УДК 004.8:004.652

https://doi.org/10.52058/2786-5274-2025-3(43)-1307-1317

Мушеник Ірина Миколаївна кандидат економічних наук, доцент кафедри інформаційних технологій, фізико-математичних та безпекових дисциплін Заклад вищої освіти «Подільський державний університет», м. Кам'янець-Подільський, тел.: (098) 975-92-85, https://orcid.org/0000-0003-4379-7358

Збаравська Леся Юріївна кандидат педагогічних наук, завідувач, доцент кафедри інформаційних технологій, фізико-математичних та безпекових дисциплін Заклад вищої освіти «Подільський державний університет», м. Кам'янець-Подільський, тел.: (097) 909-09-91, https://orcid.org/0000-0001-5802-7351

СУЧАСНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ДЛЯ МОДЕЛЮВАННЯ ТА ВИЯВЛЕННЯ РЕЛЕВАНТНОЇ ІНФОРМАЦІЇ У ВЕЛИКИХ БАЗАХ ДАНИХ

Анотація. У статті здійснено поглиблений аналіз сучасних інформаційних технологій та методів дослідження великих баз даних (Big Data) з метою побудови моделей і виявлення релевантної інформації. Зокрема, розглянуто концептуальні засади використання машинного навчання, технологій Data Mining і хмарних обчислень як інструментів для обробки великих обсягів різнорідних даних. Виокремлено основні виклики, що постають при взаємодії з гетерогенними джерелами інформації, включно з проблемами масштабованості, продуктивності та безпеки. Окреслено також можливості оптимізації процесів завдяки застосуванню інноваційних платформ, контейнеризації (Docker, Kubernetes) та сервіс-орієнтованих архітектур.

Особливу увагу приділено питанням ефективної інтеграції інтелектуальних систем у структуру баз даних для підвищення рівня автоматизації та якості ухвалення рішень у бізнесі й науково-дослідній діяльності. Детально представлено концептуальні та методичні підходи, що дозволяють поєднати аналітичні інструменти (AI/ML, Data Mining) із традиційними базами чи сховищами даних. Проаналізовано практичні кейси використання Big Data-платформ (Hadoop, Spark) та інтелектуальних методів обробки інформаційних потоків у різних галузях. Сформульовано рекомендації щодо впровадження вдосконалених технологічних рішень, які сприятимуть своєчасному отриманню релевантних даних та забезпечать точність аналітики в умовах динамічного розвитку цифрового середовища.

Ключові слова: великі бази даних, інформаційні технології, машинне навчання, релевантна інформація, хмарні обчислення, інтелектуальні системи,

обробка даних, інтерактивна візуалізація, безпека даних, доповнена та віртуальна реальність.

Mushenyk Iryna Mykolayivna Ph. D in Economic, Associate Professor of the Department of Information Technologies, Physical, Mathematical and Security Disciplines Higher Educational Institution «Podillia State University» Kamianets-Podilskyi, tel: (098) 975-92-85, https://orcid.org/0000-0003-4379-7358

Zbaravska Lesya Yuriivna Candidate of Pedagogical Sciences, Head Associate Professor of the Department of Information Technologies, Physical, Mathematical and Security Disciplines Higher Educational Institution «Podillia State University» Kamianets-Podilskyi, тел.: (097) 909-09-91, https://orcid.org/0000-0001-5802-7351

MODERN INFORMATION TECHNOLOGIES FOR MODELING AND IDENTIFYING RELEVANT INFORMATION IN LARGE DATABASES

Abstract. The article provides an in-depth analysis of modern information technologies and methods of researching large databases (Big Data) with a view to building models and identifying relevant information. In particular, the conceptual foundations of using machine learning, data mining technologies and cloud computing as tools for processing large volumes of heterogeneous data are considered. The key challenges that arise when integrating heterogeneous information sources, including scalability, security, and performance, are highlighted. Practical ways to optimize data analysis processes through innovative platforms, containerization (Docker, Kubernetes) and service-oriented architectures (SOA) are outlined.

Particular attention is paid to the effective integration of intelligent systems into the structure of large databases to increase the level of automation and quality of decision-making in business and research. Methodological approaches that enable the combination of analytical tools (AI/ML, Data Mining) with traditional warehouses or distributed databases are described in detail. Examples of the implementation of Big Data platforms (Hadoop, Spark) and intelligent methods of processing information flows in the financial sector, industry, marketing, and academic research are demonstrated.

