

АНАЛІТИКА BIG DATA ДЛЯ КРІ ТРАНСПОРТНИХ МЕРЕЖ: МЕТОДИ ЗБОРУ, ОЧИЩЕННЯ, ВІЗУАЛІЗАЦІЇ ТА GOVERNANCE

Дмитрієва О. І., д-р екон. наук, професор

Пресич Д.І., здобувач PhD

Харківський національний автомобільно-дорожній університет

Постановка проблеми. Сфера транспорту сьогодні швидко переходить у цифровий формат. Дані, які генеруються транспортними системами, стають головним ресурсом для планування, моніторингу та управління перевезеннями. Їхній обсяг постійно зростає — від показників руху громадського транспорту і навігаційних сигналів до інформації з датчиків інфраструктури, мобільних застосунків і сервісів логістики. Цей потік даних відкриває широкі можливості для глибокої аналітики, але водночас створює нові виклики, пов'язані з якістю, узгодженістю та практичним використанням інформації.

Складність полягає в тому, що більшість транспортних даних збираються розрізненими системами, у різних форматах і з різним рівнем деталізації. Частина інформації дублюється або містить похибки, а окремі показники не піддаються автоматичному аналізу без попереднього очищення. У результаті значний обсяг потенційно корисних даних залишається невикористаним. Для того щоб перетворити їх на основу ефективного управління, потрібні системні підходи до збору, оброблення, структурування й зберігання даних, що належать до різних рівнів транспортної інфраструктури.

Оцінювання ефективності транспортних мереж спирається на систему ключових показників (KPI), які дозволяють визначати не лише поточний стан системи, а й тенденції її розвитку. Проте впровадження таких індикаторів вимагає надійних, порівнюваних і своєчасних даних. Без аналітичних інструментів, здатних працювати з великими масивами інформації в режимі реального часу, цей процес залишається фрагментарним. Саме тому в центрі уваги сучасних досліджень опиняються технології Big Data, які поєднують методи збору, очищення, інтеграції та візуалізації даних у єдиній аналітичній системі.

Окремої уваги потребує аспект управління даними — так зване data governance. Для транспортної галузі це означає формування чітких правил відповідальності, доступу, безпеки та повторного використання інформації. Без цих принципів навіть найсучасніші аналітичні платформи не забезпечать необхідного рівня довіри та узгодженості між державними структурами, перевізниками й користувачами.

Отже, актуальність дослідження визначається потребою у створенні методології аналітики Big Data, яка б дозволила оцінювати ефективність транспортних мереж за допомогою KPI, забезпечуючи при цьому якість, сумісність і керованість даних. Таке поєднання технологічних та управлінських

підходів є ключовою передумовою розвитку інтелектуальних транспортних систем і переходу до сталого управління мобільністю.

Аналіз останніх досліджень і публікацій. Проблеми цифровізації транспортних систем і використання великих даних для оцінювання їх ефективності розглядаються у численних наукових працях. У міжнародному контексті ця тематика пов'язується з побудовою інтегрованих просторів даних і розробленням нормативно-технічних механізмів обміну інформацією між учасниками транспортної екосистеми. Такі аспекти відображено у працях К. Заннат, Л. Ге, Д. Цікі-Костопулу, де акцент зроблено на методах збору, очищення та інтеграції транспортних даних, а також на викликах, пов'язаних із якістю та сумісністю інформації [6–8]. Європейська комісія визначає формування Європейського простору даних мобільності як стратегічний напрям цифрової політики ЄС, що передбачає стандартизацію метаданих, розвиток довірених посередників і єдиних протоколів доступу [4; 5]. Правову основу таких процесів задають Data Governance Act і Data Act, у яких закладено принципи прозорого доступу, відповідальності та справедливого використання даних [1; 2].

В українських дослідженнях проблематика Big Data у транспортному секторі розглядається переважно крізь призму цифрової трансформації, розвитку інтелектуальних транспортних систем та удосконалення логістичних процесів. У своїх працях О. Дмитрієва доводить, що ефективна транспортна політика потребує системного створення, накопичення та використання даних у моделюванні сценаріїв розвитку галузі, а також розроблення державних механізмів управління інноваційним розвитком інфраструктури [12; 13].

І. Саяпіна досліджує питання оптимізації структури реляційних баз даних транспортних інформаційних систем, зосереджуючи увагу на підвищенні швидкодії, достовірності та узгодженості даних для подальшої аналітики [14]. С. Ключев із колегами аналізує розвиток інтелектуальних транспортних систем, наголошуючи на важливості впровадження аналітичних інструментів і технологій штучного інтелекту у керування міськими транспортними потоками [15].

О. Марків у своїх роботах розглядає можливості застосування технологій Big Data у сфері електротранспорту, підкреслюючи значення цифрових платформ для підвищення енергоефективності та надійності інфраструктури [16]. Подібні підходи формують основу нової вітчизняної школи транспортної аналітики, у межах якої поєднуються технічні рішення зі стратегічним управлінням даними.

