Електронне наукове фахове видання "Ефективна економіка" включено до переліку наукових фахових видань України з питань економіки (Категорія «Б», Наказ Міністерства освіти і науки України від 11.07.2019 № 975) www.economy.nayka.com.ua | № 6, 2021 | 24.06.2021 р.

DOI: [10.32702/2307-2105-2021.6.89](https://doi.org/10.32702/2307-2105-2021.6.89)

УДК 339.138:004.4

*В. Р. Чаплінський,
к. е. н., старший викладач кафедри економіки підприємства,
Кам'янець-Подільський національний університет імені Івана Огієнка,
м. Кам'янець-Подільський
ORCID ID: 0000-0002-3209-1475*

*О. К. Кушнір,
к. е. н., доцент, Кам'янець-Подільський національний університет імені Івана Огієнка,
м. Кам'янець-Подільський
ORCID ID: 0000-0003-2679-2782*

*О. П. Свідер,
к. е. н., доцент, Кам'янець-Подільський національний університет імені Івана Огієнка,
м. Кам'янець-Подільський
ORCID ID: 0000-0001-6330-857X*

АНАЛІЗ ВЕЛИКИХ ДАНИХ ТА ЇХ ВІЗУАЛІЗАЦІЯ ДЛЯ ПОТРЕБ БІЗНЕСУ

*V. Chaplinsky
PhD in Economics, Senior Lecturer of the Department of Business Economics,
Kamianets-Podilskyi National Ivan Ohienko University
O. Kushnir
PhD in Economics, Associate Professor,
Kamianets-Podilskyi National Ivan Ohienko University
O. Svider
PhD in Economics, Associate Professor,
Kamianets-Podilskyi National Ivan Ohienko University*

ANALYSIS OF BIG DATA AND THEIR VISUALIZATION FOR BUSINESS NEEDS

Розкрито особливості постіндустріального суспільства, та ключові фактори успіху для бізнес середовища. Відмічено проблему великих об'ємів неструктуризованих даних, та важкість роботи з ними. Розглянуто методики Big Data Analytics, а саме, методи Data Mining, технологію Text Mining, технологію MapReduce та візуалізацію даних. Відмічено, що дані методи використовують послідовність процесів накопичення у сховищі даних, аналітичну обробку та видобуток знань. Найбільш поширені методи Data Mining є: базові методи, нечітка логіка, генетичні алгоритми, нейронні мережі. Підґрунтям технології Text Mining – статистичний та лінгвістичний аналіз, методи штучного інтелекту. Ця технологія застосовується для проведення аналізу, забезпечення навігації та

пошуку в неструктурованих текстах. Методи графічного подання результатів аналізу великих даних у вигляді діаграм або анімації для спрощення інтерпретації, полегшення розуміння отриманих результатів.

Визначено, що базовою платформою для генерування ієрархічно-синергетичний комплексів автоматизованих засобів нетривіального аналізу первинних даних і візуалізації його результатів для підтримки рішень став Business Intelligence.

The peculiarities of the post-industrial society and the key success factors for the business environment are revealed. The problem of large volumes of unstructured data and the difficulty of working with them is noted. Big Data Analytics techniques are considered, namely, Data Mining methods, Text Mining technology, MapReduce technology and data visualization.

It is noted that these methods use the sequence of accumulation processes in the data warehouse, analytical processing and extraction of knowledge. The most common methods of Data Mining are: Basic methods, fuzzy logic, genetic algorithms, neural networks. The basis of Text Mining technology is statistical and linguistic analysis, methods of artificial intelligence. This technology is used for analysis, navigation and search in unstructured texts. Methods of graphical presentation of the results of big data analysis in the form of diagrams or animations to simplify interpretation, facilitate understanding of the results.

It is determined that Business Intelligence became the basic platform for generating hierarchical-synergetic complexes of automated means of non-trivial analysis of primary data and visualization of its results to support decisions. All processes aimed at working with Big Data can be grouped into three areas, namely, collecting information about the internal and external environment of the enterprise, data integration, data analysis and their unification and data visualization.

It is considered that the functionality of Business Intelligence software platforms is most appropriate to establish based on the reports of the leading consulting company "Gartner" Magic Quadrant for Business Intelligence and Analytics Platforms.

When implementing BI systems, it is important to start by defining business requirements and setting tasks. The implementation scheme differs from the scope of the enterprise, business processes and tasks, so it is advisable to form a concept for the implementation of the mechanism of implementation of business intelligence tools for visualization of marketing data in the enterprise.