It is noted that the combined use of interactive tools and self-learning models will facilitate the rapid adaptation of systems to changing environmental conditions and the emergence of new types of data (in particular, sensory, social, multimedia). The author also emphasizes the importance of security and ethical dimensions when working with large amounts of information, namely, the formation of a secure infrastructure and adherence to the principles of responsible and transparent data use.

The author outlines approaches to automated system performance monitoring and data lifecycle management that can reduce operating costs and increase the speed of obtaining analysis results. Thus, this review emphasizes the growing importance of modern IT technologies in the field of researching large databases, and provides guidelines for further development and implementation in various sectors of the economy, science, and public life.

Keywords: big databases, information technology, machine learning, relevant information, cloud computing, intelligent systems, data processing, interactive visualization, data security, augmented and virtual reality.

Постановка проблеми. У сучасному цифровому середовищі, де обсяги даних, згенерованих як організаціями, так і окремими користувачами, невпинно зростають, спроможність обробляти й аналізувати цю інформацію стає одним із ключових чинників конкурентоспроможності. Традиційні бази даних та класичні методи аналітики часто не здатні ефективно впоратися з великими обсягами й різноманіттям даних, що вимагає впровадження новітніх інформаційних технологій та більш сучасних підходів до моделювання і виявлення найважливішої релевантної інформації.

Під ***релевантною інформацією*** розуміють ті дані, які мають пряме відношення до досягнення певної мети або розв'язання конкретної задачі. Завчасне ідентифікування таких даних у сховищах великого масштабу дає змогу не лише оптимізувати навантаження на систему, а й суттєво поліпшити точність та оперативність ухвалення управлінських і технічних рішень. У контексті стрімкого зростання обчислювальної потужності та появи технологій розподіленої обробки (наприклад, Hadoop, Spark) реалізація ефективних методів відбору й аналітики релевантної інформації відкриває нові горизонти для бізнесу, науки та державного управління [2].

Зважаючи на це, значна увага приділяється інтеграції в інформаційні системи інструментів машинного навчання та штучного інтелекту, які здатні виявляти приховані закономірності та забезпечувати динамічний аналіз величезних масивів даних у режимі реального часу. Подальший розвиток цих підходів сприятиме формуванню інтелектуальної інфраструктури, де релевантні відомості автоматично відокремлюються від інформаційного шуму, а процес ухвалення рішень базується на більш точних і повних моделях. Таким чином, висока ефективність у роботі з даними стає стратегічною перевагою для будь-якої організації, яка прагне підвищити свою стійкість і гнучкість у конкурентному середовищі.

Аналіз останніх досліджень і публікацій У численних наукових публікаціях зарубіжних і вітчизняних авторів (Atif M. [1], Марченко О.О., Россада Т.В. [6], Ковальова Т.Ю. [5]) розглядається впровадження алгоритмів машинного навчання (ML) і Data Mining у процес обробки великих баз даних. Водночас залишається низка відкритих питань, що стосуються комплексної

інтеграції ML-рішень, хмарних обчислень та методів попередньої обробки даних. Гордійчук-Бублівська О. В. [4] зосереджується на проблематиці розроблення методів і засобів обробки великих даних у розподілених інформаційних системах. Попри наявність численних публікацій, присвячених упровадженню Big Data-технологій, тут акцент зроблено на створенні інтегрованого підходу до розподіленого опрацювання інформації з урахуванням показників продуктивності, масштабованості й безпеки.

Мета статті – аналіз сучасних інформаційних технологій, концепцій і підходів до моделювання та виявлення релевантної інформації у великих базах даних, а також розробка рекомендацій щодо їх застосування в різних галузях.

Виклад основного матеріалу. У сучасному інформаційному середовищі, де стрімко зростають і обсяги даних, і потреби в оперативному доступі до них, визначальною стає здатність опрацьовувати великі та різнорідні масиви інформації. Саме цим пояснюється надзвичайна актуальність новітніх технологій, спрямованих на ефективне збирання, зберігання й аналітику. До провідних рішень у цій галузі належать системи хмарних обчислень (**Cloud Computing**), що дають змогу гнучко масштабувати ресурси залежно від конкретних потреб і обсягів оброблюваних даних. Платформи на кшталт Amazon Web Services, Microsoft Azure чи Google Cloud пропонують готові інструменти, які поєднують високу продуктивність, зручність налаштування та швидке розгортання застосунків у режимі реального часу або при пакетній обробці великих обсягів інформації.