Узагальнення результатів показує, що українська наукова спільнота поступово переходить від опису окремих технологічних рішень до комплексного підходу, який поєднує технічні, управлінські та аналітичні аспекти роботи з транспортними даними. Це створює наукове підґрунтя для подальшої розробки єдиної методології аналітики Big Data у транспортній галузі.

Невирішені складові загальної проблеми. Попри значний прогрес у цифровізації транспортного сектору, система управління великими даними в

Україні залишається роз'єднаною. Більшість ініціатив реалізується у відомчих або локальних масштабах без інтеграції в єдину аналітичну інфраструктуру, що обмежує можливості комплексного моніторингу та оцінювання ефективності транспортних мереж.

Однією з ключових проблем є відсутність узгодженої методології збору, очищення та перевірки транспортних даних. Різні відомства й оператори застосовують несумісні формати та протоколи обміну, що призводить до дублювання й втрати достовірності інформації. Бракує також загальнодержавних стандартів якості даних і механізмів координації між учасниками ринку.

Не вирішеним залишається й питання управління даними (data governance). Відсутність чітких правил доступу, збереження й повторного використання інформації створює ризики фрагментації, знижує рівень довіри між постачальниками та споживачами даних і ускладнює впровадження аналітики KPI на національному рівні.

Отже, першочерговими завданнями є узгодження технічних, правових і організаційних аспектів роботи з транспортними даними, розроблення єдиних стандартів і формування національного простору транспортної аналітики, сумісного з європейськими принципами data governance.

Формулювання цілей статті. Метою статті є розроблення науково обґрунтованих підходів до аналітики великих даних (Big Data) у системі оцінювання ключових показників ефективності (KPI) транспортних мереж. Дослідження спрямоване на визначення методів збору, очищення, стандартизації та верифікації транспортних даних, а також на аналіз сучасних наукових підходів до їх використання у сфері управління транспортом. Особливу увагу приділено систематизації методів візуалізації інформації для моніторингу KPI і формуванню принципів впровадження ефективної моделі управління транспортними даними (data governance) відповідно до європейських нормативних вимог.

Виклад основного матеріалу досліджень. Сучасні транспортні системи функціонують у середовищі постійного генерування великих обсягів даних, що надходять із різномірних джерел. Збирання інформації здійснюється як у режимі реального часу, так і з використанням накопичених архівів, що дозволяє формувати комплексну картину стану транспортної мережі. В основі цього процесу лежать три ключові напрями: автоматизований збір через сенсорні системи та IoT-пристрої, моніторинг на базі GPS-технологій і облік адміністративних даних з корпоративних систем.

Автоматизовані сенсорні системи забезпечують найвищу частоту оновлення даних, що робить їх основою для моделювання транспортних потоків. Вони охоплюють детектори руху, відеокамери, індукційні петлі, датчики навантаження та погодні сенсори. Такі дані є найбільш інформативними для оцінки поточної ситуації, але часто мають проблеми синхронізації або втрату сигналу [7].

Дані з GPS-трекерів та навігаційних платформ формують другу за обсягом категорію. Вони фіксують траєкторії руху транспортних засобів,

швидкість, час затримки, відхилення від маршруту. Їхня перевага — доступність у режимі реального часу, а недолік — різна точність позиціонування, особливо у міських умовах із щільною забудовою [6].

Третю групу становлять дані з корпоративних систем — CRM, ERP, білінгових та облікових платформ підприємств-перевізників. Вони містять відомості про розклади, квиткові операції, обсяги перевезень і витрати палива. Ці дані мають високу деталізацію, однак залежать від людського фактора і часто оновлюються із запізненням [14].

Узагальнена типологія транспортних даних наведена в таблиці 1.

Таблиця 1 – Типи джерел транспортних даних та їхня характеристика

Тип джерела даних	Характеристика	Орієнтовна частка у загальному потоці, %	Типові проблеми якості
Сенсорні системи (IoT, відео, датчики)	Автоматична фіксація руху, стану доріг, навантаження	60–70	Нестача синхронізації, пропуски сигналів
GPS-трекери та навігаційні пристрої	Дані про маршрути, швидкість, затримки, відхилення	20–25	Помилки позиціонування, різна точність
Ручна звітність, CRM/ERP системи	Дані підприємств-перевізників, розклади, облік палива	10–15	Людський фактор, несвоєчасне оновлення

Побудовано автором на основі [6; 7; 8; 14].

Типологічна класифікація транспортних даних дозволяє визначити не лише їхні джерела, а й ступінь надійності та придатності до подальшого аналізу. Для ефективного використання таких даних у системі КРІ важливо забезпечити баланс між частотою оновлення, повнотою охоплення та стабільністю якості. Саме ці параметри надалі стають визначальними для процесів очищення, інтеграції та побудови аналітичних моделей.

