



Р. Б. Юринець, І. Б. Пірко

Національний лісотехнічний університет України, м. Львів, Україна

## ІННОВАЦІЙНІ МЕТОДИКИ ІНТЕГРУВАННЯ ДАНИХ ДЛЯ ОПТИМІЗАЦІЇ ПРОЦЕСУ НАПОВНЕННЯ СХОВИЩА ДАНИХ

Проаналізовано та оцінено альтернативні методики інтегрування даних, що не потребують класичного ETL-підходу (англ. *Extract, Transform, Load*). Основну увагу приділено вивченню обмежень традиційного ETL, зокрема аналізу проблем, з якими стикаються організації через їх складність та витрати, пов'язані з традиційними ETL-процесами. Проаналізовано платформи, які дають змогу інтегрувати та наповнювати сховища даних без потреби писати складний код. Окремо розглянуто, як ці системи дають можливість швидше та ефективніше реагувати на зміни в даних і бізнес-вимогах. Зосереджено увагу на стратегіях і технологіях для інкрементного завантаження даних, які можуть значно зменшити навантаження та витрати на ресурси порівняно з повним перезавантаженням даних. Також проаналізовано, як ці методи допомагають утримувати дані актуальними без постійного повного оновлення сховища. Це дослідження сприяє детальному розумінню сучасних інструментів і методів, які роблять процес оптимізації наповнення сховища даних більш ефективним та доступним для широкого кола організацій, зокрема і тих, що не мають значних технічних ресурсів. Також підкреслено значення наявності кваліфікованих розробників і їхній вплив на загальні витрати проєкту, що не може бути недооцінений. Роль якісної комунікації між командами розробників, дата аналітиків і бізнес-користувачів є критичною для успішної реалізації проєктів з інтегрування даних. Ефективний обмін інформацією та чітке визначення бізнес-вимог відіграють ключову роль у синхронізації та оптимізації процесів завантаження та оброблення даних, що в остаточному підсумку веде до кращого рівня обслуговування та зменшення часу реакції на зміни у бізнес-середовищі. Результати дослідження демонструють, що впровадження технологій Incremental Data Loading та CDC (англ. *Change Data Capture*) у великій роздрібній компанії значно покращило тривалість завантаження даних та оперативність бізнес-процесів. Порівнюючи результати впровадження з інших галузей, з'ясовано, що методики Incremental Data Loading мають універсальний характер, але потребують індивідуального підходу залежно від сектора застосування. Обговорення рекомендацій для компаній, що розглядають адаптацію цього підходу, має значний потенціал для покращення ефективності роботи, зниження витрат та підвищення якості обслуговування клієнтів, особливо у роздрібному секторі.

**Ключові слова:** сховища баз даних; платформи інтегрування даних без написання коду; інкрементне завантаження даних; захоплення змін даних.

### Вступ/ Introduction

Наразі оброблення та зберігання даних, традиційні методи їхнього інтегрування, які містять складні та високозатратні ETL-процеси, часто виявляються недостатньо гнучкими та швидкими для задоволення динамічних потреб бізнесу. Постійне зростання об'ємів даних та підвищення вимог до швидкості їх оброблення змушують організації шукати альтернативні рішення для оптимізації процесів наповнення сховищ даних.

Перехід до No-Code-платформ для інтегрування даних та застосування стратегій інкрементного завантаження даних набуває щораз більшої популярності як ефективний спосіб вирішення викликів традиційних ETL-методик. Ці інноваційні підходи дають змогу знизити завантаження робочих систем, зменшити тривалість оброблення даних та забезпечити більшу адаптивність до змін у бізнес-процесах.

Ці підходи також вирішують ключову потребу бізнесу в забезпеченні доступу до даних у режимі онлайн (або максимально наближеному до онлайн) та подальшому швидкому реагуванні на зміни. Це особливо важливо для ефективного впровадження нових бізнес-стратегій, оптимізації логістики та втілення інших важливих змін у роботі компанії, що можуть істотно підвищити її конкурентоспроможність та оперативність на ринку виробників.