Важливою тенденцією в побудові надійних та масштабованих IT-систем залишається впровадження контейнеризації й мікросервісної архітектури. Завдяки цим методам кожен компонент або процес працює у власному середовищі, що спрощує оновлення, віддалене розгортання й адміністрування. Технології **Docker** та **Kubernetes** значно полегшують розподіл навантаження між різними сервісами, дозволяючи ефективно реагувати на зміни у вимогах до пропускної здатності й продуктивності. Крім того, така підхід підвищує стійкість системи до відмов і збоїв в окремих її складових [7].

Інтеграція хмарних сервісів із контейнеризацією та мікросервісною архітектурою формує комплексну платформу для успішної роботи з великими обсягами даних. Усе це зумовлює нові можливості для підприємств та організацій — від оперативного прийняття рішень на основі актуальної аналітики до оптимізації витрат на обчислювальну інфраструктуру й утримання власних дата-центрів. Одночасно розширюються перспективи автоматизації бізнес-процесів, розвитку гнучких операційних моделей та впровадження систем штучного інтелекту. Саме тому дослідження методів, підходів і засобів використання великих баз даних у хмарному та контейнеризованому середовищі є одним із ключових напрямів у сфері інформаційних

технологій, що визначають конкурентоспроможність і стратегічний розвиток сучасних організацій [12].

Робота з великими обсягами даних у сучасних системах вимагає оперативного виявлення релевантної інформації, що робить критично важливими технології машинного навчання (ML) й інтелектуального аналізу (**Data Mining**). Алгоритми кластеризації, зокрема **K-means** чи **DBSCAN,** дозволяють групувати об'єкти на підставі подібності, водночас відокремлюючи «шумові» записи від цінної інформації. Класифікаційні методи, як-от **Random Forest** або SVM, надають можливість швидко визначати категорії нових спостережень і можуть бути застосовані в режимі реального часу. Паралельно, нейронні мережі глибинного навчання забезпечують розпізнавання складних нелінійних залежностей у даних, що особливо затребуване при опрацюванні зображень, аудіо- та текстових матеріалів.

Помітне місце серед методів Data Mining посідає асоціативний аналіз (**Apriori, FP-Growth**), метою якого є виявлення логічних зв'язків і закономірностей у транзакціях або лог-файлах. Аналіз часових рядів виступає незамінним інструментом у прогнозуванні різноманітних тенденцій і пошуку нетипових відхилень у часових вибірках, що є важливим для фінансових, технічних чи соціальних даних. Крім цього, обробка природної мови (NLP) розкриває нові можливості під час аналізу великих масивів текстової інформації, дозволяючи виокремлювати семантично значущі елементи та будувати бази знань, орієнтовані на конкретний предмет дослідження [6].

Широке коло методів ML і Data Mining у поєднанні з NLP-технологіями формують комплексний інструментарій для глибинної аналітики та якісної обробки великих масивів інформації. Подальший розвиток цих технологій спростить адаптацію систем під специфічні вимоги галузі — від онлайн-маркетингу та промислового моніторингу до наукових досліджень і державного сектору, де швидкість та точність опрацювання даних стають стратегічно важливими.

Важливим підґрунтям побудови якісної аналітичної системи є вибір методичного підходу до моделювання великих даних, і ключову роль тут відіграють дві концептуальні парадигми: **CRISP-DM (Cross-Industry Standard Process for Data Mining)** та **SEMMA (Sample, Explore, Modify, Model, Assess)**. Перший з них (CRISP-DM) пропонує повний цикл від розуміння бізнес-процесів і визначення вимог до даних, через їхнє опрацювання, побудову моделі та інтеграцію результатів, до отримання конкретних висновків і контролю якості впровадженої системи. SEMMA, у свою чергу, передбачає побудову ефективної схеми аналізу з акцентом на послідовне відбирання вибірки, дослідження й модифікацію даних, моделювання та оцінювання результатів.

Обидва підходи містять низку спільних рис, зокрема етапи збирання даних із різноманітних джерел, очищення від шуму й помилкових записів, нормалізації

та трансформації для подальшого опрацювання. Важливо, що в обох методиках моделювання не обмежується лише вибором алгоритмів: воно передбачає відстеження точності, релевантності та стійкості створених моделей у контексті реальних бізнес-вимог чи наукових завдань. Одночасно наголошується на візуалізації результатів, без якої складно об'єктивно оцінити коректність моделі та представити результати аналітики зацікавленим сторонам [9].