There are also specialized marketing software products on the software market. They are designed to process and analyze information for further use in making strategic marketing decisions, planning and forecasting the development of the enterprise.

Ключові слова: великі дані; аналіз великих даних; business intelligence; програмне забезпечення.

Keywords: big data; big data analytics; business intelligence; software.

Постановка проблеми. Процес інформатизації суспільства та економіки загалом, з кожним роком набирає обертів та проникає у всі галузі та сфери життя. Така тенденція є позитивною для розвитку та функціонування постіндустріального суспільства, однак вона наділена специфічними особливостями. «Big Data» – сукупність великої кількості неструктуризованої інформації яка зростає у геометричній прогресії щороку, та значно ускладнює процес пошуку та аналізу необхідної інформації в мережі. Разом з формуванням великих об'ємів інформації з'являється програмне забезпечення, що дозволяє обробляти та класифікувати необхідну інформацію для спеціалістів з маркетингу. Підбір такого програмного забезпечення є незамінним інструментом майбутнього для великого комплексу аналітичних дій на підприємстві та якості отриманих результатів, що уже є 50% успіху при плануванні будь яких стратегій підприємства.

Аналіз останніх досліджень і публікацій. Дослідженням проблеми систематизації «Великих даних» та аналізом існуючих інструментів бізнес-аналітики для візуалізації маркетингових даних на підприємствах

займалися такі вчені, як: В. Белов, В. Варенко, О. Костенко, О. Новак, Д. Прішниклова, Т. Шабельник. Досліджуючи роботи даних науковців, необхідно відмітити, що проблематика сучасного стану інформатизації розкрито досить широко та детально, однак питанням вибору та систематизації програмного забезпечення для маркетингу та бізнес середовища загалом приділено мало уваги, що у свою чергу створює необхідність для подальшого дослідження та формування відповідних висновків.

Цілі статті. Дослідження ринку програмного забезпечення, що має можливість комплексно систематизувати та проаналізувати великі об'єми інформації для подальшого використання у бізнес середовищі.

Виклад основного матеріалу дослідження. Сучасне бізнес середовище, як ніколи, є досить турбулентним та інформаційно перевантаженим, кількість різноманітних даних з внутрішнього та зовнішнього середовища постійно зростає, стає складнішою та менш структурованою. Існуючі підходи та методи аналізу інформації уже не виконують повноцінно свої функції та стають менш актуальними, виникає потреба пошуку нових можливостей. Найкраще з такими викликами справляються методики Big Data Analytics.

Big Data Analytics – це комплекс методик та підходів, що направлені на акумулювання, систематизацію та обробку великої кількості різної за своїми характеристиками інформації та формуванні на їх основі відповідних висновків та гіпотез. Відповідно функціональні зв'язки Big Data Analytics є досить розгалуженим та включають у себе такі елементи: технології візуалізації, статистичний аналіз, штучний інтелект, технології баз даних, технології розпізнавання образів, об'єднавши які та класифікувавши можна отримати перелік методів аналітики Великих даних.

До таких методів можна віднести: методи Data Mining, технології Text Mining, технологія MapReduce та візуалізація даних.

Для отримання обґрунтованого управлінського рішення дані проходять крізь послідовність процесів накопичення у сховищі даних, аналітичну обробку та видобуток знань. Для реалізації цього процесу використовують методи та алгоритми, які входять до складу технології Data Mining. Близько 80% роботи над Data Mining полягає в зборі та підготовці даних, що проводиться ще до запуску інструментів видобутку знань. Найбільш поширеними методами Data Mining є: Базові методи, нечітка логіка, генетичні алгоритми, нейронні мережі.

Технологія MapReduce – модель розподілених обчислень у комп'ютерних кластерах, представлена компанією Google. Згідно з цією моделлю, додаток розділяється на значну кількість однакових елементарних завдань, що виконуються на вузлах кластера і потім, природнім шляхом зводяться у кінцевий результат [1].

Технології Text Mining. Підґрунтям технології Text Mining – статистичний та лінгвістичний аналіз, методи штучного інтелекту. Ця технологія застосовується для проведення аналізу, забезпечення навігації та пошуку в неструктурованих текстах [2, с. 542]. Застосування інформаційних систем класу Text Mining дає змогу користувачам набувати нових знань.