Цей аналіз має на меті дослідити потенціал та реалії використання No-Code-платформ і методик інкрементного завантаження в контексті автоматичного наповнення сховищ даних. Аналізують переваги, виклики, а також практичні особливості імплементації цих технологій, надаючи аналіз наявних рішень та майбутніх перспектив у цій галузі. Така інформація буде корисною для організацій, які прагнуть максимізувати ефективність своїх дата-систем, покращуючи при цьому адап-

### Інформація про авторів:

Юринець Роман Богданович, аспірант, кафедра комп'ютерних наук.

Email: yurynets.roman@nltu.lviv.ua; <https://orcid.org/0009-0006-3589-5803>

Пірко Ігор Богданович, канд. фіз.-мат. наук, доцент, кафедра комп'ютерних наук.

Email: pirko@nltu.edu.ua; <https://orcid.org/0009-0008-2378-2929>

**Цитування за ДСТУ:** Юринець Р. Б., Пірко І. Б. Інноваційні методики інтегрування даних для оптимізації процесу наповнення сховища даних. Науковий вісник НЛТУ України. 2024, т. 34, № 6. С. 101–105.

**Citation APA:** Yurynets, R. B., & Pirko, I. B. (2024). Innovative data integration methodologies for optimizing data warehouse filling. *Scientific Bulletin of UNFU*, 34(6), 101–105. <https://doi.org/10.36930/40340614>

тивність та реакцію на динамічне бізнес-середовище.

*Об'єкт дослідження* – інтеграція даних як можливість оптимізації процедури наповнення сховища даних.

*Предмет дослідження* – методи аналізу та порівняльного оцінювання ефективності традиційних ETL-методів порівняно з новітніми No-Code-платформами та методиками інкрементного завантаження даних щодо їх впливу на скорочення тривалості виконання операцій та витрат за одночасного збільшення гнучкості у процесах інтегрування даних.

*Мета роботи* – оцінити ефективність й адаптивність інноваційних методів інтегрування даних на базі No-Code-платформ та особливості використання інкрементального завантаження даних для оптимізації наповнення сховищ даних, що забезпечить їхню адаптивність до змін у даних та бізнес-вимогах.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

1. Проаналізувати наявні ETL-процеси і їх обмеження, оцінити можливості No-Code-платформ і їх впливу на процеси інтегрування даних, що дасть змогу ідентифікувати ключові особливості для підвищення адаптивності та ефективності систем управління даними в динамічних бізнес-умовах.
2. Вивчити та проаналізувати стратегії інкрементного завантаження як засобу зниження навантаження на інформаційні системи, що дасть змогу оптимізувати процеси оброблення даних та забезпечити швидке реагування на зміни у них, що критично для підтримки прийняття оперативних рішень у діловій діяльності.
3. Порівняти ефективність використання традиційних та інноваційних методів інтегрування даних, виявити оптимальні практики для їх застосування у конкретних бізнес-умовах і сформулювати рекомендації для успішного інтегрування та використання нової інформаційної архітектури інтегрування даних в організаціях, що забезпечить підвищення гнучкості та оперативності бізнес-процесів, ефективне масштабування та адаптацію до змінних умов ринку.

**Аналіз останніх досліджень та публікацій.** Термін, що знаходиться за аббревіатурою ETL (англ. *Extract, Transform, Load*), є стандартним процесом у сфері оброблення даних [4, 16], який використовують для збирання даних з різних джерел, їх перетворення у формат, придатний для аналізу, і завантаження їх у сховище для подальших запитів і аналізу.