Подальший розвиток цих підходів перебачає адаптацію CRISP-DM та SEMMA до сучасних умов обробки великих і різнорідних даних, зокрема у розподілених обчислювальних середовищах. Ідеться про інтеграцію з хмарними платформами, мікросервісними архітектурами та системами контейнеризації, що дає змогу поліпшити масштабованість і забезпечити гнучкість під час швидкого розгортання аналітичних модулів. Крім того, актуальним стає впровадження методів автоматизації (AutoML) і механізмів безперервної інтеграції (CI/CD) у контур процесів CRISP-DM та SEMMA, що створює більш динамічну і стійку екосистему для діяльності у сфері великих даних [11].

Практичний підхід до обробки даних передбачає кілька ключових етапів, які забезпечують логічно впорядкований та ефективний процес отримання цінної аналітичної інформації. Насамперед виконується **збирання та інтеграція** даних із різних джерел (внутрішні бази даних, веб-сервіси, датчики IoT, зовнішні API) з уніфікацією форматів і обов'язковою верифікацією якості. Важливо, щоб уже на цьому етапі було передбачено механізми контролю версійності, формування метаданих і фіксування ланцюга перетворень (data lineage), що допоможе відстежувати походження та зміни, внесені в інформацію.

Наступний крок — **очищення і нормалізація**, коли здійснюється видалення або коригування некоректних та дубльованих записів, а також зниження шуму в наборі даних, аби підвищити точність моделювання. Після цього відбувається **трансформація та завантаження (ETL)**, де зібрані дані пристосовуються до вимог аналітичної платформи чи сховища (Data Warehouse або Data Lake). На цьому ж етапі нерідко передбачаються додаткові процедури, такі як агрегація, кодування категорійних ознак або індексація для прискорення вибірок [1].

Коли дані підготовлені, настає критичний етап **моделювання**, який передбачає вибір оптимальних методів — від традиційної класифікації й кластеризації до регресії або рекомендаційних алгоритмів, з урахуванням специфіки завдання. Сучасні підходи включають автоматизацію вибору гіперпараметрів (AutoML) і можливість паралельних експериментів для пошуку найкращої моделі. На завершення обов'язковим стає **оцінювання** продуктивності побудованих моделей.

Оскільки результати аналітики мають приносити користь у реальних сценаріях, важливо надати **зручну візуалізацію** та засоби інтерактивної

взаємодії з даними. Це полегшує інтерпретацію моделей і сприяє швидкому ухваленню рішень фахівцями у бізнесі, науці чи державному управлінні. На додачу все частіше застосовують механізми CI/CD (Continuous Integration/ Continuous Deployment) у сфері DataOps, що дозволяє підтримувати безперервне оновлення та вдосконалення аналітичних модулів, забезпечуючи їх якість і узгодженість із вимогами системи в динамічному середовищі [8].

Ця послідовність етапів набуває особливого значення в контексті розподілених і хмарних середовищ, де великі обсяги даних надходять у режимі реального часу з численних географічно розподілених джерел. У таких умовах важливо враховувати балансування навантажень, забезпечувати стійкість у разі відмов окремих вузлів і використовувати інструменти автошардингу в базах даних. Окрім цього, постає потреба в інтеграції з технологіями контейнеризації (Docker) та оркестрації (Kubernetes), які дозволяють гнучко масштабувати обчислювальні ресурси під поточні вимоги аналітичних процесів.

Зі свого боку, перехід до концепції Data Lake на додаток до звичних сховищ даних (Data Warehouse) суттєво змінює підхід до ETL-процесів, адже в Data Lake дані часто спершу зберігаються у «сирому» (raw) вигляді, а їх комплексна підготовка і фільтрація здійснюється вже під конкретне завдання. Такий підхід полегшує роботу з різноманітними типами даних (JSON, CSV, зображення, відео, текстові документи тощо) і дозволяє швидше запускати «proof-of-concept» проєкти, проте вимагає чіткіших методів організації метаданих та класифікації інформації, аби уникнути надмірного «засмічення» сховища [5].

Ключовим викликом залишається вибір методів і механізмів автоматизації та контролю якості на кожному із зазначених кроків. Чимало компаній використовують підходи DataOps, що поєднують методології DevOps із практиками обробки даних. Зокрема, йдеться про системи моніторингу (Prometheus, Grafana), pipelines для безперервної інтеграції та доставки (Jenkins, GitLab CI/CD) і стандартизовані процедури тестування аналітичних моделей, які мінімізують ризики помилок та скорочують час від ідеї до реального впровадження.