Процеси роботи з великими транспортними даними охоплюють повний життєвий цикл — від отримання інформації до її стандартизації та інтеграції в аналітичні системи. У сучасній практиці домінує модель ETL (Extract–Transform–Load), що передбачає послідовне вилучення, перетворення й завантаження даних до централізованого сховища. Такий підхід дає змогу узгодити інформацію з різних джерел — GPS-трекерів, сенсорів, білінгових систем, транспортних реєстрів — у єдиному форматі [6; 7].

Альтернативою традиційним сховищам даних стають Data Lake-платформи, які забезпечують зберігання сирих (неструктурованих) даних і дають змогу гнучко формувати аналітичні вибірки. Для транспортної аналітики це особливо важливо, адже потоки даних різномірні за частотою, типом і структурою [8; 14].

Проблема якості даних залишається ключовим бар'єром. За оцінками європейських досліджень, до 30 % транспортних записів містять дублювання або помилки позиціонування, а до 20 % — неповні часові мітки [9]. В

українських системах типові труднощі пов'язані з відсутністю уніфікованих стандартів форматів обміну, обмеженим застосуванням ISO 14813, DATEX II та GTFS-специфікацій, а також із браком механізмів автоматичної верифікації [15].

Поступове впровадження європейських практик — зокрема, застосування відкритих форматів даних та алгоритмів очищення на основі машинного навчання — дозволяє знизити обсяг шуму, скоротити час попередньої обробки та підвищити достовірність показників KPI. Досвід пілотних проєктів в Україні (електронна товарно-транспортна накладна, відкрита статистика дорожньої інфраструктури) показує, що навіть часткова автоматизація процесів очищення підвищує ефективність використання Big Data для транспортного планування [12; 16].

Візуалізація даних є ключовим інструментом для оцінки ефективності транспортних мереж. За допомогою сучасних технологій візуалізації, таких як дашборди та інтерактивні аналітичні панелі, можливо оперативно відображати основні показники ефективності (KPI), що дозволяє швидко приймати управлінські рішення на основі достовірних даних.

Для транспортної галузі найважливіші KPI включають середній час поїздки, рівень завантаження транспорту, точність прогнозування часу прибуття, кількість заторів на маршрутах та інші параметри, що безпосередньо впливають на якість перевезень і надання послуг. У результаті використання аналітики Big Data можна помітно підвищити точність прогнозів і зменшити час очікування для пасажирів.

Для демонстрації можливостей візуалізації KPI наведено умовну динаміку показників транспортної ефективності на основі узагальнених даних джерел [9; 15; 16] (рис 1). Це показує динаміку зміни одного з таких KPI — середнього часу поїздки після впровадження Big Data аналітики у транспортну систему. Як видно з графіка, середній час поїздки знизився на 18 % за три роки завдяки застосуванню інтелектуальних транспортних систем, використанню більш точних даних про рух і завантаженість шляхів, а також удосконаленню алгоритмів прогнозування.

Такі візуалізації допомагають не лише аналізувати стан транспортної мережі, а й передбачати її зміни, сприяючи розвитку «розумних» міст і раціоналізації транспортних потоків. Використання інтерактивних візуалізацій на дашбордах забезпечує зручний доступ до даних і дає змогу здійснювати їхню детальну перевірку в реальному часі. Це значно полегшує процес прийняття рішень у рамках міського або міжміського транспортного планування.

Ефективне управління транспортними даними неможливе без чіткої системи **data governance**, що охоплює принципи організації, відповідальності та контролю за використанням інформаційних ресурсів. У транспортній галузі ці принципи мають особливе значення, оскільки дані створюються численними учасниками — державними структурами, операторами перевезень, міськими сервісами та приватними платформами, — а їхня якість безпосередньо впливає на безпеку та ефективність перевезень.

