контент-аналізу в контур управління підприємствами зазначено у [1, 4, 16]. Кібернетична модель підприємства як двокомпонентної складної системи зі зворотним зв'язком для обґрунтування управлінських рішень описана у багатьох підручниках, зокрема в [17]. Структура сучасного простору всесвітньої мережі Інтернет охарактеризована у працях [9, с. 72-95; 5, с. 38-110], де наголошено на існуванні невидимого веб-простору (*invisible* або *hidden web* [5, с. 50-51]) як сукупності інформаційних ресурсів, не охоплених звичайними пошуковими машинами. Методи і моделі інформаційного пошуку у мережах висвітлено у низці публікацій, зокрема у [5, с. 135-176; 18, с. 55-112;

2, с. 350-381; 15, с. 133-173]. Теоретичні засади інформаційних потоків як основи моніторингу ЗМІ в Інтернеті та базові процедури їх агрегації описано в [5, с. 17-38, с. 336-343; 15, с. 264-307; 6, с. 85-89]. Сучасний стан ринку систем контент-аналізу Інтернет-простору оприлюднюється у серії оглядів таких систем [7, 8, 19]. Значна увага дослідників приділена проблемі інформаційного управління в соціальних мережах як сучасного засобу досягнення цілей організацій [10; 11; 9, с. 98-136]. Інформаційні війни стали характерною ознакою 21 століття, і тому моніторинг Інтернет-простору розглядається як спосіб протидії інформаційним загрозам підприємства [12, с. 12-21, с. 125-140; 9, с. 149-189; 5, с. 357-429]. Типові форми подання звітів бізнес-аналітики розглядаються у [6, с. 102-104; 14, 20]. Також є публікації з певними негативними аспектами застосування систем контент-аналізу на практиці. Зокрема, низка авторів підкреслюють обов'язковість участі фахівця із предметної області у процесі моніторингу Інтернет-простору для остаточної верифікації віднайдених документів [21, 16, 4], пропонують удосконалені моделі інформаційного пошуку на основі інтелектуалізації та онтологій [22, 23].

Формулювання цілей статті (постановка завдання). Незважаючи на значний рівень зацікавленості проблемою опрацювання неструктурованої інформації в управлінні підприємствами, низка аспектів автоматизації її опрацювання та використання для обґрунтування управлінських рішень не мають поки що задовільного вирішення. Дискусійним залишається питання ефективності застосування систем моніторингу Інтернет-простору в бізнес-аналітиці та доцільності їх вбудовування у контур управління організацією, методологічні засади використання систем контент-аналізу в економічних дослідженнях нечітко окреслені, продовжується розроблення нових методів Text Mining та індексів оцінювання рівня зацікавленості проблемою у ЗМІ і в соціальних мережах.

Отже, *метою статті* було уточнення технології моніторингу Інтернет-простору у контексті економічної чи управлінської проблеми в онлайн-режимі використання систем контент-аналізу з подальшою верифікацією доцільності їх застосування на прикладах.

Виклад основного матеріалу дослідження. Насамперед звернемо увагу на дві обставини, які стимулюють зацікавленість системами контент-аналізу в управлінні еко-

номічними об'єктами, а також під час економічних досліджень.

По-перше, аналіз типових звітів систем моніторингу ЗМІ вже окреслює певні їхні можливості. Стиль подання тексту аналітичного звіту діловий, стосується тільки досліджуваної проблеми, орієнтований на зайняту особу, насичений графічними елементами і зазвичай включає такі компоненти, як [14, 20]:

- загальна характеристика аналітичного дослідження (формулювання проблеми, часовий інтервал, опис методики дослідження, інструменти);

- тексти запитів на пошук інформації в Інтернеті мовою системи;

- аналіз динаміки релевантних проблемі документів на часовому проміжку у розрізі сегментів Інтернет-простору з урахуванням тональності публікацій;

- ідентифікація ключових інфосюжетів у контексті проблеми та аналіз їх впливу на конкурентоспроможність організації;

- оцінка репутації організації у контексті досліджуваної проблеми;

- формулювання висновків та рекомендацій для керівництва підприємства.

По-друге, серед 5 базових важелів керівника (*ресурси, фінанси, кадри, організація, інформація*) [24, с. 11] саме поліпшення використання інформації найшвидше приводить до покращення ефекту управлінських рішень. Причина – у переході до інформаційного управління, яке сьогодні розглядається як перспективний напрям менеджменту організацій у світі [25]. Інформаційне управління міняє спосіб реалізації оберненого зв'язку у відомій кібернетичній моделі фірми [17]. Особливості управлінських рішень для відо-

мих способів реалізації оберненого зв'язку охарактеризовано в табл. 1. Переваги інформаційного управління наочно продемонстровані на рис. 1.