Зрештою, ефективне дотримання вищезгаданих принципів і технологій формує міцну основу для комплексних рішень, які не лише обробляють великі обсяги інформації, а й здатні генерувати достовірні прогнозні моделі, автоматизовані рекомендаційні системи й інші інтелектуальні сервіси. Це, своєю чергою, відкриває нові перспективи впровадження штучного інтелекту та машинного навчання в різних галузях: від фінансових і маркетингових досліджень до медицини, містобудування та агросектору.

Із впровадженням інтелектуальних систем у процеси обробки даних пов’язані, з одного боку, суттєві переваги: прискорення аналітики великих масивів інформації, можливість знаходити приховані закономірності в даних і формувати висновки та поради, які враховують конкретні цілі чи вимоги.

З іншого боку, з'являються вагомі виклики, пов'язані насамперед із потребою в потужних обчислювальних ресурсах: масштабованих хмарних платформах, кластерах чи серверних фермах, які можуть динамічно задовольняти запити аналітичних систем. Крім цього, постають проблеми безпеки та захисту приватності при обробці чутливих або персональних даних, що вимагає від розробників і замовників особливої уваги до законодавчих та етичних норм. Не менш гостро стоїть і кадрове питання, оскільки індустрія відчуває брак спеціалістів, здатних органічно поєднувати компетенції у сфері машинного навчання, аналізу даних, системної архітектури та розробки програмного забезпечення [4].

Зважаючи на це, перед організаціями постає необхідність розробляти **комплексні стратегії** впровадження інтелектуальних систем, що охоплюють не лише технічний аспект, а й організаційно-правові та кадрові чинники. Ідеться про формування політик безпеки даних, планування інфраструктури під періодичні пікові навантаження, а також налагодження партнерських відносин між бізнесом, науковими установами й урядовими інституціями. Така взаємодія дозволить стимулювати освітні програми та навчальні проєкти, спрямовані на підготовку висококваліфікованих IT-фахівців.

Перспективним напрямом залишається **DataOps**, що поєднує методології DevOps із принципами обробки даних, забезпечуючи безперервний цикл розвитку й удосконалення аналітичних рішень. У контексті інтелектуальних систем це означає постійне оновлення моделей машинного навчання, систематичний збір нових даних і проведення тестів на узгодженість результатів. Окрім того, дедалі важливішим стає впровадження прозорих і етичних підходів до використання персональних даних і алгоритмів штучного інтелекту, що допоможе підвищити довіру як у суспільстві загалом, так і серед бізнес-партнерів [10].

Таким чином, перехід на новий рівень аналітики та автоматизації завдяки інтелектуальним системам приносить відчутні вигоди для бізнесу, наукових досліджень і державного управління. Проте реалізація цих можливостей вимагає зваженого підходу до вибору технологій, організації процесів і підготовки фахівців. Саме системна взаємодія між усіма зацікавленими сторонами — IT-компаніями, університетами, урядовими структурами й бізнес-організаціями — здатна створити умови, за яких сучасні методи обробки та аналізу даних перетворяться на реальний каталізатор розвитку цифрових інновацій у масштабах держави та світової спільноти.

Таким чином, запровадження хмарних обчислень, технологій контейнеризації, а також широке використання машинного навчання (ML) й інтелектуального аналізу даних (Data Mining) складають один із найважливіших векторів розвитку сучасних IT-технологій. Завдяки цим підходам вдається не лише скоротити час пошуку й опрацювання релевантної інформації, а й забезпечити оптимізацію бізнес-процесів та впровадження інноваційних

рішень із високою доданою вартістю в різних секторах економіки та наукових дослідженнях. У перспективі подальшого прогресу в цьому напрямі першорядне значення матиме розробка ще більш універсальних і гнучких інтелектуальних систем, здатних обробляти дедалі більші обсяги неоднорідних даних у реальному часі й пропонувати ефективні інструменти для ухвалення стратегічних рішень. Це, своєю чергою, передбачає продовження інтеграції передових методів штучного інтелекту, розширення можливостей автоматизації аналітики (AutoML) та активне залучення практик DataOps, що дасть змогу зменшити розрив між потребами бізнесу та наукового середовища у високоякісних, масштабованих і надійних IT-рішеннях.