Як зазначено в багатьох дослідженнях, традиційні ETL-процеси часто зазнають складнощів, які істотно ускладнюють їх впровадження та ефективне використання [2, 5, 8, 15]. Однією з основних проблем є складність та висока вартість налаштування та підтримки. Спеціалізовані знання та значні ресурси, які вимагаються для підтримки ETL-процесів, роблять цей процес дорогим складником ІТ-структури компанії. Часто витрати на підтримку команди розробників системи й аналітиків її проектування, які володіють належними знаннями про неї, перевищують вартість самої системи, на якій реалізують рішення.

Дослідження показують, що традиційні ETL-системи характеризуються жорсткістю у реагуванні на зміни [8]. Вони мають складнощі зі швидкою адаптацією до змін у джерелах даних або бізнес-процесах [2]. Необхідні для змін ресурси часто бувають недоступними, а процес впровадження змін може бути довгим і складним. Це не тільки становить проблему для впровадження змін, але й часто призводить до затримок у реалізації

проектів і збільшення загальних витрат, що негативно впливає на ефективність бізнес-операцій і може створити додаткові обтяження для команди, залученої у процес. А це, водночас, може мати серйозні наслідки для оперативності прийняття рішень і загалом знижувати ефективність бізнесу.

## Результати дослідження та їх обговорення / Research results and their discussion

*Аналіз проблеми в контексті сучасних вимог.* У динамічному цифровому бізнес-середовищі потреба у швидкому і гнучкому доступі до даних постійно зростає. Традиційні ETL-системи часто не можуть забезпечити потрібну швидкість і масштабування. Наприклад, у сферах, де дані швидко змінюються, як фінансові служби або онлайн-ритейл, затримка інформації може призводити до втрати прибутків і можливостей.

Тому в процесі впровадження таких систем виникають питання розгляду альтернативних підходів, таких як No-Code-платформи і стратегії інкрементного завантаження, можуть пропонувати рішення для тих бізнесів, що прагнуть до більшої гнучкості та оперативності в інтегруванні та аналізі даних.

**Перехід до використання No-Code Data Integration Platforms.** No-Code Data Integration Platforms – це інноваційні рішення, які дають можливість користувачам без глибоких технічних знань створювати та керувати процесами інтегрування даних [14]. Такі платформи пропонують графічні інтерфейси з перетягуванням компонентів (drag-and-drop), які забезпечують візуалізацію та налаштування потоків даних. Вони є відповіддю на потребу в більшій гнучкості та оперативності, даючи можливість швидко реагувати на зміни у джерелах даних і бізнес-процесах.

**Аналіз наявних платформ і методик інтегрування даних.** Серед популярних No-Code Data Integration Platforms можна виділити такі рішення, як Zapier, Microsoft Power Automate та Workato [11, 14]. Ці платформи дають можливість користувачам інтегрувати, фільтрувати, перетворювати та автоматизувати потоки даних між різними джерелами та призначеннями без написання коду. Платформи також містять підтримку багатьох популярних вебсервісів і API, що забезпечує широкі можливості для інтегрування.

Переваги та потенційні недоліки No-Code-інтегрування охоплюють кілька ключових аспектів. Переваги містять зниження бар'єру входження, оскільки користувачам не потрібні спеціальні технічні знання для роботи з даними. Дата аналітики та архітектори рішень можуть швидко зрозуміти і визначити, які дані найцінніші для бізнесу, і вибірково інтегрувати ці дані для оптимізації процесів. Швидкість реалізації також є значною перевагою, адже No-Code-рішення дають можливість швидко налаштовувати і запускати інтеграційні процеси. Еластичність цих систем дає змогу їм швидко адаптуватися до змін у даних або бізнес-процесах без необхідності перепрограмування. Що стосується недоліків, то основними є обмеження функціональності.

Попри гнучкість, деякі спеціалізовані або складні завдання можуть бути недосяжними для платформ без коду. Також існує недолік у вигляді залежності від постачальника, оскільки використання No-Code-платформ може створити певну залежність від конкретного виробника програмного забезпечення, обмежуючи можли-

вості міграції до інших систем або сторінних рішень. Це може змушувати компанію підтримувати команду розробників, які мають знання про специфіку використаного програмного забезпечення, для забезпечення стабільності і безперебійності роботи системи.