На рис. 1 прямокутниками окреслено віднайдені для відповідного запиту дані як фрагменти інформаційного простору, стрілка \Rightarrow відображає процес формування управлінського рішення R_d^β . Оскільки рішення приймається на основі інформації, що характеризує проблему, то на схемі відображено три типові сценарії його обґрунтування. В ідеалі використовується повна і достовірна інформація про ситуацію в організації і ринковому середовищі. Тоді залежно від способу реалізації оберненого зв'язку можна отримати рішення або R_K^{id} , або R_a^{id} , або R_{opt}^{id} відповідно для принципів *класичного кібернетичного, адаптивного чи оптимального*. Другому сценарію (1-й запит) відповідають управлінські рішення R_K^1, R_a^1, R_{opt}^1 , які обґрунтовуються повністю не релевантній до проблеми інформації. Ситуація 2-го запиту ілюструє сценарій формування управлінських рішень на підставі частково релевантної до проблеми інформації. Такі рішення позначено на рис. 1 як R_K^2, R_a^2, R_{opt}^2 .

Якщо позначити через $E(R_d^\beta)$ оцінку ефекту управлінського рішення R_d^β , то на підставі табл. 1 і рис. 1 з очевидністю впливає справедливість відношень

$$E(R_{opt}^{id}) \geq E(R_a^{id}) \geq E(R_K^{id}), \quad E(R_{opt}^2) \geq E(R_a^2) \geq E(R_K^2), \\ E(R_{opt}^1) \geq E(R_a^1) \geq E(R_K^1). \quad (1)$$

Як наслідок з (1), можна стверджувати, що

$$E(R_{opt}^{id}) \geq E(R_{opt}^2) \geq E(R_{opt}^1), \quad E(R_a^{id}) \geq E(R_a^2) \geq E(R_a^1), \\ E(R_K^{id}) \geq E(R_K^2) \geq E(R_K^1). \quad (2)$$

Таблиця 1

Реалізації оберненого зв'язку в кібернетичній моделі організації

Назва	Суть реалізації
Принцип зворотного зв'язку	Раціональне управлінське рішення на підставі аналізу відхилень фактичних показників від планових за сталої структури організації з метою виведення процесу на планову траєкторію розвитку.
Принцип адаптивного управління	Раціональне управлінське рішення відповідно до поточної ринкової ситуації на підставі аналізу відхилення фактичних показників від планових та адаптації структури організації з метою виведення процесу на планову траєкторію розвитку.
Принцип оптимального управління	Оптимальне управлінське рішення відповідно до поточної ситуації в ринковому середовищі з урахуванням відхилення від планової траєкторії та реорганізації структури організації з метою виведення процесу на оптимальну траєкторію.
Принцип інформаційного управління	Управлінське рішення відповідно до інформаційного опису ринкової ситуації та стану і структури організації з метою коригування траєкторії розвитку процесу згідно зі стратегією організації.

Джерело: складено на підставі [17]

Нерівності (2) можна трактувати як серйозний аргумент на користь застосування систем контент-аналізу для обґрунтування управлінських рішень та в економічних дослідженнях. Проте технологічний процес їх використання описаний у фахових джерелах дуже поверхово. Приміром, у [5, с. 339] процес їх застосування включає 3 загальні етапи – *моніторинг, адаптивне агрегування, узагальнення*. Автори [22] пошук потрібної управлінцю інформації розбили на два процеси, такі як пошук *серед структурованих даних та пошук у неструктурованій інформації за допомогою методів Text Mining, а у методичці [14, с. 4] вказано 2 етапи*. Тому потрібно уточнити технологічний процес моніторингу ЗМІ в Інтернеті у контексті заданої проблеми з метою конкретизації дій аналітика.

Оскільки основним режимом використання аналітиком систем контент-аналізу є діалог із системою, то саме застосування їх у режимі онлайн і було нами розглянуте. Уточнений процес обґрунтування управлінського рішення на засадах інформаційного управління з використанням систем контент-аналізу Інтернет-простору складається з таких послідовних етапів, як (див. рис. 2):

- постановка проблеми моніторингу Інтернет-простору;
- підготовка запиту на пошук потрібної для обґрунтування рішення інформації за допомогою системи контент-аналізу;
- моніторинг інформаційних джерел вибраної БД системи контент-аналізу;

- оцінка відповідності результатів пошуку проблемі;
- уточнення запиту на пошук у вибраній БД системи за потребою;
- консолідація змісту множини документів, отриманих з інформаційних джерел вибраної БД системи;
- пошук документів з проблеми в інших БД системи контент-аналізу;
- обґрунтування рішення на підставі віднайдених документів;
- завершення процесу.

На першому етапі (блок № 1) потрібно чітко окреслити проблему, для вирішення якої розробляється управлінське рішення. Приміром, чи можливе розширення ринків збуту підприємства, яка динаміка іміджу організації у ЗМІ, наскільки актуальною є певна тематика серед фахівців тощо.