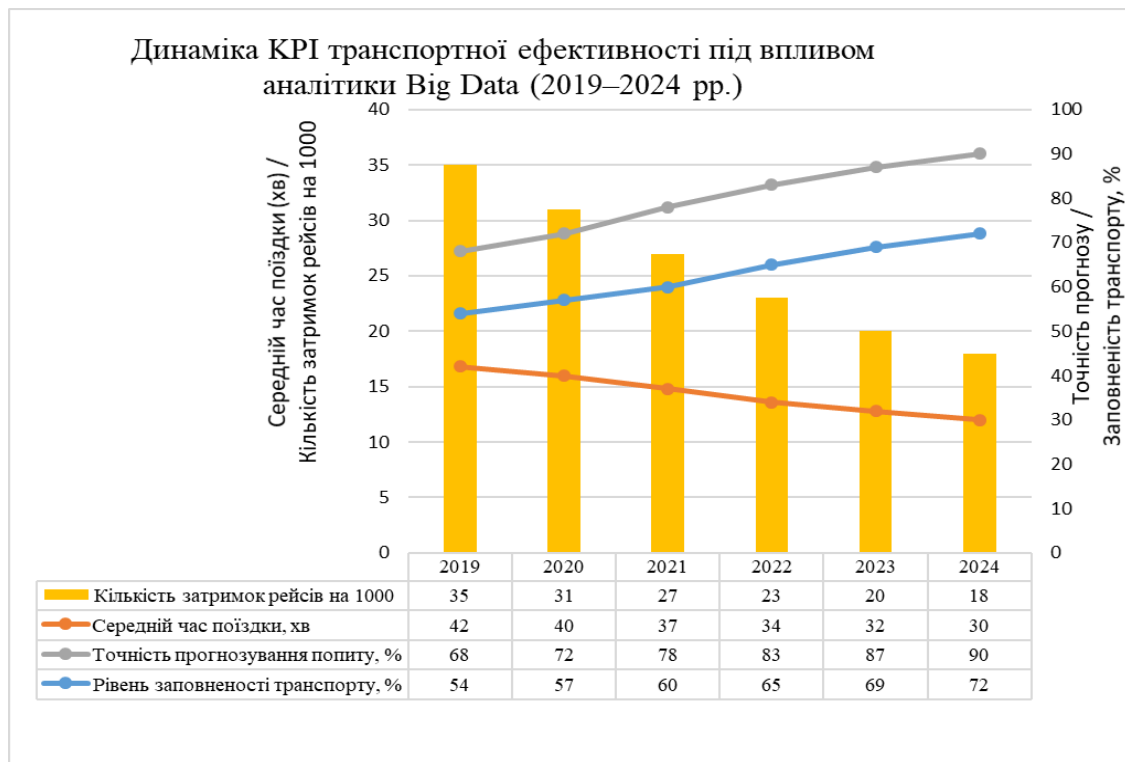


Рис. 1. Зміна середнього часу поїздки в транспортній мережі завдяки Big Data аналітиці

Джерело: побудовано автором на основі даних [9; 15; 16].

Примітка: показники наведено для ілюстрації узагальнених тенденцій, виявлених у джерелах [9; 15; 16], які демонструють вплив Big Data-аналітики на середній час поїздки, точність прогнозування та рівень заповненості транспорту.

Європейський підхід до управління даними визначено регламентами **Data Governance Act (DGA)** та **Data Act**, які закладають основи для створення прозорості, довірчої та інтероперабельної інфраструктури обміну інформацією між секторами. Ці акти встановлюють вимоги до сертифікації посередників даних, захисту комерційної та персональної інформації, а також до забезпечення повторного використання даних у публічних цілях [1; 2; 4].

У контексті транспортної аналітики концепція **data governance** передбачає визначення ролей і повноважень усіх учасників обміну: **власники даних** встановлюють політику доступу та збереження; **постачальники даних** відповідають за їхню якість і технічну сумісність; **аналітичні платформи** здійснюють інтеграцію, обробку та візуалізацію інформації;

— **органи нагляду** гарантують дотримання правових і етичних норм [5; 8].

В українських умовах питання управління транспортними даними поки не врегульоване належним чином. Відсутність централізованої інституції, що координує політику data governance, призводить до дублювання функцій та нерівномірності доступу до даних. Водночас позитивним кроком є розвиток **Національної транспортної стратегії України до 2030 року**, де передбачено створення цифрової екосистеми транспортних сервісів і впровадження принципів прозорого обміну даними між державою та бізнесом [12; 13].

Запровадження національної моделі **data governance** на основі європейських принципів створить умови для формування **єдиного простору транспортної аналітики**, у межах якого можлива інтеграція публічних і приватних даних, підвищення довіри до інформаційних процесів і забезпечення ефективного управління KPI транспортних мереж.

Аналітика великих даних у транспортному секторі спирається на широкий спектр технологічних рішень, які забезпечують збирання, зберігання, оброблення та інтерпретацію інформації у масштабних обсягах. Залежно від функцій, ці інструменти можна поділити на п'ять основних груп: Big Data-платформи, IoT-рішення, технології штучного інтелекту (AI), блокчейн-технології та логістичні аналітичні платформи.

Big Data-платформи, такі як Hadoop, Spark або Flink, застосовуються для паралельної обробки великих масивів транспортних даних, зокрема маршрутних та телеметричних потоків. Рішення IoT інтегрують дані з сенсорів, трекерів і розумних пристроїв, створюючи основу для моніторингу транспортних процесів у реальному часі [6; 7]. Алгоритми штучного інтелекту використовуються для прогнозування попиту, розрахунку оптимальних маршрутів і виявлення аномалій у транспортних потоках [9; 15].

Блокчейн-технології дедалі частіше впроваджуються у сфері транспортної логістики для забезпечення прозорості транзакцій, автентичності даних і відстеження ланцюгів постачання [10]. Логістичні аналітичні платформи — такі як Tableau, Power BI, Google Data Studio — виконують роль інструментів візуалізації та управління KPI, дозволяючи приймати рішення на основі інтегрованої аналітики.