Організації використовують No-Code-платформи для автоматизації рутинних завдань, інтегрування нових сервісів і джерел даних, а також для ширшої демократизації доступу до даних серед нетехнічного персоналу. Вони значно спрощують такі процеси, як звітність, моніторинг, аналітика, роблячи дані більш доступними для широкого кола співробітників і сприяючи швидшому прийняттю інформованих рішень.

Перехід до No-Code Data Integration Platforms у такий спосіб стає могутньою стратегією для тих організацій, які прагнуть оптимізувати витрати на управління даними, підвищити продуктивність та поліпшити якість обслуговування клієнтів завдяки більш оперативному та гнучкому доступу до інформації.

No-Code Data Integration Platforms можуть слугувати як довгострокове рішення для наповнення сховищ даних, але їх ефективність та придатність залежать від певних факторів і потреб специфічного бізнесу. Оцінювання їх як довгострокового рішення вимагає розгляду деяких аспектів (швидкість адаптації та впровадження, зниження витрат на власність та експлуатацію, демократизація доступу до даних, також треба брати до уваги описані вище обмеження).

Incremental Data Loading (інкрементне завантаження даних) – це процес, що дає змогу завантажувати тільки ті дані, які зазнали змін або були додані з моменту останньої синхронізації [2, 13]. Цей підхід сприяє оптимізації процесів оброблення даних, зниженню навантаження на мережеві ресурси та зберігаючі системи, а також зменшенню часу, необхідного для завантаження даних.

1. Timestamp-based: дані фільтрують та завантажують на підставі міток часу створення або останньої модифікації. Ця техніка ефективна, коли джерело даних підтримує точний часовий штамп оновлень.
2. Latest Watermark: використовують змінний ідентифікатор, який відображає останню відслідковану зміну. Наприклад, можна використовувати найбільше значення ID або ключове значення, знайдене у попередній загрузці, завантажувати тільки записи, які перевищують це значення.
3. Change Data Capture (CDC): техніка, яка містить використання логів бази даних для відстеження змін і відбір тільки змінених даних для завантаження [2, 12]. CDC може бути налаштований для роботи в реальному часі, що виділяє його серед інших технік incremental loading.
4. Snapshot Diff: порівняння порції поточних даних із попередньо збереженою копією (snapshot) для ідентифікації і завантаження тільки змінених або нових записів.

Виклики та обмеження інкрементного завантаження даних містять низку аспектів [7, 12, 13], які можуть ставити під загрозу ефективність цього методу. Одним з основних викликів є додаткове навантаження на джерело даних, зокрема техніки, як Change Data Capture (CDC), вимагають читання транзакційних логів, що може збільшити навантаження на сорсову систему. Також важливою проблемою є синхронізація даних, оскільки підтримання цілісності даних між джерелом та призначенням може бути складним, особливо у разі збоїв або переривань процесу завантаження.

Окрім цього, комплексність реалізації також є значним викликом: налаштування та підтримка інкремент-

них завантажень потребують додаткових зусиль і технічної експертизи, особливо при використанні складних методик, таких як CDC. Необхідність кваліфікованого персоналу є також критичною, оскільки ці підходи вимагають постійного оновлення та завантаження тільки змінених або нових даних, що є складним і потребує глибоких знань у сферах баз даних, SQL та оброблення поточних даних. Фахівці повинні вміти налаштовувати детекцію змін у датасетах, забезпечувати надійність та цілісність даних під час їх оброблення.

Спеціалісти відіграють ключову роль і повинні мати не тільки технічні навички, але й вміти ефективно взаємодіяти з власниками даних та іншими зацікавленими сторонами, щоб обговорювати параметри інтегрування даних, інформувати про потенційні проблеми та співпрацювати над знаходженням рішень. Такі комунікаційні навички є невід'ємною складовою частиною успішного управління даними, особливо коли інкрементне завантаження може стикатися з такими проблемами, як порушення цілісності даних у разі помилок завантаження, проблем зі зберіганням і великих об'ємів даних.