Висновки. Проведене дослідження підтверджує, що сучасні інформаційні технології, зокрема хмарні обчислення, технології контейнеризації, а також методи машинного навчання і Data Mining, є ключовими інструментами для ефективного опрацювання великих обсягів даних. Завдяки впровадженню цих підходів вдалося досягти своєчасного виявлення релевантної інформації, що сприяє оптимізації бізнес-процесів та підвищенню якості ухвалення рішень у режимі реального часу.

Розглянуті концептуальні моделі (CRISP-DM та SEMMA) доводять свою ефективність, забезпечуючи послідовну обробку даних — від їхнього збору та інтеграції до побудови, оцінки й візуалізації аналітичних моделей. Такий підхід дозволяє не тільки зменшити обчислювальні витрати, але й підвищити точність аналізу завдяки автоматизації процесів і використанню передових алгоритмів.

Проте дослідження також виявило низку викликів, серед яких потреба в потужних обчислювальних ресурсах, забезпеченні безпеки та конфіденційності даних, а також нестача кваліфікованих спеціалістів, здатних інтегрувати комплексні системи штучного інтелекту в реальне виробництво. У зв’язку з цим, подальший розвиток в цьому напрямі має зосереджуватися на впровадженні методологій DataOps і AutoML, що сприятимуть більшій адаптивності та ефективності аналітичних систем.

Отже, інтеграція хмарних технологій, контейнеризації, ML та Data Mining створює міцну основу для побудови сучасних інформаційних систем, що здатні задовольнити зростаючі вимоги до обробки даних у різних галузях економіки та науки. Результати дослідження мають практичне значення для впровадження інноваційних рішень, які сприятимуть стратегічному розвитку підприємств, удосконаленню управлінських процесів і підвищенню конкурентоспроможності організацій у цифрову епоху.

Література:

1. Atif M**.** Data mining / M. Atif // *International Journal of Communication and Information Technology*. 2022. Vol. 3, No. 1. P. 37–40.DOI:10.33545/2707661X.2022.v3.i1a.44

2. Bzai, J.; Alam, F.; Dhafer, A.; Bojović, M.; Altowaijri, S.M.; Niazi, I.K.; Mehmood, R. Machine Learning-Enabled Internet of Things (IoT): Data, Applications, and Industry Perspective. Electronics 2022. №11.

3. Вакшинська Н. Ю., Щандрівська О. Є. Специфіка розвитку ринку BIG DATA для потреб відновлення економіки України в поствоєнний період. *Менеджмент та підприємництво в Україні: етапи становлення і проблеми розвитку.* 2023. № 2 (9). С.244-256.

4. Гордійчук-Бублівська О. В. Методи та засоби опрацювання великих даних в розподілених інформаційних системах : дис. д-ра філософії : 122 – Комп'ютерні науки / О. В. Гордійчук-Бублівська. – 2024. – 150 с.

5. Ковальова Т. Ю. Математичні моделі та методи DATA MINING для аналізу поведінки студентів у веб-базованих освітніх системах. *Радіоелектроніка та молодь у XXI столітті :* матеріали 28-го Міжнар. молодіж. форуму, 16-18 квітня 2024 р. – Харків : ХНУРЕ, 2024. – Т. 7. – С. 214-215.

6. Марченко О.О., Россада Т.В. Актуальні проблеми Data Mining: Навчальний посібник для студентів факультету комп'ютерних наук та кібернетики. Київ. 2017. 150 с.

7. Мушеник І.М., Марчук Н.А. Інформаційні технології та бази даних як інструмент для використання надлишкової інформації в задачах стратегічного управління логістичними підприємствами. *Наука і техніка сьогодні.* 2025. № 2(43).

8. Пархоменко О. Ю. Застосування хмарних обчислень та великих даних для оптимізації ланцюгів постачання в агропромисловому секторі. *Продовольча безпека України в умовах війни і післявоєнного відновлення: глобальні та національні виміри.* (м. Миколаїв, 30-31 травня 2024 р.) / Міністерство освіти і науки України; Миколаївський національний аграрний університет. Миколаїв : МНАУ, 2024. С. 319-321. DOI: https://doi.org/10.31521/978-617-7149-78-0-105.

9. Шандрівська О. Є., Кириленко А. А. Особливості ідентифікації ризиків ринку BIG DATA. *Менеджмент та підприємництво в Україні: етапи становлення і проблеми розвитку.* 2021. 3 (1), 82 – 95.

