

УДК 004.8:330.14

[https://doi.org/10.52058/2786-6025-2025-13\(54\)-2060-2074](https://doi.org/10.52058/2786-6025-2025-13(54)-2060-2074)

**Сльчанінов Дмитро Борисович** кандидат технічних наук, доцент, доцент кафедри комп'ютерної математики і аналізу даних, Національний технічний університет «Харківський політехнічний інститут», м. Харків, <https://orcid.org/0000-0002-5163-9117>

**Андрейчиков Олександр Олегович** аспірант за спеціальністю F3 (122) «Комп'ютерні науки», кафедра кібербезпеки та інформаційних технологій, Харківський національний економічний університет імені Семена Кузнеця, м. Харків, <https://orcid.org/0009-0009-8496-6139>

## МОДЕЛЮВАННЯ ІНТЕЛЕКТУАЛЬНОГО КАПІТАЛУ ІТ-КОМПАНІЙ ЗАСОБАМИ ГРАФОВИХ БАЗ ДАНИХ

**Анотація.** У статті розв'язано актуальну науково-прикладну задачу підвищення ефективності управління інтелектуальним капіталом ІТ-компаній на основі застосування технологій графових баз даних. Проведено порівняльний аналіз підходів при моделюванні корпоративних знань, у результаті якого виявлено суттєві обмеження реляційної парадигми та стандартів Semantic Web (RDF/OWL) щодо гнучкості схеми, продуктивності та придатності до аналізу динамічних, слабо структурованих зв'язків у корпоративних середовищах. Обґрунтовано доцільність використання моделі маркованого графа властивостей (Labeled Property Graph, LPG) як базової парадигми репрезентації інтелектуального капіталу. Запропоновано формалізовану графову модель, яка інтегрує людський, структурний та капітал відносин ІТ-компанії в єдину семантичну мережу з атрибутивно збагаченими зв'язками, що дозволяє зберігати контекст компетенцій, досвіду та організаційної взаємодії. Здійснено програмну реалізацію запропонованої архітектури в середовищі системи керування графовими базами даних Neo4j. На основі тестового набору даних, що емулює структуру середньостатистичної ІТ-компанії, експериментально підтверджено ефективність використання механізму безіндексної суміжності для виконання складних аналітичних запитів. Зокрема, продемонстровано практичну реалізацію аналізу ризиків концентрації знань («bus factor») та алгоритм ідентифікації унікальних експертів засобами мови запитів Cypher. Визначено перспективи подальшого розвитку запропонованого підходу шляхом використання вбудованих векторних типів даних та інтеграції графової бази знань із великими мовними

моделями в межах архітектури Graph Retrieval-Augmented Generation (GraphRAG), що створює основу для побудови інтелектуальних асистентів і гібридних систем пошуку експертизи в корпоративному середовищі.

**Ключові слова:** графові бази даних, інтелектуальний капітал, bus factor, Cypher, Labeled Property Graph, Neo4j.

**Yelchaninov Dmytro** Candidate of Technical Sciences, Docent, Associate Professor at the Department of Computer Mathematics and Data Analysis, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, <https://orcid.org/0000-0002-5163-9117>

**Andreichikov Oleksandr** Postgraduate Student in the specialty F3 (122) «Computer Science», Department of Cyber Security and Information Technologies, Simon Kuznets Kharkiv National University of Economics, Kharkiv, <https://orcid.org/0009-0009-8496-6139>

## MODELING INTELLECTUAL CAPITAL USING GRAPH DATABASES

**Abstract.** The paper addresses the relevant scientific and applied problem of enhancing the efficiency of intellectual capital management in IT companies through the application of graph database technologies. A comparative analysis of corporate knowledge modeling approaches was conducted, revealing significant limitations of the relational paradigm and Semantic Web standards (RDF/OWL) regarding schema flexibility, performance, and suitability for analyzing dynamic, weakly structured relationships within corporate environments. The expediency of utilizing the Labeled Property Graph (LPG) model as the foundational paradigm for representing intellectual capital is substantiated. A formalized graph model is proposed that integrates the human, structural, and relational capital of an IT company into a unified semantic network with attributively enriched relationships, thereby allowing for the preservation of the context of competencies, experience, and organizational interactions. The software implementation of the proposed architecture was carried out within the Neo4j graph database management system environment. Based on a test dataset emulating the structure of a typical IT company, the efficiency of using the index-free adjacency mechanism for executing complex analytical queries was experimentally confirmed. Specifically, the practical implementation of knowledge concentration risk analysis («bus factor») and an algorithm for identifying unique experts using the Cypher query language are demonstrated. Prospects for the further development of the proposed approach are outlined through the use of native vector data types and the integration of the graph knowledge base with Large Language