Таблиця 2. – Класифікація інструментів Big Data-аналітики у транспортних системах

Категорія технологій	Приклади платформ / інструментів	Основні функції	Очікуваний ефект
Big Data-платформи	Hadoop, Spark, Flink	Масова паралельна обробка даних	Прискорення аналітичних розрахунків
IoT-рішення	MQTT, AWS IoT, Azure IoT Hub	Моніторинг сенсорних даних у реальному часі	Підвищення точності та оперативності даних
Технології AI/ML	TensorFlow, Scikit-learn	Прогнозування трафіку, виявлення аномалій	Оптимізація маршрутів і навантаження
Blockchain	Hyperledger, Ethereum	Прозорий облік і відстеження даних	Підвищення довіри та безпеки обміну
Логістичні аналітичні платформи	Tableau, Power BI, Data Studio	Візуалізація KPI та інтерактивна аналітика	Підтримка управлінських рішень

Побудовано автором на основі [6; 9; 10; 15; 16].

Застосування цих інструментів у поєднанні з принципами **data governance** створює умови для формування цілісної аналітичної екосистеми, у якій технологічні рішення підсилюють один одного: Big Data забезпечує масштабність, IoT — актуальність, AI — прогностичність, а блокчейн — довіру. Це відкриває можливість переходу до моделі **data-driven mobility**, коли рішення щодо розвитку транспортної інфраструктури ухвалюються на основі реальних аналітичних показників.

Проведений аналіз засвідчив, що аналітика великих даних у транспортних мережах є комплексною системою, яка поєднує технічні, управлінські й правові компоненти. Її ефективність визначається не лише технологічними можливостями, а й якістю даних, дотриманням принципів data governance і наявністю інтегрованої інфраструктури для оброблення та візуалізації інформації. Використання сучасних інструментів Big Data, IoT, AI та блокчейну формує основу для переходу до моделі data-driven управління транспортом, де ключові рішення приймаються на основі достовірних аналітичних показників KPI. Ці результати створюють наукове підґрунтя для подальших практичних досліджень і впровадження системної аналітики у вітчизняну транспортну політику.

Висновки з проведеного дослідження. Проведене дослідження дозволило досягти поставленої мети та реалізувати завдання, спрямовані на систематизацію підходів до використання великих даних (Big Data) у транспортних мережах. У роботі визначено основні методи збору, очищення, інтеграції та візуалізації транспортних даних, а також проаналізовано принципи управління ними в контексті сучасних вимог data governance.

Узагальнення наукових і практичних джерел показало, що ефективність транспортної аналітики безпосередньо залежить від якості та сумісності даних, застосування відкритих стандартів обміну й автоматизації процесів оброблення. Розглянуті моделі ETL і Data Lake довели свою придатність для інтеграції різноманітних джерел інформації, тоді як використання технологій IoT, AI та Big Data-платформ забезпечує високу швидкість та точність оцінювання KPI транспортних мереж.

Встановлено, що ключовими умовами підвищення ефективності аналітики є розбудова національної інфраструктури транспортних даних, уніфікація форматів і створення системи управління якістю на основі принципів data governance. Український досвід свідчить про поступовий перехід від локальних ініціатив до інтегрованої моделі аналітики, проте цей процес потребує інституційної підтримки та узгодження з європейськими стандартами.

Узагальнені результати підтверджують, що ефективна аналітика Big Data для KPI транспортних мереж можлива лише за поєднання методів збору, очищення, візуалізації та принципів data governance, що забезпечує повний цикл управління транспортними даними.

Результати дослідження формують наукове підґрунтя для розроблення єдиної методології Big Data-аналітики в транспортній сфері, здатної забезпечити підвищення точності моніторингу, прогнозування попиту та прийняття управлінських рішень на основі достовірних даних. Подальші дослідження доцільно спрямувати на розробку національної моделі аналітичної взаємодії між

державними та приватними учасниками транспортного ринку з урахуванням принципів прозорості, безпеки та стійкого розвитку.