Отже, наявність кваліфікованого персоналу й ефективна комунікація з власниками даних є критично важливими для успішного застосування технік Incremental Data Loading. Це забезпечує не тільки технічну коректність процесів, але й їхню відповідність бізнес-логіці та цілям організації.

Незважаючи на можливі труднощі, інкрементне завантаження даних пропонує значні переваги для бізнесу, які містять ефективність ресурсів, скорочення тривалості оброблення та підвищення актуальності даних. Завантаження тільки нових або змінених даних знижує вартість та покращує продуктивність сховищ, що дає можливість економити на ресурсах й одночасно підвищувати ефективність системи. Цей підхід також значно скорочує тривалість оброблення, оскільки менший об'єм даних потребує менше часу для оброблення, забезпечуючи більш своєчасний доступ до актуалізованої інформації. Окрім цього, завдяки постійному оновленню даних у сховищі, інкрементний підхід може поліпшити їх актуальність, що є критично важливим у динамічному бізнес-середовищі.

Застосування технік Incremental Data Loading у контексті сучасних бізнес-вимог і технологій дата-інжинірингу дає змогу компаніям підтримувати високу продуктивність і гнучкість в управлінні даними, що є вирішальним для досягнення успіху на сучасному ринку.

У таблиці наведено порівняння різних підходів для інтегрування даних. Ця таблиця допомагає наочно виділити ключові особливості кожного методу і підходу, вказуючи на їхні сильні та слабкі сторони з погляду технічних вимог, складності впровадження, швидкості реагування, а також вартості та підтримки великих обсягів даних.

Проведемо аналіз кейсів IT-департаменту великої роздрібною компанією, яка стикалася з проблемами затримки інформації у своїй діяльності через використання традиційних методів завантаження даних. Компанія вирішила імплементувати техніку Incremental Data Loading з використанням технології Change Data Capture.

Підготовка плану імплементації:

- вибір технології: вибрано рішення, яке підтримує CDC та інтегрують з наявними системами компанії;
- залучення Stakeholders: у кейсі активно брали участь дата-овнери та відповідальні за бази даних.

- Збір детальної інформації:
- технічний аналіз: систематичні звіти про IT-інфраструктуру і поточні процеси збирання даних;
  - запити/відгуки користувачів: збір інформації про задоволення користувачів наявною системою оброблення даних.
- Аналіз та інтерпретація результатів:

- зменшення затримки даних: завдяки CDC, потрібні дані стали доступними майже в реальному часі;
- підвищення продуктивності: спостерігалось зниження тривалості завантаження даних на 60-80 %;
- відгуки користувачів: покращення загальної задоволеності користувачів щодо допомоги у прийнятті бізнес-рішень.

**Таблиця.** Порівняння різних підходів для інтегрування даних / Comparison of different approaches for data integration

| Параметр/Метод                                  | Traditional ETL | No-Code Platforms       | Incremental Data Loading  | Change Data Capture (CDC)        |
|-------------------------------------------------|-----------------|-------------------------|---------------------------|----------------------------------|
| Основні технології                              | Скрипти, SQL    | Візуальне програмування | Диференційне завантаження | Захоплення змін у реальному часі |
| Технічні вимоги                                 | Високі          | Низькі                  | Середні                   | Високі                           |
| Складність впровадження                         | Висока          | Низька                  | Середня                   | Висока                           |
| Час реагування на зміни даних                   | Повільний       | Швидкий                 | Швидкий                   | Найшвидший                       |
| Підтримка великих обсягів даних                 | Так             | Обмежено                | Так                       | Так                              |
| Вартісна ефективність                           | Змінна          | Висока                  | Середня                   | Змінна                           |
| Підтримка реального часу                        | Ні              | Обмежено                | Ні                        | Так                              |
| Легкість оновлення та технічного обслуговування | Складна         | Проста                  | Середня                   | Складна                          |

