

Облікова картка дисертації

I. Загальні відомості

Державний обліковий номер: 0826U001123

Особливі позначки: відкрита

Дата реєстрації: 20-04-2026

Статус: Запланована

Реквізити наказу МОН / наказу закладу:



II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Малий Роман Михайлович

2. Roman M. Malyi

Кваліфікація:

Ідентифікатор ORCID ID: 0000-0002-2255-1132

Вид дисертації: доктор філософії

Аспірантура/Докторантура: так

Шифр наукової спеціальності: 121

Назва наукової спеціальності: Інженерія програмного забезпечення

Галузь / галузі знань:

Освітньо-наукова програма зі спеціальності: Інженерія програмного забезпечення

Дата захисту:

Спеціальність за освітою: Інженерія програмного забезпечення

Місце роботи здобувача: Національний університет "Львівська політехніка"

Код за ЄДРПОУ: 02071010

Місцезнаходження: вул. Степана Бандери, Львів, 79013, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

III. Відомості про організацію, де відбувся захист

Шифр спеціалізованої вченої ради (разової спеціалізованої вченої ради): PhD 12900

Повне найменування юридичної особи: Національний університет "Львівська політехніка"

Код за ЄДРПОУ: 02071010

Місцезнаходження: вул. Степана Бандери, Львів, 79013, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

IV. Відомості про підприємство, установу, організацію, в якій було виконано дисертацію

Повне найменування юридичної особи: Національний університет "Львівська політехніка"

Код за ЄДРПОУ: 02071010

Місцезнаходження: вул. Степана Бандери, Львів, 79013, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

V. Відомості про дисертацію

Мова дисертації: Українська

Коди тематичних рубрик: 20.54.07

Тема дисертації:

1. Методи еволюції схеми даних для супроводу документоорієнтованих баз даних
2. Data schema evolution methods for the maintenance of document-oriented databases

Реферат:

1. Сучасна інженерія програмного забезпечення, що базується на гнучких методологіях та принципах неперервної інтеграції, стає причиною постійної еволюції вимог, що неминуче призводить до частих змін у структурі даних. Некерований процес таких змін стає джерелом накопичення технічного боргу, збільшує витрати на супровід та створює значні ризики пошкодження даних, що ставить під загрозу надійність та довготривалу життєздатність програмного забезпечення. У зв'язку з цим, у дисертаційній роботі розв'язано актуальну науково-прикладну задачу з інженерії програмного забезпечення – удосконалення процесу супроводу еволюції схеми даних, зосереджуючись на двох ключових, але принципово різних архітектурних парадигмах: широко розповсюджених документоорієнтованих NoSQL базах даних та системах, побудованих на основі незмінних даних. Розроблено метод для кількісного вибору стратегії міграції у системах зі змінними даними, що враховує ключові характеристики проекту; метод порівняльного тестування продуктивності, спеціально розроблений для оцінювання баз даних у контексті ПДП (патерн джерела подій), що заповнює прогалину у наявних еталонних тестах; а також удосконалений метод застосування великих

мовних моделей, що дає можливість для автоматизації прийняття архітектурних рішень у найскладніших сценаріях еволюції незмінних даних. Дисертаційна робота складається зі вступу, чотирьох розділів, висновків, списку літературних джерел та додатків. У першому розділі проведено комплексний аналіз сучасного стану проблеми та обґрунтовано актуальність дослідження. Підтверджено, що відсутність формалізованих підходів до керування еволюцією схеми документоорієнтованих баз даних є однією з ключових перешкод для довготривалого супроводу програмного забезпечення. Встановлено, що гнучкість безсхемних баз даних, де схема є неявною та визначається кодом застосунку, створює значні виклики для узгодженості даних протягом життєвого циклу системи. Особливу увагу приділено етапу супроводу програмного забезпечення, оскільки саме він є основним рушієм постійних модифікацій структури даних у відповідь на нові вимоги чи зміну середовища. У другому розділі розроблено новий метод для вибору стратегії еволюції схеми даних у документоорієнтованих NoSQL базах. На відміну від наявних підходів, які часто ґрунтуються на досвіді розробників або орієнтовані на складні системи самоадаптації в реальному часі, запропонований метод є прагматичним інструментом підтримки прийняття рішень. Результати тестування продемонстрували, що застосування розробленого методу дає змогу скоротити обсяг операцій з міграції даних на 40% порівняно з традиційним підходом невідкладної міграції, що підтверджує його практичну ефективність та доцільність. У третьому розділі розроблено метод для порівняльного аналізу продуктивності баз даних у контексті застосування ПДП. Обґрунтовано, що наявні еталонні тести, як-от YCSB чи TPC-C, не враховують специфіку ПДП, зокрема домінуючі операції дозапису та критично важливий показник затримки створення проєкцій. Запропонований метод складається зі спеціалізованих тестових сценаріїв, що оцінюють максимальну пропускну здатність операцій запису під паралельним навантаженням, вимірюють затримку оновлення проєкцій та аналізують деградацію продуктивності зі збільшенням обсягу даних. У четвертому розділі розроблено та експериментально перевірено удосконалений метод використання великих мовних моделей (ВММ) для класифікації міграційної стратегії еволюції схеми для незмінних даних. Метою було перетворити процес вибору архітектурної стратегії з інтуїтивного, схильного до помилок судження, на детерміновану, автоматизовану процедуру. Для усунення недоліків евристичних підходів запропоновано формальну методологію, що ґрунтується на атомарній таксономії з 15 гранулярних операцій. Експериментальне порівняння на шести ВММ показало, що розроблений атомарний метод продемонстрував виняткову ефективність для моделей GPT-5, Gemini-2.5-pro, досягнувши 100% точності у виборі міграційної стратегії. Встановлено прямий зв'язок між архітектурним масштабом ВММ та стратегією формування запиту, що надає практичні рекомендації для побудови надійних систем підтримки архітектурних рішень.