Метою другого етапу, що включає блоки № 2-4, є підготовка запиту для системи контент-аналізу. Для цього спочатку запускається вибрана система в режимі on-line (як правило) – блок № 2. Свідченням успішного запуску є вивід на екран дисплея вікна-запрошення, після чого потрібно зареєструватися, і тоді система готова до виконання запитів користувача. Другим кроком має бути формалізація запиту на моніторинг Інтернет-простору. Полягає у формуванні тексту запиту пошуковою мовою вибраної системи контент-аналізу (блок № 3).

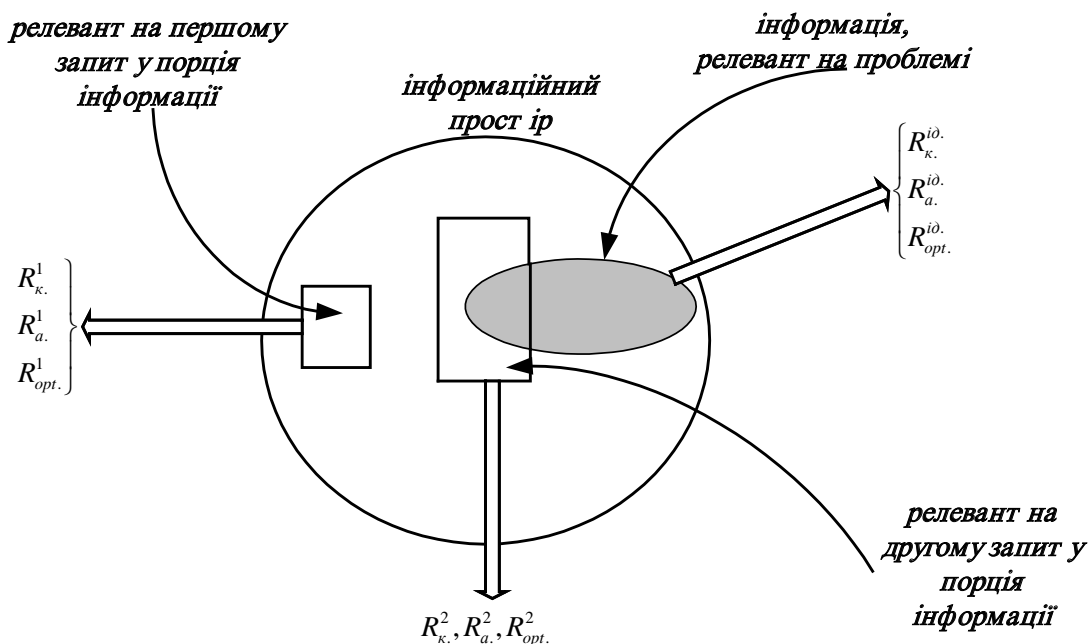


Рис. 1. Ілюстрація залежності управлінського рішення від результатів пошуку релевантної інформації, авторська розробка

Є важливою операцією, оскільки від адекватності запиту проблемі суттєво залежить якість множини віднайдених системою документів (див. рис. 1). Завершує цей етап процедура передачі тексту запиту системі. Реалізована традиційно у формі введення попередньо

сформованого тексту у вікно системи (блок № 4). При цьому ще додатково задаються опції пошуку (БД системи, усунення дублювання, морфологія, часовий проміжок тощо).