**Обговорення результатів дослідження.** Спробуємо обговорити результати дослідження та висвітлити ключові особливості імплементації Incremental Data Loading та технологій CDC у великій роздрібній компанії та їхнє порівняння з традиційними ETL-процесами. Використання цих технологій привело до скорочення тривалості завантаження даних та підвищення оперативності бізнес-процесів, що є особливо цінним в динамічних ринкових умовах, що підтверджують дослідження про вплив Incremental RFM моделі на оптимізацію часу (60-80 %) та підвищення точності прогнозування покупок комодитів у роздрібній торгівлі [18].

Додатково, як доводить дослідження алгоритмів розроблення та валідації ETL-процесу для німецьких реальних даних у медичній сфері, Incremental loading виявився ефективнішим за щоденне масивне завантаження (87 %), що значно скорочує час виконання та забезпечує актуальність даних для підтримки клінічних досліджень [6]. Стратегія інкрементного переходу з використанням технологій CDC, обговорена у дослідженні про міграцію до подієво-орієнтованих мікросервісів, вказала на значення реплікації даних та зв'язку між мікросервісами, покращуючи продуктивність систем [3].

Обмеження, пов'язані із залежністю від якості даних у джерельних системах також важливі для обговорення щодо пришвидшення процесу оброблення та точності великих даних, вказують на можливості покращення цих аспектів через оптимізацію та застосування індексційних технік [17]. На підставі аналізу можна розробити рекомендації для компаній щодо адаптації Incremental Data Loading та використання No-Code-платформ для залучення нетехнічного персоналу до процесів інтегрування та аналізу даних. Це поліпшує гнучкість та адаптивність компаній до змін, як це відображено в дослідженні використання великих даних у роздрібній індустрії для трансформації організаційної практики [9].

Важливість залучення стейкхолдерів і дата аналітиків у процеси інтегрування даних й оптимізації бізнес-процесів є критичною. Наприклад, дослідження управління складськими системами показало, що аналітика даних відіграє ключову роль в оптимізації логістичних каналів [1]. Також дослідження візуалізаційного підходу для моніторингу оброблення замовлень в електронній комерції вказує на, що візуальна аналітика може допомогти менеджерам складів аналізувати та покращувати ефективність оброблення замовлень у реальному ча-

сі, що необхідно для успішної реалізації та експлуатації систем управління даними [10]. Отже, взаємодія зі стейкхолдерами та аналітиками на різних рівнях управління є істотною для успішного впровадження аналітичних інновацій в компанії.

Отже, за результатами виконаної роботи можна сформулювати такі наукову новизну та практичну значущість результатів дослідження.

*Наукова новизна отриманих результатів дослідження* – удосконалено методику аналізу впливу No-Code-платформ порівняно з традиційними ETL-процесами, що дає змогу виділити оптимальні сценарії їх застосування в різноманітних бізнес-середовищах, містить інтеграцію стратегій інкрементного завантаження даних для зниження навантаження на системи та скорочення тривалості відгуку і в такий спосіб забезпечить гнучкість і оперативність у динамічних ринкових умовах.

*Практична значущість результатів дослідження* – отримані результати дослідження можна застосовувати для оптимізації методів інтегрування даних через використання No-Code-платформ й інкрементного завантаження, що приведе до істотного скорочення тривалості та витрат на оброблення та управління даними в компаніях, забезпечуючи ефективніше розподілення ресурсів і швидшу реакцію на зміни в даних і бізнес-умовах. Це також поліпшує якість прийнятих рішень та знижує ризики, пов'язані з використанням застарілих даних завдяки пришвидшеному оновленню та кращій інтегрування даних, при цьому зниження технічного бар'єру для взаємодії з No-Code-платформами сприяє полегшенню доступу більшому колу співробітників до аналізу даних і забезпечує більшу гнучкість організації, що, як наслідок, підвищує конкурентоспроможність компаній і знижує витрати на IT-інфраструктуру.