2. Modern software engineering, based on agile methodologies and the principles of continuous integration, provokes a constant evolution of requirements, which inevitably leads to frequent changes in data structures. An unmanaged process of such changes becomes a source of accumulating technical debt, increases maintenance costs, and creates significant risks of data corruption, jeopardizing the reliability and long-term viability of software. In this regard, the dissertation solves a pressing scientific and applied problem in software engineering – improving the maintenance process of data schema evolution, focusing on two key but fundamentally different architectural paradigms: the widely used document-oriented NoSQL databases and systems built on the basis of immutable data. A method has been developed for the quantitative selection of a migration strategy in systems with mutable data, which considers key project characteristics; a performance comparison testing method specifically designed for evaluating databases in the context of the Event Sourcing (ES) pattern, filling a gap in existing benchmarks; and an improved method for applying large language models, providing the possibility to automate architectural decision-making in the most complex scenarios of immutable data evolution. The dissertation consists of an introduction, four chapters, conclusions, a list of references, and appendices. In the first chapter, a comprehensive analysis of the current state of the problem is conducted, and the relevance of the research is substantiated. It is confirmed that the lack of formalized approaches to managing schema evolution in document-oriented databases is one of the key obstacles to the long-term maintenance of software. It is established that the flexibility of schemaless databases, where the schema is implicit and defined by the application

code, creates significant challenges for data consistency throughout the system's lifecycle. Particular attention is paid to the software maintenance stage, as it is the primary driver of constant modifications to the data structure in response to new requirements or changes in the environment. In the second chapter, a new method for selecting a data schema evolution strategy in document-oriented NoSQL databases is developed. Unlike existing approaches, which are often based on developer experience or oriented toward complex real-time self-adaptation systems, the proposed method is a pragmatic decision-support tool. Testing results demonstrated that applying the developed method allows for reducing the volume of data migration operations by 40% compared to the traditional eager migration approach, confirming its practical effectiveness and expediency. In the third chapter, a method for the comparative analysis of database performance in the context of applying the ES pattern is developed. It is substantiated that existing benchmarks, such as YCSB or TPC-C, do not account for the specifics of ES, particularly the dominant append-only operations and the critically important metric of projection creation latency. The proposed method consists of specialized test scenarios that evaluate maximum write throughput under parallel load, measure projection update latency, and analyze performance degradation as data volume increases. In the fourth chapter, an improved method for using large language models (LLMs) to classify the migration strategy of schema evolution for immutable data is developed and experimentally validated. The goal was to transform the process of selecting an architectural strategy from an intuitive, error-prone judgment into a deterministic, automated procedure. To eliminate the shortcomings of heuristic approaches, a formal methodology is proposed, based on an atomic taxonomy of 15 granular operations. Experimental comparison on six LLMs showed that the developed atomic method demonstrated exceptional effectiveness for the GPT-5 and Gemini-2.5-pro models, achieving 100% accuracy in selecting the migration strategy. A direct link was established between the architectural scale of an LLM and the prompting strategy, providing practical recommendations for building reliable architectural decision support systems.