Моніторинг інформаційних джерел вибраної БД системи контент-аналізу – це третій

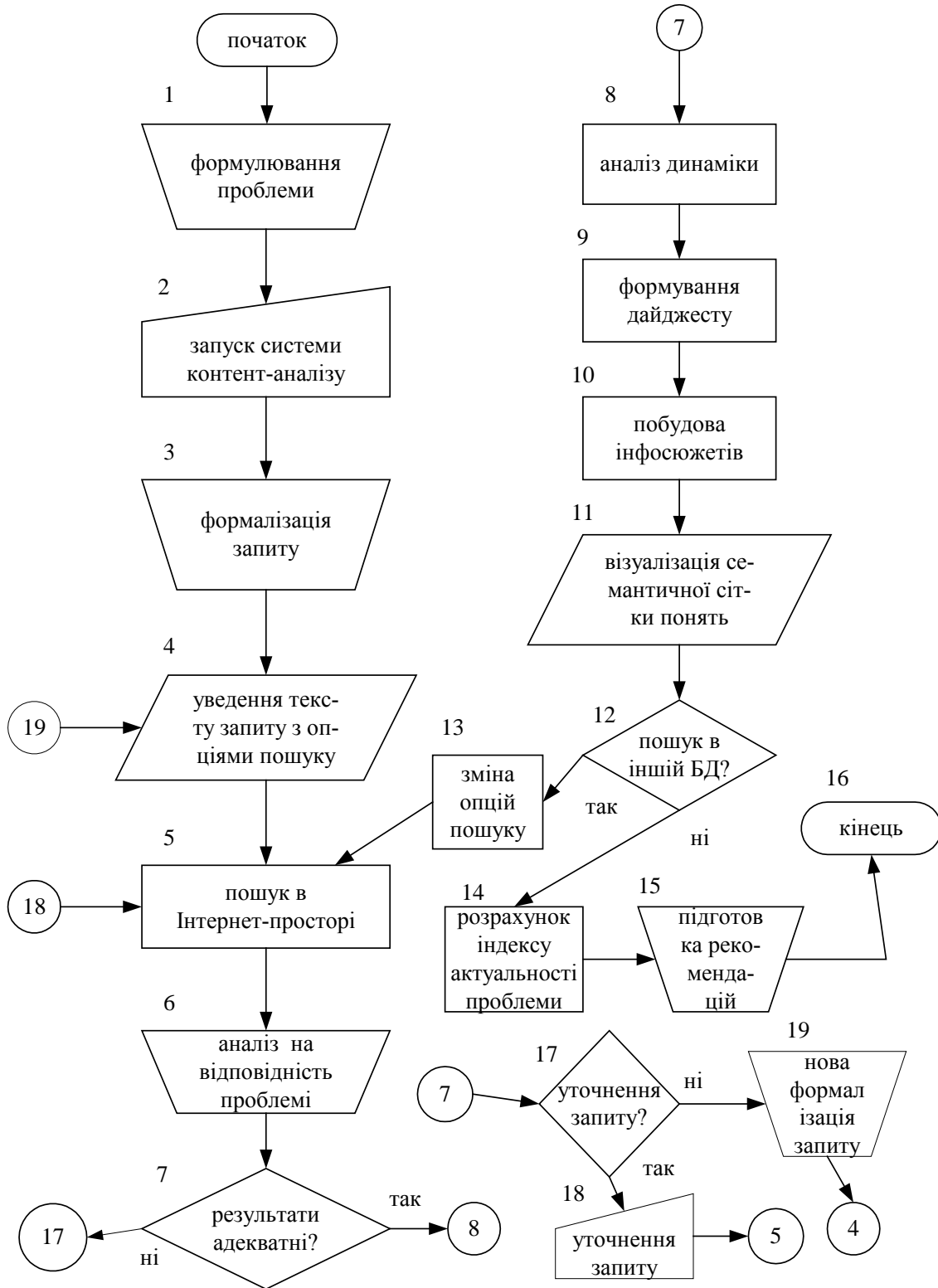


Рис. 2. Блок-схема автоматизації моніторингу Інтернет-простору системою контент-аналізу в режимі online у контексті проблеми, авторська розробка

етап методики, призначенням якого є автоматизація пошуку сервером системи контент-аналізу документів, релевантних запиту, серед інформаційних джерел зазначеної в ньому БД за вказаний часовий проміжок (блок № 5).

На четвертому етапі потрібно здійснити оцінку відповідності результатів пошуку проблемі (блоки № 6, № 7). Під час формалізації запиту на попередніх етапах можливі певні упущення, що може привести до неточного подання контенту проблеми текстом пошукового запиту. Крім того, якість контекстного пошуку у кожній системі різна і загалом бажає кращого. Тому дуже важливо під час використання таких систем здійснювати постійний контроль релевантності результатів пошуку.

Уточнення запиту на пошук у вибраній БД системи – це 5-й етап методики, який виконується у разі невідповідності результатів пошуку контексту проблеми (блоки № 17-19). Уточнення можливе двома шляхами (блок № 17):

- з використанням інформаційного портрету запиту, що уможлиблює швидке коригування тексту запиту за допомогою включення або виключення понять у портреті та дає змогу уточнити контент запиту (блок № 18), після чого здійснюється перехід на блок № 5;
- через нову формалізацію запиту, тобто уточнення тексту запиту пошуковою мовою системи (блок № 19) із подальшим переходом до блоку № 4.

На шостому етапі здійснюється консолідація змісту множини документів, отриманих з інформаційних джерел вибраної БД системи (блоки № 8-11). Традиційно системи контент-аналізу реалізують такі процедури агрегації відібраної множини документів, як:

- аналіз та візуалізація динаміки контенту запиту на часовому проміжку в інформаційних джерелах вибраної БД системи (блок № 8), зокрема і з урахуванням тональності [14];
- формування дайджесту [6, 14] на множині відібраних документів, що може використовуватися як список джерел інформації з проблеми (блок № 9);
- формування інфосюжетів [15, с. 264-307; 14] як метод групування документів за змістом у контексті проблеми (блок № 10);
- візуалізація взаємозв'язків між концептами документів у формі семантичної сітки [14] (блок № 11).

Ще часто за їх допомогою можна відстежити нові події в інформаційному просторі мережі Інтернет, що стосуються досліджуваної проблеми.