## Висновки / Conclusions

Оцінено ефективність й адаптивність інноваційних методів інтегрування даних на базі No-Code-платформ та особливості використання інкрементального завантаження даних для наповнення сховищ даних, що забезпечують їхню адаптивність до змін у даних та бізнес-вимогах. За результатами проведеного дослідження можна зробити такі основні висновки.

1. З'ясовано, що результати впровадження інноваційних методичних підходів, зокрема Incremental Data Loading та технології Change Data Capture (CDC), виявились ефективними в контексті зниження часових затримок,

оптимізації IT-інфраструктури і дата-процесів, а також поліпшення якості прийняття рішень у великій роздрібній компанії. Це вказує на значення наявності кваліфікованих розробників, які можуть ефективно імплементувати та підтримувати ці технології.

2. Наведено перспективи використання зазначених вище технологій для оптимізації бізнес-операцій і стратегічного планування у різних галузях, що відкриває нові можливості для технологічного розвитку індустрій.
3. Надано рекомендації щодо розвитку внутрішніх кваліфікацій інженерів даних і аналітиків для ефективної адаптації до нових технологій інтегрування даних, що забезпечить подальше поліпшення та розширення технологій Incremental Data Loading та CDC і їх вплив на технологічний розвиток.
4. Виявлено, що No-Code платформи пропонують значні переваги у швидкості і зручності розробки рішень для інтегрування даних, знижуючи поріг технічних знань, необхідних для реалізації проєктів, що важливо на початкових етапах проєкту.
5. Підтверджено важливість ролі висококваліфікованих розробників у забезпеченні успішного впровадження інноваційних методів інтегрування даних, які підвищують швидкість і точність процесів наповнення сховищ даних та потужно впливають на бізнес-стратегії.

## References

1. Andiyappillai, N. (2019). Data Analytics in Warehouse Management Systems (WMS) Implementations – A Case Study. *International Journal of Computer Applications*. <https://doi.org/10.5120/IJCA2019918542>
2. Awiti, J., Vaisman, A., & Zimányi, E. (2020). Design and implementation of ETL processes using BPMN and relational algebra. *Data & Knowledge Engineering*, 129. <https://doi.org/10.1016/j.datak.2020.101837>
3. Bashtovyi, A., & Fechan, A. (2023). Change Data capture for migration to event-driven microservices Case Study. *2023 IEEE 18th International Conference on Computer Science and Information Technologies (CSIT)*, 1–4. <https://doi.org/10.1109/CSIT61576.2023.10324262>
4. El-Sappagh, S., Hendawi, A., & El-Bastawissy, A. (2011). A proposed model for data warehouse ETL processes. *Journal of King Saud University – Computer and Information Sciences*, 23, 91–104. <https://doi.org/10.1016/J.JKSUCI.2011.05.005>
5. Gueddoudj, E., & Chikh, A. (2023). Towards a Scalable and Efficient ETL. *International Journal of Computing and Digital Systems*. <https://doi.org/10.12785/ijcds/140195>
6. Henke, E., Peng, Y., Reinecke, I., Zoch, M., Sedlmayr, M., & Bathelt, F. (2023). An Extract-Transform-Load Process Design for the Incremental Loading of German Real-World Data Based on FHIR and OMOP CDM: Algorithm Development and Validation. *JMIR Medical Informatics*, 11. <https://doi.org/10.2196/47310>
7. Kleppmann, Martin. (2017). Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems, 614. URL: <https://www.oreilly.com/library/view/designing-data-intensive-applications/9781491903063/>
8. Oliveira, B., Oliveira, O., & Belo, O. (2023). Overcoming traditional ETL-systems architectural problems using a service-oriented approach. *Expert Systems*, 40. <https://doi.org/10.1111/exsy.13442>
9. Santoro, G., Fiano, F., Bertoldi, B., & Ciampi, F. (2019). Big data for business management in the retail industry. *Management Decision*. <https://doi.org/10.1108/MD-07-2018-0829>
10. Tang, J., Zhou, Y., Tang, T., Weng, D., Xie, B., Yu, L., Zhang, H., & Wu, Y. (2021). A Visualization Approach for Monitoring Order Processing in E-Commerce Warehouse. *IEEE Transactions on Visualization and Computer Graphics*, 1–10. <https://doi.org/10.1109/TVCG.2021.3114878>
11. Website. (2024). Automation software, the 6 best: Zapier, Workato, etc. URL: <https://lapala.io/en/automation-software/>
12. Website. (2024). Change Data Capture (CDC). URL: <https://www.geeksforgeeks.org/change-data-capture-cdc/>
13. Website. (2024). Incremental Data Loading from Databases for ETL. URL: <https://www.linkedin.com/pulse/incremental-data-loading-from-databases-etl-dhiraj-patra-fxdwc>
14. Website. (2024). The Best No-Code Integration Tools For Businesses. URL: <https://exalate.com/blog/no-code-integration/>
15. Website. (2024). What are the common challenges faced by ETL developers? URL: <https://moldstud.com/articles/p-what-are-the-common-challenges-faced-by-etl-developers>
16. Website. (2024). What is ETL (Extract Transform Load). URL: <https://aws.amazon.com/what-is-etl/>
17. Wei, F., Jiang, S., & Wang, K. (2023). A survey on fast and incremental technologies for big data, 12605, 126050U–126050U-6. <https://doi.org/10.1117/12.2673376>
18. Zuo, H., Yang, S., Wu, H., Guo, W., Wang, L., Chen, X., & Su, Y. (2022). A Data-Driven Customer Profiling Method for Offline Retailers. *Computational Intelligence and Neuroscience*, 1. <https://doi.org/10.1155/2022/8069007>