10. Ясінецька І.А., Мушеник І.М. Використання блокчейну для забезпечення прозорості та безпеки земельних правочинів. *Інноваційний потенціал сучасної освіти та науки: зб. матеріалів І Міжнародної наук.-практ.інтернет-конф.,* (25 квітня 2024 року. Кам'янець-Подільський). - Кам'янець-Подільський Заклад вищої освіти «Подільський державний університет», 2024. С.133-137.

References:

1. Atif, M. (2022). Data mining. International Journal of Communication and Information Technology, 3(1), 37–40.

2. Bzai, J., Alam, F., Dhafer, A., Bojović, M., Altowaijri, S. M., Niazi, I. K., & Mehmood, R. (2022). Machine Learning-Enabled Internet of Things (IoT): Data, Applications, and Industry Perspective. Electronics, 11.

3. Vakshynska, N. Yu., & Shandrivska, O. Ye. (2023). Spetsyfika rozvytku rynku BIG DATA dlia potreb vidnovlennia ekonomiky Ukrainy v postvoiennyi period [The specifics of the development of the BIG DATA market for the needs of Ukraine's economic recovery in the post-war period]. Menedzhment ta pidpryiemnytstvo v Ukraini: etapy stanovlennia i problemy rozvytku – Management and Entrepreneurship in Ukraine: Stages of Formation and Development Issues, 2(9), 244–256 [in Ukrainian].

4. Hordiichuk-Bublivska, O. V. (2024). Metody ta zasoby opratsiuvannia velykykh danukh v rozpodilenykh informatsiinykh systemakh [Methods and tools for big data processing in distributed information systems]. Doctoral thesis. [in Ukrainian].

5. Kovaliova, T. Yu. (2024). Matematychni modeli ta metody DATA MINING dlia analizu povedinky studentiv u web-bazovanykh osvitnikh systemakh [Mathematical models and methods of DATA MINING for analyzing student behavior in web-based educational systems]. Proceedings from REYM '24: 28th International Youth Forum «Radioelectronics and Youth in the 21st Century» (pp. 214–215). Kharkiv: KhNURE [in Ukrainian].

6. Marchenko, O. O., & Rossada, T. V. (2017). Aktualni problemy Data Mining [Current issues of Data Mining]. Kyiv: [Publisher not specified], 150 p. [in Ukrainian].

7. Mushenyk, I. M., & Marchuk, N. A. (2025). Informatsiini tekhnolohii ta bazy danukh yak instrument dlia vykorystannia nadlyshkovoi informatsii v zadachakh stratehichnoho upravlinnia lohistychnymy pidpryiemstvamy [Information technologies and databases as a tool for utilizing redundant information in strategic management tasks of logistics enterprises]. Nauka i tekhnika sohodni – Science and Technology Today, 2(43) [in Ukrainian].

8. Parkhomenko, O. Yu. (2024). Zastosuvannia khmarnykh obchyslen ta velykykh danukh dlia optymizatsii lantsiuhiv postachannia v ahropromyslovomu sektori [Application of cloud computing and big data for optimizing supply chains in the agro-industrial sector]. Proceedings from *Food Security of Ukraine in Wartime and Post-War Recovery: Global and National Dimensions* (Mykolaiv, May 30–31, 2024) (pp. 319–321). Mykolaiv: MNAU [in Ukrainian].

9. Shandrivska, O. Ye., & Kyrylenko, A. A. (2021). Osoblyvosti identyfikatsii ryzykiv rynku BIG DATA [Peculiarities of risk identification in the BIG DATA market]. *Menedzhment ta* pidpryiemnytstvo v Ukraini: etapy stanovlennia i problemy rozvytku – Management and Entrepreneurship *in Ukraine: Stages of Formation and Development Issues*, 3(1), 82–95 [in Ukrainian].

10. Yasinetska, I. A., & Mushenyk, I. M. (2024). Vykorystannia blokcheinu dlia zabezpechennia prozorosti ta bezpeky zemelnykh pravochnyniv [Using blockchain to ensure transparency and security of land transactions]. Proceedings from *Innovative Potential of Modern Education and Science: Collection of Materials of the I International Scientific and Practical Internet Conference* (April 25, 2024, Kamianets-Podilskyi) (pp. 133–137). Kamianets-Podilskyi: Podilskyi State University [in Ukrainian].