Сьомий етап методики (блоки № 12, № 13) виконується, якщо є потреба дослідити проблему у сегменті інформаційного простору, що окреслюється іншою БД системи контент-аналізу. Зазвичай такі системи грубують джерела інформації за певними ознаками і подають групи як бази консолідованих ресурсів відповідного спрямування. Для охоплення

Таблиця 2

Характеристика результатів інформаційного пошуку в базах даних системи InfoStream у контексті проблеми економічного дослідження

Проблема	Дата пошуку	Опції пошуку				Обсяг результату		Оцінка якості пошуку		
		БД	дублі	морфологія	запит	документів	сюжетів	релевантних Док.	частка у %	індекс актуальності
Реінжиніринг бізнес-процесів банківської установи	24.05.2017	головна	ні	так	№ 1	509	359	175	34,4	[0.889;1]
		форуми	ні	так	№ 1	25	17	13	52,0	{-1, +1}
Інформатизація Львівщини	09.05.2017	головна	ні	так	№ 2	24	10	16	66,7	1,0
		форуми	ні	так	№ 2	0	0	-	-	-
	27.05.2017	головна	ні	так	№ 2	38	16	23	60,5	1,0
		головна	так	так	№ 2	59	16	32	54,2	1,0
		головна	ні	так	№ 3	453589	-	-	-	[0.658; 0.718]

№ 1: (реінжиніринг/1/бізнес-процесів) | (реинжиниринг/1/бизнес-процессов)

№ 2: (програма~/0/електронна~/0/львівщина) | (программа~/0/электронная~/0/львовщина)

№ 3: (індекс/1/інформаційного/0/суспільства) | (индекс/1/информационного/0/общества)

Джерело: складено на підставі звітів системи InfoStream

більшій кількості джерел необхідно здійснити інформаційний пошук у кількох БД із переліку системи моніторингу Інтернет-простору. Після прийняття рішення щодо пошуку в іншій БД системи потрібно змінити деякі опції запиту (блок № 13), залишивши текст запиту незмінним, після чого виконати повний цикл описаних вище процедур.

На восьмому етапі здійснюється підготовка рекомендацій із вирішення проблеми. Включає оцінювання її актуальності в інформаційному просторі за допомогою спеціальних індексів [13, 20] (блок № 14) у розрізі БД системи та узагальнених за їх сукупністю з подальшим обґрунтуванням управлінського рішення (блок № 15). Форми подання такого обґрунтування можуть бути різними – від простого переліку пропозицій до аналітичного звіту із проблеми.

Завершальний (дев'ятий) етап методики зображено блоком № 16 схеми на рис. 2. Зазвичай потрібно коректно завершити роботу з системою, можливо зберегти результати для подальшого використання у пам'яті комп'ютера користувача, зберегти тексти запитів у пам'яті сервера системи [14].

Описана вище методика була апробована під час виконання магістерських і курсових робіт магістрантами спеціальності «Консолідована інформація» економічного факультету ЛНУ імені Івана Франка. Пошук потрібних

джерел інформації здійснювався за допомогою системи InfoStream [14] моніторингу ЗМІ в Інтернеті. В табл. 2 (див. с. 598) наведено оцінки якості окремих результатів такого моніторингу у контексті проблем економічних досліджень студентів.

Запити мовою InfoRes [14, с. 31-35], які формалізують контент відповідної проблеми, наведені під таблицею. Оскільки БД системи InfoStream охоплюють головно джерела пострадянських держав, то базові концепти проблем у запиті подано двома мовами. Остаточна релевантність віднайдених документів контексту проблеми встановлювалась вручну через зіставлення їхнього змісту чи анотації та тематики досліджуваної проблеми. Прочерками відзначено відсутність оцінок. Нагадаємо, що головна БД є консолідованим ресурсом ЗМІ, а БД «Форуми» включає соціальні мережі та інші групи за інтересами [14]. Оскільки ця система використовувалася згідно з пакетом послуг BizOnline [14], то часовий проміжок для усіх моніторингів обмежувався 6 місяцями і охоплював інтервал із листопада 2016 р. по травень 2017 р.

В останній графі табл. 2 зазначено оцінку актуальності відповідної проблеми через призму активності її обговорень в інформаційному просторі Інтернет. Для цього скористалися репутаційним індексом – Net sentiment [20]. Цей показник обчислюється за формулою [20]:

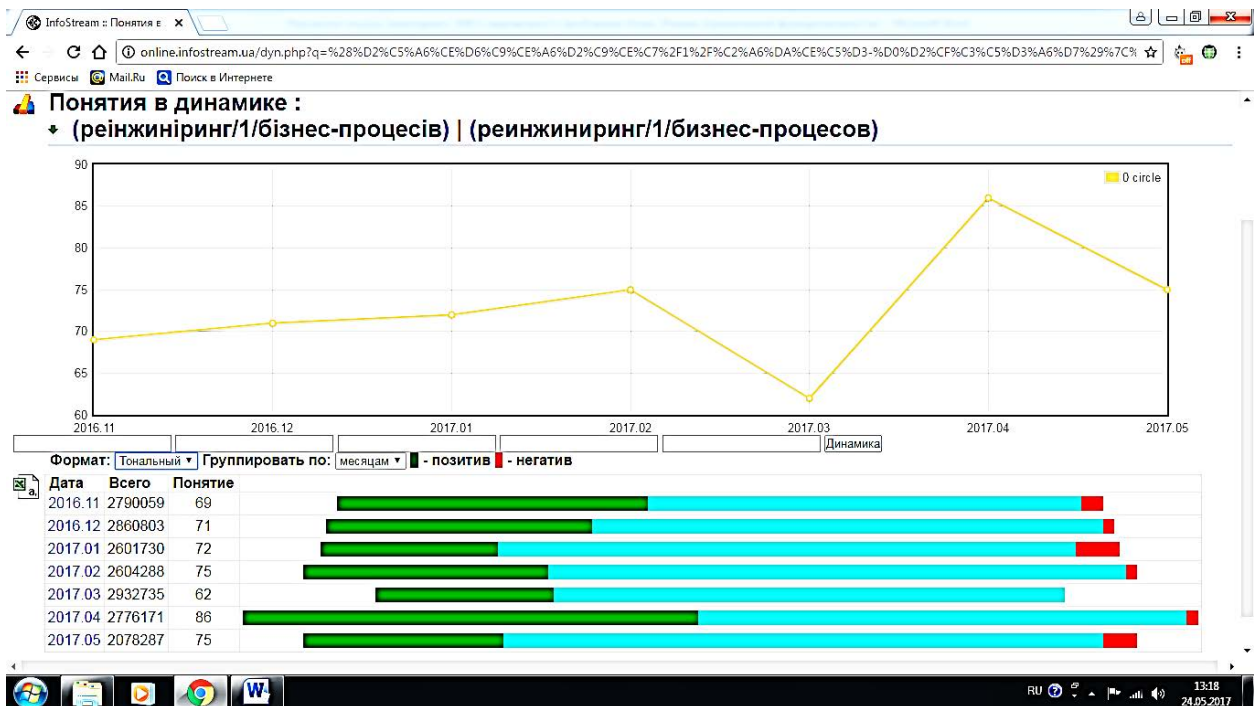


Рис. 3. Тенденції зміни зацікавленості бізнес-реінжинірингом на підставі головної БД системи InfoStream з урахуванням тональності змісту документів у розрізі місяців, авторська розробка

$$I(t) = [k_1(t) + k_2(t) - k_3(t)] / K(t), \quad (3)$$

$$K(t) = k_1(t) + k_2(t) + k_3(t)$$

де t – часовий період, для якого розраховується індекс; $I(t)$ – репутаційний індекс для періоду t ; $k_1(t)$, $k_2(t)$, $k_3(t)$ – кількість релевантних запиту документів за період t , зміст яких відповідно позитивний, нейтральний чи негативний; $K(t)$ – кількість релевантних запиту документів у період t . Отримати необхідні дані для обчислення індексу $I(t)$ можна активізувавши після пошуку операцію «Динамика» і задавши опції агрегації (див. рис. 3 на с. 599).

Вимога тональності змісту документів означає, що горизонтальні стовпчики на діаграмі розбиваються на 3 підгрупи: а) позитивне ставлення до проблеми реінжинірингу бізнес-процесів (зелений колір); б) нейтральне відношення (світло-синій); в) негативне відношення до бізнес-реінжинірингу (червоний колір). Сервер InfoStream забезпечує користувачу можливість отримати обсяги кожної групи документів для вибраного часового періоду. Для цього досить навести курсор на відповідний фрагмент гістограми часового періоду – і над ним виводиться відповідне значення.

Значення репутаційного індексу згідно з (3) може змінюватися в інтервалі $[-1; +1]$. Очевидно, що чим ближче до одиниці значення $I(t)$, тим об'єкт має кращу репутацію в Інтернеті. І навпаки, чим ближче величина $I(t)$ до мінус одиниці, тим гірша репутація об'єкта моніторингу у період t .

Проілюструємо застосування цього індексу для оцінювання актуальності проблеми реінжинірингу бізнес-процесів в Україні (див. рис. 4). Підписи рядів даних на рис. 4 означають: *індекс NS проблеми в засобах ЗМІ*, *індекс NS проблеми в форумах*, *індекс NS проблеми в Інтернеті* – розраховані значення індексу актуальності бізнес-реінжинірингу на підставі відповідно головної БД, БД «Форуми» та узагальнений за цими двома базами системи InfoStream. Оскільки значення розрахованого узагальненого індексу цієї проблеми близьке до 1, то можна констатувати стійкий у часі позитивний характер зацікавленості проблематикою бізнес-реінжинірингу в Україні.

Висновки. На підставі проведеного дослідження проблеми впровадження систем контент-аналізу у практику бізнес-аналітики можна дійти декількох висновків.

