

**ХЕРСОНСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ**  
(повне найменування вищого навчального закладу)  
**ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА ДИЗАЙНУ**  
(повне найменування інституту, назва факультету (відділення))  
**КАФЕДРА ПРОГРАМНИХ ЗАСОБІВ І ТЕХНОЛОГІЙ**  
(повна назва кафедри (предметної, циклової комісії))

**Пояснювальна записка**

до кваліфікаційної роботи

магістра  
(освітній рівень)

на тему: «Розробка програмних засобів трансформації моделі баз  
даних»

Виконала: студентка групи 6ПР  
спеціальності  
121 - «Інженерія програмного забезпечення»  
(шифр і назва спеціальності)

Єсіп Анна Михайлівна  
(прізвище та ініціали)

Керівник д.т.н., професор Шерстюк В.Г.  
(прізвище та ініціали)

Рецензент \_\_\_\_\_  
(прізвище та ініціали)

Хмельницький - 2023

Херсонський національний технічний університет

(повне найменування вищого навчального закладу)

Факультет, відділення Інформаційних технологій та дизайну  
Кафедра Програмних засобів і технологій  
Освітній рівень магістр  
Спеціальність 121 – Інженерія програмного забезпечення  
(шифр і назва)

**ЗАТВЕРДЖУЮ**

в.о. завідувача кафедри

Програмних засобів і технологій

к.т.н. доц. О.Є. Огнєва

“ \_\_\_\_\_ ” \_\_\_\_\_ 2023 р.

**З А В Д А Н Н Я**

**НА КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТУ**

Єсіп Анні Михайлівні

(прізвище, ім'я, по батькові)

1. Тема роботи «Розробка програмних засобів трансформації моделі баз даних»

керівник роботи д.т.н., професор Шерстюк В.Г.

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджена наказом вищого навчального закладу від 29.09.2023 р. №509-С

2. Строк подання студентом роботи 20.12.2023

3. Вихідні дані до роботи літературні та періодичні джерела, матеріали переддипломної практики, теоретико-математичні дослідження, методи трансформації баз даних з ER-моделі в реляційну модель

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити):

1) огляд загальних принципів організації даних та комунікацій в мікросервісній архітектурі;

2) дослідження та оптимізація методів здійснення розподілених транзакцій в мікросервісній архітектурі;

3) обробка розподілених транзакцій в мікросервісній архітектурі.

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)

1) Схема мікросервісної архітектури;

2) Схема оркестрації транзакцій;

3) Використання кластерів Apache Kafka;

4) Схема дуплікації даних;

5) Схема алгоритму;

6) Скріншоти програми.

## 6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

7. Дата видачі завдання 29.09.2023**КАЛЕНДАРНИЙ ПЛАН**

№	Назва етапів виконання роботи	Термін виконання етапів роботи	Примітки
1.	Отримання завдання	29.09.2023	Виконано
2.	Підбір літератури	05.10.2023	Виконано
3.	Аналіз предметної області	19.10.2023	Виконано
4.	Розробка та обґрунтування завдання	26.10.2023	Виконано
5.	Розробка концептуальної моделі	05.11.2023	Виконано
6.	Розробка алгоритму	12.11.2023	Виконано
7.	Проектування програми	26.11.2023	Виконано
8.	Розробка інтерфейсу програми	30.11.2023	Виконано
9.	Тестування програми	05.12.2023	Виконано
10.	Оформлення пояснювальної записки	10.12.2023	Виконано
11.	Захист кваліфікаційної роботи	20.12.2023	Виконано

Студент

Єсіп А.М.  
(підпис) (прізвище та ініціали)

Керівник роботи

Шерстюк В.Г.  
(підпис) (прізвище та ініціали)

## РЕФЕРАТ

**Пояснювальна записка:** 104 сторінки, 36 рисунків, 13 таблиць, 1 додаток, 55 джерел.

**Об'єкт дослідження:** реляційні та нереляційні бази даних.

**Предмет дослідження:** методи та алгоритми обробки даних у реляційних та ER-моделях баз даних.

**Мета роботи** полягає у створенні програмного забезпечення для реалізації методів та алгоритмів трансформації баз даних з ER-моделі в реляційну модель.

**Методи дослідження** базуються на основних принципах теорії баз даних, а саме методах теоретичного моделювання, теорії графів, теорії множин, теоретичних основах проектування реляційних та нереляційних баз даних.