Візуалізація (*Visualization*). Методи графічного подання результатів аналізу великих даних у вигляді діаграм або анімації для спрощення інтерпретації, полегшення розуміння отриманих результатів. Візуалізація аналітичних даних – зображення інформації у вигляді рисунків, графіків, схем і діаграм з використанням інтерактивних можливостей та анімації для результатів, а також вихідних даних для подальшого аналізу [3, с. 173–210].

Аналізуючи особливості Big Data Analytics, а саме, розгалуженість, багатофункціональність та складність існуючих методик, можна чітко стверджувати, що даний інструментарій є перевантаженим для практичного використання на підприємствах. Бізнес середовище не може витратити багато часу на усі етапи аналізу, а тому, потребує автоматизації даних процесів. Відповідно на рику почало з'являтися програмне забезпечення, функції якого дають змогу виконувати ряд специфічних функцій, що до обробки великих даних.

Програмні засоби великих даних можна класифікувати за задачами, які вони вирішують. У процес створення рішення повинні бути інтегровані різні засоби для зберігання, управління та аналізу великих даних. Інструменти великих даних відповідно до їх задач включають наступні групи:

- ПЗ зберігання великих даних;
- ПЗ управління великими даними;
- ПЗ обробки великих даних;
- методи та засоби візуалізації великих даних;
- методи та засоби аналітики великих даних.

Повноцінне програмне забезпечення по обробці великих даних повинен відображати інструменти для зберігання, управління та обробки такої інформації, інструменти та методи аналітики, візуалізації та оцінювання у різні етапи процесу побудови рішення [4].

Досліджуючи такий ринок програмного забезпечення, необхідно відмітити, що за останні п'ять років він значно збільшився і почав кластеризуватись. В процесі таких згруповань створилась платформа Business Intelligence. Business Intelligence (BI) – це термін-метафора, який не має дослівного тлумачення та означає

«ієрархічно-синергетичний комплекс автоматизованих засобів нетривіального аналізу первинних даних і візуалізації його результатів для підтримки рішень (Decision Support)» [5].

У бізнес середовищі за допомогою ВІ можна виконувати такі завдання:

- класифікація споживачів;
- виявлення асоціативних правил у споживчому попиті та їх використання для збільшення продажів;
- багатомірний аналіз обсягів продажів, маркетингових затрат засобами OLAP;
- оптимізація асортименту;
- прогнозування обсягів продажів та інших показників за допомогою методу регресійного аналізу;
- сегментування ринку за допомогою кластерного аналізу
- оцінка ефективності та оптимізація маркетингових кампаній;
- оптимізаційне управління ціновою політикою [6, с. 264].

Основними операціями, які проводяться з інформацією в бізнес середовищі це – накопичення, аналіз, та побудова на її основі прогнозів та виявлення тенденцій (Табл. 1).

Таблиця 1.
Програмне забезпечення, що використовується при роботі з Big Data у бізнес середовищі

Призначення	Продукт
Для збору інформації про внутрішнє та зовнішнє середовище	Marketing Geo, Mapinfo, ArcGI “Infostreamcorporate”, “Stikler” KonSi-Competitive Intelligence&Benchmarking CRM-системи ERP-системи: 1C, SAP ERP, Галактика ERP.
Інтеграція даних	ERP Integration, ETL Integration, Portal Inte-gration, CRM Integration, MS Office Applica- tions and Big-Data Connectors
Аналіз даних та їх уніфікація	SAS Business Intelligence, Microsoft BI, IBM Cognos BI, SAP Business Intelligence, Oracle Business Intelligence. Tableau 9.0, Qlik Sense 2.0 i Microsoft Power BI
Візуалізація даних	Visual Querying, Storyboarding, Geospatial Integration, Autocharting and Animations

Джерело: складено на основі джерела [7,8]

Аналізуючи вище наведені програми та методи, необхідно відмітити, що усі вони є дієвими та ефективними, та мають місце на практичне застосування, їхня актуальність для бізнесу є досить індивідуальною та залежить від особливостей підприємства та очікуваних результатів від їхнього використання.

Висновки

В умовах розбудови постіндустріального суспільства, ключовим благом якого є інформація та інтелект, неможливо вести бізнес без роботи з цифровими даними та різноманітними інформаційними технологіями. Актуальність та вагомість, на сьогодні програмного забезпечення, яке використовується для аналізу «Big Data» є очевидним, оскільки дає змогу ефективно та в короткі строки опрацювати та систематизувати великі дані та значно полегшити роботу при планування, аналізі та прогнозуванні.

Список літератури.