Перелік посилань

1. Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act). *Official Journal of the European Union*. 2022. L 152. P. 1–44. URL: <https://eur-lex.europa.eu/eli/reg/2022/868/oj> (дата звернення: 23.10.2025).
2. Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data (Data Act). *Official Journal of the European Union*. 2023. URL: <https://eur-lex.europa.eu/eli/reg/2023/2854/oj/eng> (дата звернення: 23.10.2025).
3. Regulation (EU) 2020/1056 of the European Parliament and of the Council of 15 July 2020 on electronic freight transport information. *Official Journal of the European Union*. 2020. URL: <https://eur-lex.europa.eu/eli/reg/2020/1056/oj/eng> (дата звернення: 23.10.2025).
4. European Commission. Creating a common European mobility data space. URL: https://transport.ec.europa.eu/transport-themes/smart-mobility/creating-common-european-mobility-data-space_en (дата звернення: 23.10.2025).
5. European Commission. Mobility data policy page “Unlocking the potential of mobility data”. URL: <https://digital-strategy.ec.europa.eu/en/policies/mobility-data> (дата звернення: 23.10.2025).
6. Zannat K. E., Choudhury C. F. Emerging Big Data Sources for Public Transport Planning: A Systematic Review on Current State of Art and Future Research Directions. *Smart and Sustainable Built Environment*. 2019. DOI: 10.1007/s41745-019-00125-9. URL: <https://link.springer.com/article/10.1007/s41745-019-00125-9> (дата звернення: 23.10.2025).
7. Ge L., Fukuda D., Nakamura T., Kurauchi F. Review of Transit Data Sources: Potentials, Challenges, and Opportunities. *Sustainability*. 2021. 13(20):11450. DOI: 10.3390/su132011450. URL: <https://www.mdpi.com/2071-1050/13/20/11450> (дата звернення: 23.10.2025).
8. Tzika-Kostopoulou D., Mitroglou K., Polydoropoulou A. Big data in transportation: a systematic literature analysis and topic classification. *Knowledge and Information Systems*. 2024. DOI: 10.1007/s10115-024-02112-8. URL: <https://link.springer.com/article/10.1007/s10115-024-02112-8> (дата звернення: 23.10.2025).
9. Nesmachnow S., et al. Big Data Analysis for Travel Time Characterization in Public Transportation Systems. *Sustainability*. 2023. 15(19):14561. DOI: 10.3390/su151914561. URL: <https://www.mdpi.com/2071-1050/15/19/14561> (дата звернення: 23.10.2025).
10. Myronenko S., et al. A Case Study of Public Transport in Ukraine. *Applied Sciences*. 2023. 13(3):67. DOI: 10.3390/app13030067. URL: <https://www.mdpi.com/2624-6511/6/3/67> (дата звернення: 23.10.2025).
11. ISO 8000-1:2022. Data quality. Part 1: Overview. URL: <https://www.iso.org/standard/81745.html> (дата звернення: 23.10.2025).

12. Дмитрієва О. І. Оцінювання інноваційного розвитку транспортної інфраструктури. *Вісник Київського національного університету технологій та дизайну. Серія: Економічні науки*. 2019. № 5(139). С. 8–20. URL: <https://er.knutd.edu.ua/handle/123456789/15465> (дата звернення: 23.10.2025).
13. Дмитрієва О. І. Державне регулювання інноваційного розвитку транспортної інфраструктури. Дисертація д-ра екон. наук. Харківський національний автомобільно-дорожній університет. Харків. 2020. URL: https://kart.edu.ua/wp-content/uploads/2020/11/dmytriieva_dis.pdf (дата звернення: 23.10.2025).
14. Саяпіна І. І. Оптимізація роботи реляційної бази даних інформаційної системи на транспорті. *Transport Systems and Technologies*. 2022. № 40. С. 23–31. DOI: 10.32703/2617-9040-2022-40-23. УДК 004.65. URL: <https://tst.duit.in.ua/index.php/tst/article/view/359> (дата звернення: 23.10.2025).
15. Ключев С. О., Цимбал С. В., Сігонін А. Є. Розвиток інтелектуальних транспортних систем. *Вісник машинобудування та транспорту*. 2023. № 2(18). С. 80–86. DOI: 10.31649/2413-4503-2023-18-2-80-86. URL: <https://doi.org/10.31649/2413-4503-2023-18-2-80-86> (дата звернення: 23.10.2025).
16. Марків О., Ришковець Ю. Використання технологій Big Data в галузі електротранспорту. *Вісник Національного університету «Львівська політехніка»*. 2024. С. 423–433. URL: <https://science.lpnu.ua/sites/default/files/journal-paper/2024/aug/35692/maket2402951-423-433.pdf> (дата звернення: 23.10.2025).
17. ISO/IEC 38505-1:2017. Information technology — Governance of data — Part 1: Application of ISO/IEC 38500 to the governance of data. URL: <https://www.iso.org/standard/56639.html> (дата звернення: 23.10.2025).

References.

1. Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act), *Official Journal of the European Union, L 152*, 1–44, Retrieved October 23, 2025, available at : <https://eur-lex.europa.eu/eli/reg/2022/868/oj>
2. Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data (Data Act), *Official Journal of the European Union, L 374*. Retrieved October 23, 2025, available at : <https://eur-lex.europa.eu/eli/reg/2023/2854/oj/eng>
3. Regulation (EU) 2020/1056 of the European Parliament and of the Council of 15 July 2020 on electronic freight transport information, *Official Journal of the European Union*, Retrieved October 23, 2025, available at : <https://eur-lex.europa.eu/eli/reg/2020/1056/oj/eng>
4. European Commission, *Creating a common European mobility data space*, Retrieved October 23, 2025, available at : https://transport.ec.europa.eu/transport-themes/smart-mobility/creating-common-european-mobility-data-space_en
5. European Commission, *Mobility data policy page “Unlocking the potential of mobility data”*, Retrieved October 23, 2025, available at : <https://digital-strategy.ec.europa.eu/en/policies/mobility-data>