Державний реєстраційний номер ДіР:

Пріоритетний напрям розвитку науки і техніки: Інформаційні та комунікаційні технології

Стратегічний пріоритетний напрям інноваційної діяльності: Розвиток сучасних інформаційних, комунікаційних технологій, робототехніки

Підсумки дослідження: Нове вирішення актуального наукового завдання

Публікації:

- Fedushko S., Malyi R., Syerov Y., Serdyuk P. NoSQL document data migration strategy in the context of schema evolution. *Data and Knowledge Engineering*. 2024. Vol. 154.
- Malyi R., Serdyuk P. Developing a performance evaluation benchmark for event sourcing databases. *Вісник Національного університету "Львівська політехніка"*. Серія: Інформаційні системи та мережі. 2024. Вип. 15. С. 159–168.
- Serdyuk P., Malyi R. Development of a rule-based LLM prompting method for high-accuracy event-schema evolution. *Технологічний аудит та резерви виробництва*. 2025. Vol. 5, No. 2 (85) : Information and control systems. P. 13–19.

Наукова (науково-технічна) продукція: програмні продукти, програмно-технологічна документація

Соціально-економічна спрямованість:

Охоронні документи на ОПВ:

Впровадження результатів дисертації: Впроваджено

Зв'язок з науковими темами: 0125U002820

VI. Відомості про наукового керівника/керівників (консультанта)

Власне Прізвище Ім'я По-батькові:

1. Сердюк Павло Віталійович
2. Pavlo V. Serdiuk

Кваліфікація: к. т. н., доц., 05.13.12

Ідентифікатор ORCID ID: 0000-0002-2677-3170

Додаткова інформація:

Повне найменування юридичної особи: Національний університет "Львівська політехніка"

Код за ЄДРПОУ: 02071010

Місцезнаходження: вул. Степана Бандери, Львів, 79013, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

VII. Відомості про офіційних опонентів та рецензентів

Офіційні опоненти

Власне Прізвище Ім'я По-батькові:

1. Мороз Борис Іванович
2. Boris Moroz

Кваліфікація: д.т.н., професор, 05.13.06

Ідентифікатор ORCID ID: 0000-0002-5625-0864

Додаткова інформація:

Повне найменування юридичної особи: Національний технічний університет "Дніпровська політехніка"

Код за ЄДРПОУ: 02070743

Місцезнаходження: проспект Дмитра Яворницького, Дніпро, Дніпровський р-н., 49005, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Власне Прізвище Ім'я По-батькові:

1. Мазурова Оксана Олексіївна
2. Oksana O. Mazurova

Кваліфікація: к. т. н., доц., 01.05.04

Ідентифікатор ORCID ID: 0000-0003-3715-3476

Додаткова інформація:

Повне найменування юридичної особи: Харківський національний університет імені В. Н. Каразіна

Код за ЄДРПОУ: 02071205

Місцезнаходження: майдан Свободи, Харків, Харківський р-н., 61022, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Рецензенти

Власне Прізвище Ім'я По-батькові:

1. Левус Євгенія Василівна

2. Yevgeniya V. Levus

Кваліфікація: к. т. н., доц., 05.13.12

Ідентифікатор ORCID ID: 0000-0001-5109-7533

Додаткова інформація:

Повне найменування юридичної особи: Національний університет "Львівська політехніка"

Код за ЄДРПОУ: 02071010

Місцезнаходження: вул. Степана Бандери, Львів, 79013, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Власне Прізвище Ім'я По-батькові:

1. Яковина Віталій Степанович

2. Vitalii S. Yakovyna

Кваліфікація: д. т. н., професор, 01.05.03

Ідентифікатор ORCID ID: 0000-0003-0133-8591

Додаткова інформація:

Повне найменування юридичної особи: Національний університет "Львівська політехніка"

Код за ЄДРПОУ: 02071010

Місцезнаходження: вул. Степана Бандери, Львів, 79013, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

VIII. **Заключні відомості**

**Власне Прізвище Ім'я По-батькові
голови ради**

Федасюк Дмитро Васильович

**Власне Прізвище Ім'я По-батькові
головуючого на засіданні**

Федасюк Дмитро Васильович

**Відповідальний за підготовку
облікових документів**

Мельник Наталія Богданівна

Реєстратор

Юрченко Тетяна Анатоліївна

**Керівник відділу УкрІНТЕІ, що є
відповідальним за реєстрацію наукової
діяльності**



Юрченко Тетяна Анатоліївна