

SCI-CONF.COM.UA

SCIENCE AND EDUCATION: SYNERGY OF INNOVATION



**PROCEEDINGS OF XIII INTERNATIONAL
SCIENTIFIC AND PRACTICAL CONFERENCE
AUGUST 10-12, 2026**

**BERLIN
2026**

SCIENCE AND EDUCATION: SYNERGY OF INNOVATION

Proceedings of XIII International Scientific and Practical Conference

Berlin, Germany

10-12 August 2026

Berlin, Germany

2026

UDC 001.1

The 13th International scientific and practical conference “Science and education: synergy of innovation” (August 10-12, 2026) MDPC Publishing, Berlin, Germany. 2026. 204 p.

ISBN 978-3-954753-01-7

The recommended citation for this publication is:

Ivanov I. Analysis of the phaunistic composition of Ukraine // Science and education: synergy of innovation. Proceedings of the 13th International scientific and practical conference. MDPC Publishing. Berlin, Germany. 2026. Pp. 21-27. URL: <https://sci-conf.com.ua/xiii-mizhnarodna-naukovo-praktichna-konferentsiya-science-and-education-synergy-of-innovation-10-12-08-2026-berlin-nimechchina-arhiv/>.

Editor

Komarytskyy M.L.

Ph.D. in Economics, Associate Professor

Collection of scientific articles published is the scientific and practical publication, which contains scientific articles of students, graduate students, Candidates and Doctors of Sciences, research workers and practitioners from Europe, Ukraine and from neighbouring countries and beyond. The articles contain the study, reflecting the processes and changes in the structure of modern science. The collection of scientific articles is for students, postgraduate students, doctoral candidates, teachers, researchers, practitioners and people interested in the trends of modern science development.

e-mail: berlin@sci-conf.com.ua

homepage: <https://sci-conf.com.ua>

©2026 Scientific Publishing Center “Sci-conf.com.ua” ®

©2026 MDPC Publishing ®

©2026 Authors of the articles

11.	<i>Довженко Т. П., Сторчак К. П.</i> НАСКРІЗНЕ ВИЯВЛЕННЯ ВІЗУАЛЬНИХ АНОМАЛІЙ ЗА ОБМЕЖЕНИХ НАВЧАЛЬНИХ ДАНИХ: ПОРІВНЯННЯ HYBRID AWRED V4 ІЗ CNN+ONE-CLASS SVM	63
12.	<i>Моторнюк С. О.</i> АВТОМАТИЗОВАНЕ ОЦІНЮВАННЯ ЯКОСТІ ВІДКРИТИХ ДАНИХ ЗАСОБАМИ ШТУЧНОГО ІНТЕЛЕКТУ	67
PHYSICAL AND MATHEMATICAL SCIENCES		
13.	<i>Тимошенко А. А.</i> ЧИСЕЛЬНИЙ ПІДХІД ДО АНАЛІЗУ ЦІНИ ІГРОВОГО NFT	74
GEOLOGICAL AND MINERALOGICAL SCIENCES		
14.	<i>Zimmerman Ava, Buynevich Ilya Val, Boles Zachary, Smierciak Krisley, Leipert Mark, Lacovara Kenneth</i> EXPERIMENTAL FOSSIL EXCAVATION BACKFILL SEQUENCE: ULTRA-HIGH RESOLUTION (2.3 GHz GEORADAR) IMAGING OF GLAUCONITE-RICH STRATA, NEW JERSEY, USA	83
PEDAGOGICAL SCIENCES		
15.	<i>Ананьян Е. Л.</i> ДИСКУРСИВНА КОМПЕТЕНТНІСТЬ ЯК ОСНОВА ФОРМУВАННЯ НАВИЧОК АКАДЕМІЧНОГО ПИСЬМА У ЗДОБУВАЧІВ ВИЩОЇ ОСВІТИ У ПРОЦЕСІ ВИВЧЕННЯ ІНОЗЕМНОЇ МОВИ ЗА ПРОФЕСІЙНИМ СПРЯМУВАННЯМ	88
16.	<i>Баиштова К. К.</i> МОДЕЛІ УПРАВЛІННЯ ОСВІТНІМ ПРОЦЕСОМ В УНІВЕРСИТЕТАХ ЛІГИ ПЛЮЩА	94
17.	<i>Велієва М. В., Гуренко А. Л.</i> DATA SCIENCE НА СТАДІОНІ: ВИКОРИСТАННЯ МАТЕМАТИЧНОЇ СТАТИСТИКИ ТА ФІТНЕС-ТРЕКЕРІВ ДЛЯ ОПТИМІЗАЦІЇ ФІЗИЧНОГО СТАНУ ЗДОБУВАЧІВ ОСВІТИ	99
18.	<i>Іотковська М. О.</i> ПЕРСПЕКТИВНІ НАПРЯМИ РОЗВИТКУ НАУКОВОЇ ДІЯЛЬНОСТІ КАФЕДР ДОШКІЛЬНОЇ ПЕДАГОГІКИ В СУЧАСНОМУ ОСВІТНЬОМУ ПРОСТОРІ: РЕГІОНАЛЬНИЙ КОНТЕКСТ ОДЕСЬКОЇ ТА ДОНЕЦЬКОЇ ОБЛАСТЕЙ	104
19.	<i>Куць Є. І., Овчаренко Л. Р.</i> ФОРМУВАННЯ ІНШОМОВНОЇ КОМУНІКАТИВНОЇ КОМПЕТЕНТНОСТІ СТАРШОКЛАСНИКІВ: МЕТОДИЧНИЙ АСПЕКТ	111