6. Zannat, K. E., & Choudhury, C. F. (2019), Emerging big data sources for public transport planning: A systematic review on current state of art and future research directions, *Smart and Sustainable Built Environment*. DOI: <https://doi.org/10.1007/s41745-019-00125-9>
7. Ge, L., Fukuda, D., Nakamura, T., & Kurauchi, F. (2021), Review of transit data sources: Potentials, challenges, and opportunities. *Sustainability*, 13(20), 11450. DOI: <https://doi.org/10.3390/su132011450>
8. Tzika-Kostopoulou, D., Mitroglou, K., & Polydoropoulou, A. (2024), Big data in transportation: A systematic literature analysis and topic classification, *Knowledge and Information Systems*. DOI: <https://doi.org/10.1007/s10115-024-02112-8>
9. Nesmachnow, S., et al. (2023), Big data analysis for travel time characterization in public transportation systems, *Sustainability*, 15(19), 14561. DOI: <https://doi.org/10.3390/su151914561>
10. Myronenko, S., et al. (2023), A case study of public transport in Ukraine, *Applied Sciences*, 13(3), 67. <https://doi.org/10.3390/app13030067>
11. International Organization for Standardization (2022), *ISO 8000-1:2022. Data quality. Part 1: Overview*, Retrieved October 23, 2025, available at : <https://www.iso.org/standard/81745.html>
12. Dmytriieva, O. I. (2019), Assessment of innovative development of transport infrastructure [Otsiniuvannia innovatsiinoho rozvytku transportnoi infrastruktury], *Bulletin of Kyiv National University of Technologies and Design. Series: Economic Sciences*, 5(139), 8–20. Retrieved October 23, 2025, available at : <https://er.knutd.edu.ua/handle/123456789/15465> [in Ukrainian]
13. Dmytriieva, O. I. (2020), State regulation of innovative development of transport infrastructure [Derzhavne rehuliuvannia innovatsiinoho rozvytku transportnoi infrastruktury] (Doctoral dissertation, Kharkiv National Automobile and Highway University), Retrieved October 23, 2025, available at : https://kart.edu.ua/wp-content/uploads/2020/11/dmytriieva_dis.pdf [in Ukrainian]
14. Saiapina, I. I. (2022), Optimization of a relational database of the transport information system [Optymizatsiia roboty reliatsiinoi bazy danykh informatsiinoi systemy na transporti], *Transport Systems and Technologies*, 40, P. 23–31. DOI: <https://doi.org/10.32703/2617-9040-2022-40-23> [in Ukrainian]
15. Kliuiev, S. O., Tsymbal, S. V., & Sihonin, A. Ye. (2023), Development of intelligent transport systems [Rozvytok intelektualnykh transportnykh system], *Visnyk Machine Building and Transport*, 2(18), P. 80–86. DOI: <https://doi.org/10.31649/2413-4503-2023-18-2-80-86> [in Ukrainian]
16. Markiv, O., & Ryshkovets, Yu. (2024), Application of Big Data technologies in the field of electric transport [Vykorystannia tekhnolohii Big Data v haluzi elektrotransportu], *Bulletin of Lviv Polytechnic National University*, P. 423–433, Retrieved October 23, 2025, available at : <https://science.lpnu.ua/sites/default/files/journal-paper/2024/aug/35692/maket2402951-423-433.pdf> [in Ukrainian]
17. International Organization for Standardization (2017), *ISO/IEC 38505-1:2017. Information technology — Governance of data — Part 1: Application of ISO/IEC 38500 to the governance of data*, Retrieved October 23, 2025, available at : <https://www.iso.org/standard/56639.html>

РЕФЕРАТИ ABSTRACTS

УДК 004.65:656; JEL Classification: L91, O33, C55, R41

Дмитрієва О.І., Пресич Д.І. АНАЛІТИКА BIG DATA ДЛЯ КРІ ТРАНСПОРТНИХ МЕРЕЖ: МЕТОДИ ЗБОРУ, ОЧИЩЕННЯ, ВІЗУАЛІЗАЦІЇ ТА GOVERNANCE

Мета. Дослідження спрямоване на формування цілісної моделі використання аналітики великих даних (Big Data) для оцінювання ключових показників ефективності (KPI) транспортних мереж. У роботі визначено методи збору, очищення, інтеграції та візуалізації транспортних даних, а також принципи управління ними відповідно до сучасних вимог data governance.

Методика дослідження. У процесі дослідження застосовано системний та аналітико-узагальнювальний методи для визначення структури транспортних даних і джерел їх надходження. Використано методи порівняльного аналізу — для зіставлення європейських і національних підходів до управління даними; метод логічного моделювання — для побудови схеми аналітичного циклу Big Data у транспортних мережах; статистико-аналітичний метод — для узагальнення результатів досліджень з оцінки впливу аналітики даних на показники ефективності перевезень. Емпіричною базою стали публікації Європейської комісії, наукові праці з питань Big Data у транспорті та українські практики цифровізації галузі.