1. Стоун-брейкер М. MapReduce и параллельные СУБД: друзья или враги? [Електронний ресурс] / Майкл Стоунбрейкер, Дэниэль Абади, Дэвид Девитт, Сэм Мэдден, Эрик Паулсон, Эндрю Павло, Александр Разин ; пер. с англ. Сергей Кузнецов // Communications of the ACM, vol. 53, no. 1, January 2010. – Режим доступу: http://citforum.ru/database/articles/mr_vs_dbms-2/
2. Плескач В. Л. Інформаційні системи і технології на підприємствах : підручник / В. Л. Плескач, Т. Г. Затонацька. – К. : Знання, 2011. – 718 с
3. Паклин, Н. Б. Бизнес-аналитика: от данных к знаниям [Текст] : учеб. пособ. / Н. Б. Паклин, В. И. Орешков. – 2-е изд., испр. — СПб. : Питер, 2013. – 702 с.
4. Big Data: Survey, Technologies, Opportunities, and Challenges. Nawsher Khan, Ibrar Yaqoob, Ibrahim Abaker Targio Hashem, Zakira Inayat, Waleed Kamaleldin Mahmoud Ali, Muhammad Alam, Muhammad Shiraz and Abdullah Gani. The Scientific World Journal. January 2014.
5. Бізнесова аналітика. Електрон. дані(283 331 байт) // Вікіпедія. URL: <http://ru.wikipedia.org/wiki>.
6. Данько Т., Китова О. Вопросы развития цифрового маркетинга. Проблемы современной экономики.

2013. № 3 (47). URL: <https://cyberleninka.ru/article/n/voprosy-razvitiya-tsifrovogo-marketinga>.

7. Сонько С.П. Сучасні геоінформаційні системи як середовище управління маркетинговими базами даних / С.П. Сонько, І.В. Довгаль // Комп'ютерне моделювання та інформаційні технології в науці, економіці та освіті : збірник наукових праць. – К. : KEI KHEU, 2008. – С. 148 – 151.

8. Business Intelligence [BI]: Matrix view. SelectHub. Denver, 2018. URL: <https://app.selecthub.com/community/reports/580037b2deba89470db76486160c8f0e9c902db7c92f95c582e2436c15d2272f>.

References.

1. Stounbrejker, M. and others (2010), “MapReduce and Parallel DBMSs: Friends or Foes?”, *Communications of the ACM*, vol. 53, issue. 1, [Online], available at: http://citforum.ru/database/articles/mr_vs_dbms-2/ (Accessed 14 June 2021).

2. Pleskach, V. L. and Zatonatska, T. H. (2011), *Informatsiini systemy i tekhnolohii na pidpriemstvakh: pidruchnyk* [Information systems and technologies at enterprises: a textbook], Znannia, Kyiv, Ukraine, P. 718.

3. Paklin, N. B. and Oreshkov, V. I. (2013), *Biznes-analitika: ot dannyh k znaniyam* [Business analytics: from data to knowledge], 2nd ed., Piter, SPb., Russia, P. 702.

4. Big Data: Survey, Technologies, Opportunities, and Challenges. Nawsher Khan, Ibrar Yaqoob, Ibrahim Abaker Targio Hashem, Zakir Inayat, Waleed Kamaleldin Mahmoud Ali, Muhammad Alam, Muhammad Shiraz and Abdullah Gani. The Scientific World Journal. January 2014.

5. Wikipedia, Business analytics. Electron. data (283 331 bytes), available at: <http://ru.wikipedia.org/wiki> (Accessed 14 June 2021).

6. Dan'ko, T. and Kitova, O. (2013), “Digital marketing development issues”, *Problemy sovremennoj jekonomiki*, vol. 3 (47), [Online], available at: <https://cyberleninka.ru/article/n/voprosy-razvitiya-tsifrovogo-marketinga> (Accessed 14 June 2021).

7. Sonko, S.P. and Dovhal, I.V. (2008), “Modern geoinformation systems as a management environment for marketing databases”, *Kompiuterne modeliuvannia ta informatsiini tekhnolohii v nauci, ekonomitsi ta osviti: zbiryk naukovykh prats*, pp. 148 – 151.

8. Business Intelligence [BI]: Matrix view. SelectHub. Denver, 2018. [Online], available at: <https://app.selecthub.com/community/reports/580037b2deba89470db76486160c8f0e9c902db7c92f95c582e2436c15d2272f> (Accessed 14 June 2021).

Стаття надійшла до редакції 20.06.2021 р.