**R. B. Yurynets, I. B. Pirko**

*Ukrainian National Forestry University, Lviv, Ukraine*

## INNOVATIVE DATA INTEGRATION METHODOLOGIES FOR OPTIMIZING DATA WAREHOUSE FILLING

Research enables the analysis and evaluation of alternative data integration methodologies that do not require the classical ETL approach. Primary focus is given to examining the limitations of traditional ETL, including the analysis of the problems organizations face due to the complexity and costs associated with traditional ETL processes. A detailed review of platforms that allow data integration and enrichment of data stores without the need for complex coding has been conducted. Additionally, the ways these systems enable faster and more efficient responses to changes in data and business requirements was examined. Special attention is dedicated to strategies and technologies for incremental data loading, which can significantly reduce the workload and resource expenses compared to full data reloading. The research also analyzed how these methods help keep data up-to-date without constant full refreshes of the storage. This research contributes to an in-depth understanding of modern tools and methods that make the process of optimizing data store filling more efficient and accessible for a wide range of organizations, including those with limited technical resources. The importance of having qualified developers and their impact on the overall project costs, which cannot be underestimated, is also highlighted. The role of quality communication between development teams, data analysts, and business users is critical for the successful implementation of data integration projects. Effective information exchange and clear definition of business requirements play a key role in the synchronization and optimization of data loading and processing processes, ultimately leading to improved service levels and reduced response time to changes in the business environment. The research results demonstrate that the implementation of Incremental Data Loading and CDC technologies in a large retail company significantly improved data loading times and operational efficiency of business processes. Comparing the implementation results from other industries, it is clear that Incremental Data Loading methodologies have a universal character but require a tailored approach depending on the application sector. Discussing recommendations for companies considering adapting this approach holds significant potential for improving work efficiency, reducing costs, and enhancing customer service quality, especially in the retail sector.

**Keywords:** data warehouses; no-code data integration platforms; incremental data loading; CDC (Change Data Capture).