Результати. Визначено основні джерела транспортних даних — сенсорні системи, GPS-трекери, білінгові платформи та корпоративні облікові системи. Систематизовано методи очищення даних (ETL, Data Lake) і виявлено типові проблеми їх якості — дублювання, відсутність часових міток, розбіжності форматів. Розроблено типологічну схему структури транспортних даних і запропоновано підхід до оцінювання їхньої надійності для розрахунку KPI. Наведено приклад візуалізації динаміки транспортної ефективності, що демонструє зниження середнього часу поїздки на 18 % після впровадження Big Data аналітики. Обґрунтовано, що впровадження принципів data governance підвищує рівень інтеграції транспортних даних і сприяє створенню єдиної аналітичної інфраструктури.

Наукова новизна. Полягає у системному підході до використання аналітики Big Data для моніторингу KPI транспортних мереж, що поєднує технічні, аналітичні та управлінські аспекти. Запропоновано структурну модель інтеграції процесів збору, очищення, візуалізації та управління транспортними даними в єдиному аналітичному контурі, адаптовану до українських умов.

Практична значущість. Результати дослідження можуть бути використані органами державного управління, міськими транспортними службами, операторами перевезень і IT-компаніями для побудови аналітичних систем управління транспортними потоками. Запровадження комплексної аналітики Big Data дозволяє підвищити точність прогнозів, ефективність планування маршрутів, знизити витрати ресурсів і забезпечити прозорість управлінських рішень у транспортній галузі України.

Ключові слова: Big Data, транспортна аналітика, KPI, data governance, транспортні мережі, очищення даних, ETL, візуалізація, інтелектуальні транспортні системи, цифрова інфраструктура.

UDC 004.65:656; JEL Classification: L91, O33, C55, R41

Dmytriieva O.I., Presych D.I. BIG DATA ANALYTICS FOR KPI OF TRANSPORT NETWORKS: METHODS OF COLLECTION, CLEANING, VISUALIZATION AND GOVERNANCE

Purpose. The purpose of the study is to develop an integrated model for applying Big Data analytics to assess key performance indicators (KPI) of transport networks. The paper defines the main methods of collecting, cleaning, integrating and visualizing transport data, as well as the principles of managing them in accordance with modern data governance requirements. **Research methodology.** The study employs systemic and analytical–generalization methods to identify the structure and sources of transport data. Comparative analysis was used to contrast European and national approaches to data management, while logical modeling was applied to construct the analytical cycle of Big Data in transport networks. Statistical and analytical methods were employed to summarize the results of studies evaluating the impact of data analytics on transport performance. The empirical basis includes publications of the European Commission, scientific papers on Big Data in transport, and Ukrainian practices of digitalization in this sector. **Results.** The main sources of transport data were identified as sensor systems, GPS trackers, billing platforms, and corporate information systems. Methods of data cleaning (ETL, Data Lake) were systematized, and the main quality issues — duplication, lack of time stamps, and format inconsistencies — were revealed. A typological scheme of transport data structure was developed, and an approach to assessing their reliability for KPI calculations was proposed. The study presents an example of visualization of transport efficiency dynamics, showing an 18% reduction in average travel time after the introduction of Big Data analytics. It is substantiated that the implementation of data governance principles enhances the integration level of transport data and supports the creation of a unified analytical infrastructure. **Scientific novelty.** The novelty lies in a systemic approach to applying Big Data analytics for monitoring transport network KPIs, which combines technical, analytical, and managerial dimensions. A structural model integrating processes of data collection, cleaning, visualization, and governance into a single analytical framework adapted to Ukrainian conditions has been proposed. **Practical significance.** The findings can be applied by public authorities, municipal transport agencies, operators, and IT companies in developing analytical systems for transport flow management. The introduction of comprehensive Big Data analytics improves forecast accuracy, enhances route planning efficiency, reduces resource costs, and ensures transparency of decision-making in Ukraine’s transport sector.

Keywords: Big Data, transport analytics, KPI, data governance, transport networks, data cleaning, ETL, visualization, intelligent transport systems, digital infrastructure.

Відомості про авторів / About the Authors

Дмитрієва Оксана Іллівна – завідувачка кафедри економіки та підприємництва, доктор економічних наук, професор, Харківський національний автомобільно-дорожній університет, професор кафедри економіки і підприємництва, м. Харків, Україна; e-mail: oksanahnadu@gmail.com; ORCID: <https://orcid.org/0000-0001-9314-350X>. Моб. (063) 353-79-98.

Dmytriieva Oksana – Head of the Department of Economics and Entrepreneurship, Doctor of Economics, Professor, Kharkiv National Automobile and Highway University, Professor of the Department of Economics and Entrepreneurship, Kharkiv, Ukraine.

Пресич Дмитро - здобувач PhD кафедри економіки та підприємництва, Харківський національний автомобільно-дорожній університет, м. Харків, Україна.

Presych Dmytro – Postgraduate student Department of Economics and Entrepreneurship, Kharkiv national automobile and highway university, Kharkiv, Ukraine.