**Новизна** отриманих результатів полягає у збільшенні структури та функціональності ядра CASE-системи. Створено динамічну бібліотеку, яка містить класи для трансформації схем з ERM-моделі в реляційну, розроблений виконуваний прототип модуля в середовищі Data Atlas Modeler.

**Практична цінність** результатів роботи полягає у розробці прототипу модуля, який може вбудовуватись в середовище Data Atlas Modeler для отримання реляційних схем, на базі яких можна створювати SQL-скрипти для побудови різних баз даних.

**Перелік ключових слів:** модель даних, трансляція схем, ER-модель, реляційна модель, CASE-система, SQL-скрипт.

## АНОТАЦІЯ

Кваліфікаційна робота магістра складається зі вступу, трьох розділів, висновку, списку використаних джерел та додатку.

В роботі було досліджено реалізацію ефективного алгоритму трансформації баз даних з ER-моделі в реляційну модель.

В першому розділі роботи проаналізовано існуючі підходи до трансляції баз даних з одного формату в інший, а також підходи до оптимізації структури NOSQL БД та оптимізації запитів до них.

В другому розділі розроблено метод оптимізації структури бази даних DataAtlas Modeler для моделі "зовнішній ключ" та формалізований метод, заснований на теорії множин, що дозволяє автоматизувати процес побудови оптимальної структури бази даних за заданою сукупністю властивостей об'єктів та з урахуванням їх входження до запитів до бази даних.

В третьому розділі розроблено метод побудови схеми бази даних «сутність – зв'язок» ER-моделі та ER-діаграми з урахуванням інформаційного графа запиту та розподіленої структури реляційної бази даних, що дозволяє побудувати на шардах та репліках оптимальну за швидкістю виконання запитів структуру колекцій.

Розроблено програму для реалізації методів та алгоритмів трансформації баз даних з ER-моделі в реляційну модель, яка може застосовуватися для побудови моделі даних методом послідовних уточнень початкових діаграм. Програма, що реалізує методи та алгоритми трансформації баз даних з ER-моделі в реляційну модель, в значній мірі забезпечує економію часу та ресурсів.

## ABSTRACT

The master's thesis consists of an introduction, three sections, a conclusion, a list of literature and an appendix.

The thesis investigated the implementation of an effective database transformation algorithm from the ER model to the relational model.

The first section of the paper analyzes existing approaches to translation of databases from one format to another, as well as approaches to optimizing the NOSQL database structure and optimizing queries to them.

In the second section, a method for optimizing the structure of the DataAtlas Modeler database for the "foreign key" model and a formalized method based on set theory are developed, which allows automating the process of building an optimal database structure based on a given set of object properties and taking into account their inclusion in queries to databases.

In the third section, a method of building an "essence-relationship" database scheme of an ER-model and an ER-diagram, taking into account the information graph of the request and the distributed structure of the relational database, is developed, which allows you to build a structure of collections that is optimal for the speed of query execution on shards and replicas.

A program has been developed for the implementation of methods and algorithms for the transformation of databases from the ER model into a relational model, which can be used to build a data model by the method of successive refinements of initial diagrams. The program, which implements methods and algorithms for database transformation from the ER model to the relational model, significantly saves time and resources.

## ЗМІСТ

ЗМІСТ .....	7
ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ .....	9
ВСТУП.....	10
РОЗДІЛ 1. МЕТОДИ ТА АЛГОРИТМИ ТРАНСФОРМАЦІЇ БАЗ ДАНИХ З ER-МОДЕЛІ В РЕЛЯЦІЙНУ МОДЕЛЬ .....	12
1.1. Введення та класифікації баз даних .....	12
1.2. Модель даних ER, основні поняття .....	20
1.3. Об'єктно-орієнтовані бази даних .....	24
1.4. Висновки до першого розділу.....	28
РОЗДІЛ 2. DATA ATLAS MODELER ВИКОРИСТАННЯ ER-МОДЕЛІ ДЛЯ СТРУКТУРИ ДАНИХ.....	30
2.1. Проектування інформаційного забезпечення. Створення ER-моделі.....	30
2.2. ER–метод проектування реляційних баз даних .....	37
2.3. Перетворення ER-моделі на реляційну модель .....	42
2.4. Висновки до другого розділу .....	49
РОЗДІЛ 3. ПРОЕКТУВАННЯ ER-ДІАГРАМИ .....	50
3.1. Аналіз методів логічного проектування під СКБД MongoDB .....	50
3.2. Аналіз методів логічного проектування під СКБД Neo4j .....	54
3.3. Планування експериментального дослідження .....	55
3.3.1. Аналіз та моделювання предметної області.....	55
3.3.2. Розробка логічних моделей на основі обраних методів .....	59
3.3.3. Вибір критеріїв та обмежень для експериментального дослідження .....	65
3.3.4. Розробка запитів для експериментального дослідження.....	67