АВТОМАТИЗОВАНЕ ОЦІНЮВАННЯ ЯКОСТІ ВІДКРИТИХ ДАНИХ ЗАСОБАМИ ШТУЧНОГО ІНТЕЛЕКТУ

Моторнюк Сергій Олександрович

аспірант кафедри
кібербезпеки та інформаційних технологій
Харківський національний економічний
університет імені Семена Кузнеця
м. Харків, Україна

Вступ. / Introductions. Відкриті дані є одним із ключових ресурсів цифрового врядування, розвитку електронних сервісів, публічної аналітики та інноваційних рішень на основі штучного інтелекту. Їхнє значення особливо зростає в умовах переходу до врядування на основі даних, коли управлінські рішення дедалі частіше ґрунтуються на результатах аналізу великих масивів публічної інформації. Водночас сам факт оприлюднення набору даних на національному чи локальному порталі відкритих даних не гарантує його практичної придатності для повторного використання.

На практиці відкриті набори даних часто містять пропущені значення, дублікати, некоректні формати, неіменовані або слабо описані атрибути, структурні неузгодженості, застарілі ресурси, проблеми з доступністю файлів та недостатньо інформативні метадані. У сучасних дослідженнях якість відкритих даних розглядається як багатовимірна характеристика, що охоплює не лише точність і повноту, а й доступність, інтероперабельність, актуальність, семантичну узгодженість та придатність до повторного використання [1; 2].

Такі дефекти знижують достовірність подальшого аналізу, ускладнюють інтеграцію даних у цифрові платформи та можуть спричиняти помилки в алгоритмах машинного навчання. Особливо критичною ця проблема є для відкритих державних даних, які використовуються для моніторингу публічної політики, оцінювання ефективності управлінських рішень, екологічного контролю, аналізу соціально-економічних процесів та створення сервісів для громадян. Традиційні підходи до контролю якості відкритих даних переважно

базуються на ручній перевірці, формальній валідації форматів або експертному аналізі метаданих. Проте такі методи є трудомісткими, погано масштабуються та не завжди дозволяють виявляти приховані дефекти, зокрема аномальні значення або латентні структурні порушення. У зв'язку з цим актуальним є розроблення автоматизованих інструментів, які поєднують формальні правила перевірки, статистичне профілювання та методи штучного інтелекту для комплексного оцінювання якості відкритих даних [3].

Мета роботи. / Aim. Метою роботи є обґрунтування та практична реалізація підходу до автоматизованого оцінювання якості відкритих даних засобами штучного інтелекту на основі поєднання правил формальної валідації, статистичного аналізу та алгоритмів виявлення аномалій.