По-перше, процес використання систем контент-аналізу для аналітичних досліджень не є простим, про що свідчить блок-схема методики на рис. 2. Його не можна повністю автоматизувати, потрібна активна участь аналітика, а для цього необхідно мати певні навички.

По-друге, загалом застосування систем моніторингу Інтернет-простору в економічних дослідженнях має сенс, оскільки уможливує суттєве зменшення затрат часу дослідника на пошук і опрацювання релевантної проблеми інформації. Засоби візуалізації результатів інформаційного пошуку таких систем врахо-

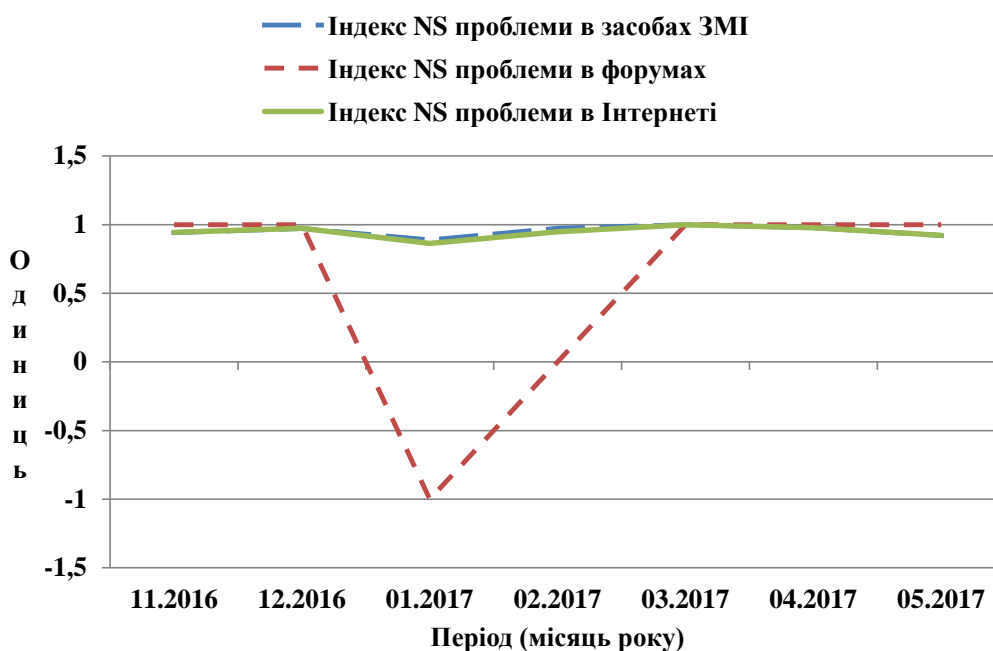


Рис. 4. Динаміка зміни індексу зацікавленості проблемою реінжинірингу бізнес-процесів в розрізі БД InfoStream і місяців року, авторська розробка

вують вимоги бізнес-аналітики. Водночас відзначимо залежність якості віднайдених множин документів від кваліфікації користувача. Критичними процедурами в цьому аспекті є уміння адекватно формалізувати проблему запитом мовою системи та здійснювати уточнення його опцій за допомогою інформаційного портрета.

По-третє, використання системи InfoStream для пошуку в Інтернеті джерел інформації з проблематики економічних досліджень засвідчило вищий відсоток релевантних документів у множині віднайдених, аніж у результатах пошукових машин: від 34,2% до 66,7%. Чим чіткіше сформульована проблема і чим вона конкретніша, тим якіснішою є множина

отриманих документів. Також спостерігаємо вплив дати моніторингу на якість множини релевантних документів, особливо, коли обмежується горизонт моніторингу. Варто задавати опцію усунення дублів під час пошуку документів, оскільки обсяг виданих системою документів зростає (у нас на понад 35%), а кількість релевантних – не змінюється.

По-четверте, доречно здійснити порівняльний аналіз ефективності інформаційного пошуку в Інтернеті різними системами контент-аналізу у контексті однакових аналітичних досліджень. Такі емпіричні оцінки ефекту застосування систем моніторингу ЗМІ в Інтернеті уможливають певні рекомендації щодо їх вибору, які зараз відсутні.