3.4. Проектування програмного забезпечення для експерименту .....	68
3.5. Розробка фізичних моделей під СКБД MongoDB.....	71
3.6. Розробка фізичних моделей під СКБД Neo4j.....	74
3.7. Реалізація запитів для експериментального дослідження.....	76
3.7.1 Реалізація запитів під СКБД MongoDB.....	76
3.7.2. Реалізація запитів під СКБД Neo4j .....	79
3.8. Експериментальне дослідження для MongoDB.....	81
3.9. Порівняння продуктивності методів та СКБД.....	87
3.10. Розробка рекомендацій щодо використання методів .....	92
ВИСНОВКИ .....	96
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ .....	98
ДОДАТОК А. КОД ПРОГРАМИ .....	103



**ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ**

БД	–	бази даних;
ОУ	–	об’єкт управління;
ПЗ	–	програмне забезпечення;
РСУБД	-	Реляційна система керування базами даних;
СУБД	-	Система управління базами даних;
ООСУБД	-	Об’єктно-орієнтована система управління базами даних;
НФБК	-	Нормальна форма Бойса-Кодда
ЗНФ	-	Третя нормальна форма
ЕОМ	-	Комплекс пристроїв для обробки інформації
СП	–	спектральне перетворення;
ШП	–	швидке перетворення;
РС	–	personal computer.

## ВСТУП

**Актуальність теми.** Отримання достовірної інформації в сучасну епоху є серйозною проблемою, з якою стикаються організації. Це завдання вимагає швидкого доступу або до єдиного загального джерела інформації, або до добре організованої системи збирання даних з різних джерел. В останньому випадку проблемою є те, що кожне з джерел інформації зазвичай дає змогу отримати вузькоспеціалізовану конкретну інформацію, що зберігається в ньому, і це, як наслідок, тягне за собою втрату уявлення про запитуваний об'єкт загалом або спотворення інформації шляхом, наприклад, неякісної синхронізації даних. Тому завдання консолідації даних, особливо текстових, з різних незалежних джерел є актуальним.

**Об'єкт досліджень:** бази даних різного типу, включно з реляційними та ER-моделями баз даних.

**Предмет досліджень:** методи та алгоритми обробки даних у реляційних та ER-моделей баз даних.

**Мета дослідження:** виявлення тенденцій формування та проблем розвитку цифрової економіки в Україні в умовах розвитку четвертої індустріальної революції. У переліку завдань - аналіз глобальних процесів, що розгортаються, які ведуть до зміни технологічного укладу, і моделей економічного зростання; узагальнення теоретичних підходів до понять "інтелектуальна економіка" і "цифрова економіка"; аналіз стану та визначення перспектив розвитку цифрової економіки в Україні.

**Методи дослідження.** Метод теоретичного дослідження і Експеримент з використанням комп'ютера. Методологія основи дослідження є основними фундаментальним положенням для розробки програмного забезпечення, наукового дослідження вітчизняних та закордонних компаній та науковців у галузі комп'ютеризованих систем.

**Новизна запропонованих рішень.** Запропоновано визначення ситуації як усвідомленого знання ER-модель яка являє собою формальну конструкцію, сама по собі не наказує ніяких графічних засобів її візуалізації. Як стандартна

графічної нотації, за допомогою якої можна візуалізувати ER-модель, була запропонована діаграма сутність-зв'язок (ER-діаграма) є основою для побудови семантичної моделі взаємозв'язків між сутностями зв'язку. Для формалізації знань у сфері ситуаційного управління запропоновані класифікація ситуацій на основі їх типів і модальностей, онтологічна модель представлення такої класифікації.

**Практичне значення.** Розроблено інформаційну технологію, метод проектування інформаційно-аналітичних систем, що базуються на основі об'єктно-реляційної моделі даних. Результати дипломної роботи використано та впроваджено при розробці інформаційної системи, при розробці та реалізації інформаційно-аналітичної системи документообігу.