Для досягнення поставленої мети визначено такі завдання: узагальнити основні типи дефектів якості відкритих даних; сформувати архітектуру інструменту автоматизованого аналізу; реалізувати прототип системи оцінювання якості відкритих даних; апробувати запропонований підхід на реальних наборах даних, опублікованих на платформі відкритих даних; визначити практичну придатність системи для підтримки процесів управління якістю даних.

Матеріали та методи. / Materials and methods. У межах дослідження використано реальні набори відкритих даних, опубліковані на національній платформі відкритих даних України, що функціонує на базі SKAN. Платформа SKAN є одним із найпоширеніших програмних рішень для публікації, каталогізації та управління відкритими даними, що забезпечує доступ до наборів і ресурсів через програмні інтерфейси [4]. Обрані набори охоплювали різні предметні області, зокрема екологічні, соціально-інфраструктурні, адміністративні та економічні дані. Такий підхід дозволив перевірити працездатність запропонованого інструменту в умовах структурної неоднорідності та різного рівня підготовки даних.

Методичну основу дослідження становить data-centric підхід, відповідно до якого основна увага приділяється не лише побудові моделей штучного

інтелекту, а насамперед якості даних як вихідного ресурсу для подальшої аналітики. Запропонований підхід передбачає кілька послідовних етапів: завантаження даних із зовнішнього джерела, автоматичне профілювання, перевірка повноти й форматної валідності, виявлення структурних дефектів, статистичний аналіз числових ознак, ідентифікація аномальних спостережень та формування аналітичного звіту.

Для реалізації прототипу використано мову програмування Python та стандартні інструменти обробки табличних даних. Взаємодія з SKAN-платформою здійснюється через API, що дозволяє отримувати метадані наборів, перелік ресурсів і самі файли для подальшого аналізу. У межах системи передбачено можливість роботи як з окремими ресурсами, так і з набором даних загалом.

Оцінювання якості даних здійснюється за такими групами показників: повнота значень, коректність типів даних, наявність неіменованих або технічних стовпців, дублювання записів, доступність ресурсу, наявність статистично нетипових значень та загальна інтегральна оцінка якості. Для формального контролю застосовуються rule-based перевірки, тоді як для виявлення прихованих відхилень використовуються алгоритмічні методи аналізу аномалій. Подібне поєднання статистичних і машинно-навчальних підходів є перспективним для аналізу динамічних відкритих державних даних [5]. Узагальнену логіку функціонування запропонованого підходу наведено на мал. 1

Результати та обговорення. / Results and discussion. У результаті дослідження розроблено прототип інструменту автоматизованого оцінювання якості відкритих даних, який умовно можна представити як послідовний аналітичний конвеєр. На першому етапі система отримує посилання на набір даних або окремий ресурс. Далі виконується завантаження даних, визначення формату, перевірка доступності файлу та попереднє структурування. Якщо ресурс є недоступним або не може бути коректно оброблений, система фіксує це як окремий дефект якості, що має інфраструктурний характер.



Мал. 1. Архітектура автоматизованого оцінювання якості відкритих даних

На другому етапі здійснюється профілювання даних. Система автоматично визначає кількість рядків і стовпців, типи атрибутів, частку пропущених значень, кількість унікальних значень, базові статистичні характеристики числових полів і наявність потенційно проблемних стовпців. Особлива увага приділяється виявленню неіменованих атрибутів, які часто виникають унаслідок некоректного експорту таблиць із редакторів електронних таблиць або порушення структури CSV-файлів.

На третьому етапі застосовуються правила контролю якості. Зокрема, перевіряються надмірна частка пропусків, наявність повністю порожніх стовпців, дублювання записів, невідповідність очікуваних типів даних фактичним значенням, а також структурні порушення. Кожен виявлений дефект супроводжується коротким описом, зазначенням проблемного атрибуту

та рекомендацією щодо виправлення.

На четвертому етапі застосовується модуль виявлення аномалій. Його завдання полягає у фіксації нетипових записів, які не обов'язково порушують формальні правила, але статистично відрізняються від основної маси даних. Такий підхід є важливим, оскільки частина помилок у відкритих даних не має очевидного синтаксичного характеру. Наприклад, числове значення може бути формально коректним, але суттєво відхилятися від очікуваного діапазону або не відповідати логіці інших атрибутів.