ЛІТЕРАТУРА:

1. Черняк Л. Аналітика неструктурованих даних / Л. Черняк // Открытые системы. – 2012. – № 6. – [Електронний ресурс]. – Режим доступу: <http://www.osp.ru/os/2012/06/13017038/>
2. Барсеян А.А. Аналіз даних і процесів: учеб. посіб. / А.А. Барсеян [та ін.]. – 3-е изд., перераб. и доп. – СПб.: БХВ-Петербург, 2009. – 512 с.
3. Кудрявцев Д.В. Системы управления знаниями и применение онтологий: учеб. пособие / Д.В. Кудрявцев. – СПб.: Изд-во Политехн. ун-та, 2010. – 344 с.
4. Шатров Д. Аналітичний відділ на робочому столі / Д. Шатров, А. Дикарев. // Директор по безпеці. – 2012. – № 12. – С. 83-87.
5. Додонов А.Г. Комп'ютерні мережі та аналітичні дослідження / А.Г. Додонов, Д.В. Ландэ, В.Г. Путятин. – К.: ИПРИ НАН України, 2014. – 486 с.
6. Матвієнко О.В. Консолідована інформація [текст]: навч. посіб. / О.В. Матвієнко, М.Н. Цивін. – К.: «Центр навчальної літератури», 2014. – 134 с.
7. Обзор поисковых систем. – [Електронний ресурс]. – Режим доступу: <http://novikov.com.ua/service/node/6293>
8. Базенков Н.И. Обзор информационных систем анализа социальных сетей / Н.И. Базенков, Д.А. Губанов // Управление большими системами. – 2013. – Вып. 41. – С. 357-394.
9. Додонов А.Г. Конкурентная разведка в компьютерных сетях / А.Г. Додонов, Д.В. Ландэ, В.В. Прищеп, В.Г. Путятин. – К.: ИПРИ НАН України, 2013. – 250 с.
10. Губанов Д.А. Модели репутации и информационного управления в социальных сетях / Д.А. Губанов, Д.А. Новиков, А.Г. Чхартишвили // Математическая теория игр и ее приложения. – 2009. – Т. 1. – Вып. 2. – С. 14-37.
11. Губанов Д.А. Социальные сети: модели информационного влияния, управления и противоборства / Под ред. чл.-корр. РАН Д.А. Новикова / Д.А. Губанов, Д.А. Новиков, А.Г. Чхартишвили. – М.: Изд-во физ.-мат. литературы, 2010. – 228 с.
12. Горбулін В.П. Інформаційні операції та безпека суспільства: загрози, протидія, моделювання: монографія / В.П. Горбулін, О.Г. Додонов, Д.В. Ландэ. – К.: Інтертехнологія, 2009. – 164 с.
13. Городниченко Ю. Говорить президент: як відрізняються виступи Порошенка в Україні та за кордоном / Ю. Городниченко, О. Талавера, Х. Сью // Європейська правда. – 16.08.2016. – [Електронний ресурс]. – Режим доступу: <http://www.eurointegration.com.ua/articles/2016/08/16/7053438/>
14. Григорьев А.Н. InfoStream. Мониторинг новостей из Интернет: технология, система, сервис: научно-метод. пособие / А.Н. Григорьев, [та ін.]. – К., ООО «Старт-98», 2007. – 40 с.
15. Додонов О.Г. Інформаційні потоки в глобальних комп'ютерних мережах / О.Г. Додонов, Д.В. Ландэ, В.Г. Путятин. – К.: Наукова думка, 2009. – 295 с.
16. Ландэ Д. Динамика информационных потоков и информационные операции / Д. Ландэ // Информационные технологии для менеджмента. – 2010. – № 10. – С. 22-26.
17. Інформаційні системи та технології на підприємстві: конспект лекцій / І.О. Ушакова, Г.О. Плеханова. – Харків: Вид. ХНЕУ, 2009. – 128 с. (Укр. мов.)

18. Снарский А.А. Моделирование сложных сетей: учебное пособие / А.А. Снарский, Д.В. Ландэ. – К.: НТУУ «КПИ», 2015. – 212 с.
19. Софт і сервіси для професійного пошуку. – [Електронний ресурс]. – Режим доступу: <http://isearch.kiev.ua/uk/searchpractice/methodsinstruments/1405softwareandservicesforprofessionalsearch>
20. Мониторинг СМИ от Медиалогии. Мониторинг прессы. – [Електронний ресурс]. – Режим доступу: <http://www.mlg.ru/solutions/pr/monitoring/>
21. Шепелев В.П. Text Mining как высокая технология словоблудия / В.П. Шепелев // Компьютерра. – 2003. – № 43. – С. 40-45.
22. Бабін Д.В. Підвищення ефективності здобуття знань на основі інтелектуального аналізу і структурування інформації / Д.В. Бабін, С.М. Вороной, Є.В. Малащук // Штучний інтелект. – 2005. – № 3. – С. 259-264.
23. Нгуен Ба Нгюк, Тузовский А.Ф. Модель информационного поиска на основе семантических метаописаний / Ба Нгюк Нгуен, А.Ф. Тузовский // Управление большими системами. – 2013. – Вып. 41. – С. 51-92.
24. Ахромеева Т.С. Синергетика и сетевая реальность / Т.С. Ахромеева, [та ін.] // Препринты ИПМ им. М.В. Келдыша. – 2013. – № 34. – 32 с.
25. Соловьев И.В. Информационное управление и информационный менеджмент / И.В. Соловьев // Перспективы Науки и Образования. Международный электронный научный журнал. – 2015. – № 5(17). – С. 114-118.