Апробація прототипу на реальних відкритих наборах даних підтвердила здатність системи виявляти різні типи дефектів якості. Серед найпоширеніших проблем було зафіксовано пропущені значення в окремих атрибутах, наявність неіменованих стовпців, структурні порушення у файлах, недоступність окремих ресурсів, а також статистично аномальні записи. Отримані результати свідчать, що навіть формально опубліковані відкриті дані можуть мати обмежену аналітичну придатність без попереднього автоматизованого контролю.

Важливою перевагою запропонованого підходу є поєднання інтерпретованості rule-based перевірок із гнучкістю методів штучного інтелекту. Формальні правила дозволяють чітко пояснити виявлені дефекти користувачеві, тоді як алгоритми аналізу аномалій розширюють можливості системи щодо виявлення нетипових випадків. Таким чином, система виконує не лише функцію технічної перевірки, а й забезпечує формування аналітичного уявлення про придатність даних до повторного використання.

Результати аналізу подаються у вигляді структурованого звіту, який містить загальну оцінку якості, перелік проаналізованих ресурсів, кількісні характеристики набору, опис виявлених дефектів, приклади проблемних значень та рекомендації щодо покращення. Такий формат дозволяє використовувати звіт як для технічного аудиту, так і для управлінського прийняття рішень щодо пріоритизації наборів даних, які потребують доопрацювання.

Практична значущість запропонованого підходу полягає в тому, що він

може бути інтегрований у роботу порталів відкритих даних, зокрема платформ на базі SKAN. Це дозволяє перейти від періодичної ручної перевірки до регулярного автоматизованого моніторингу якості даних. У перспективі такий інструмент може використовуватися органами публічної влади, адміністраторами порталів відкритих даних, дослідниками та розробниками цифрових сервісів.

Висновки. / Conclusions. У роботі запропоновано підхід до автоматизованого оцінювання якості відкритих даних засобами штучного інтелекту, який поєднує формальні правила перевірки, статистичне профілювання та алгоритми виявлення аномалій. Розроблений прототип дозволяє автоматично аналізувати відкриті набори даних, виявляти дефекти повноти, структурної узгодженості, форматної валідності, доступності ресурсів та наявності аномальних значень.

Результати апробації на реальних наборах відкритих даних підтвердили практичну придатність запропонованої системи для первинного аудиту якості даних. Встановлено, що автоматизований аналіз дозволяє швидко локалізувати проблемні ресурси, сформувати кількісну оцінку їх якості та надати рекомендації щодо покращення. Це створює передумови для підвищення аналітичної придатності відкритих даних і розширення можливостей їх повторного використання.

Подальші дослідження доцільно спрямувати на розширення алгоритмів автоматизованого виправлення дефектів, удосконалення методів семантичного аналізу метаданих, інтеграцію explainable AI для пояснення результатів виявлення аномалій, а також створення повноцінного SKAN-плагіна для регулярного моніторингу якості відкритих даних у публічних інформаційних системах.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Carvalho A. M., Oliveira R., Monteiro J., Silva C. Data Quality: revisiting dimensions towards new framework development. *Procedia Computer Science*. 2025. Vol. 253. P. 247–256. URL: <https://doi.org/10.1016/j.procs.2025.01.088>.

2. González-Vidal A., Ramallo-González A. P., Skarmeta A. F. Intrinsic and extrinsic quality of data for open data repositories. *ICT Express*. 2022. URL: <https://doi.org/10.1016/j.icte.2022.06.001>.
3. Rahal M., Abdallah M., Farhat F., El Hajj A. Enhancing Machine Learning Performance through Intelligent Data Quality Assessment: An Unsupervised Data-centric Framework. *Heliyon*. 2025. URL: <https://doi.org/10.1016/j.heliyon.2025.e42777>.
4. CKAN code architecture – CKAN 2.11 documentation. URL: <https://docs.ckan.org/en/2.11/contributing/architecture.html>.
5. Karamanou A., Tarabanis K., Peristeras V., Tambouris E. Exploring the Quality of Dynamic Open Government Data Using Statistical and Machine Learning Methods. *Sensors*. 2022. Vol. 22, No. 24. Article 9684. URL: <https://doi.org/10.3390/s22249684>.