Models (LLMs) within the Graph Retrieval-Augmented Generation (GraphRAG) architecture. This integration establishes a foundation for building intelligent assistants and hybrid expertise search systems in a corporate environment.

**Keywords:** bus factor, Cypher, graph databases, intellectual capital, Labeled Property Graph, Neo4j.

**JEL Classification:** O34, M15, M12, L86.

**Постановка проблеми.** В умовах становлення Індустрії 4.0 та переходу до концепції суспільства знань інтелектуальний капітал (ІК) поступово трансформується з допоміжного ресурсу в домінуючий фактор економічного зростання, особливо у сфері інформаційних технологій. Для сучасних ІТ-компаній ІК перестає бути лише сукупністю індивідуальних компетенцій персоналу й набуває статусу стратегічного нематеріального активу, який визначає ринкову вартість бізнесу, інноваційний потенціал та здатність ІТ-компаній до адаптації в умовах динамічних технологічних змін.

Водночас інтелектуальний капітал ІТ-компаній характеризується швидким старінням технологій, постійною зміною технологічних стеків та значною часткою неявних знань, які є персоналізованими, контекстно залежними та слабо формалізованими. Це ускладнює процеси їх формалізації, акумуляції, передачі та повторного використання в межах ІТ-компаній.

Специфіка ІТ-галузі зумовлює наявність комплексу критичних управлінських ризиків щодо ІК, які детально систематизовані у роботі [1]. Передусім йдеться про ризики, пов'язані з людським капіталом, зокрема високу мобільність кваліфікованих фахівців, що в умовах дистанційної зайнятості призводить до втрати ключових носіїв знань і зниження інноваційного потенціалу ІТ-компаній. Додатково проблема посилюється неефективним обміном знаннями та їх документуванням, внаслідок чого критична експертиза щодо архітектури програмних продуктів або бізнес-логіки концентрується у обмеженого кола осіб, формуючи ризики так званого «фактору автобуса» («bus factor»). У мультипроектному середовищі ці чинники ускладнюють формування збалансованих проектних команд та підвищують залежність результатів діяльності ІТ-компаній від окремих вузлів експертних знань. Крім того, багатовимірні взаємозв'язки між гетерогенними компонентами інтелектуального капіталу, зокрема носіями знань, технологічними стеками, програмними артефактами (вихідним кодом, документацією), клієнтами, галузевої специфікою проектів, формують складну мережеву топологію, якій притаманні нелінійність, динамічність та емерджентні властивості.

Традиційні інформаційні системи управління підприємством (ERP, HRM), а також інструменти підтримки життєвого циклу програмного забезпечення (трекери задач, системи контролю версій), які переважно

ґрунтуються на реляційній моделі даних, виявляються обмеженими у здатності забезпечити цілісне представлення інтелектуального капіталу. Їхня фрагментована, «силосна» (siload) архітектура призводить до розпорошення даних: відомості про програмний код зберігаються у репозиторіях коду (GitHub, GitLab), про задачі у трекерах задач (Jira, Trello), а про компетенції у формалізованих окремих документах (резюме) або неструктурованих знаннях менеджерів. Проблема посилюється ще й тим, що строгі схеми реляційних баз даних (БД) не дозволяють адекватно працювати з семантикою відносин, контекстними залежностями та еволюцією зв'язків між елементами ІК.

У результаті формується науково-прикладна проблема розриву між значними обсягами накопичених гетерогенних даних та обмеженими можливостями їхнього інтегрованого аналізу з метою підтримки управлінських рішень. Це обумовлює об'єктивну потребу у пошуку нових підходів до моделювання ІК, здатних відобразити його як цілісну систему взаємопов'язаних сутностей. У цьому контексті перспективним є застосування графових баз даних та семантичних технологій, які забезпечують природне представлення знань у вигляді зв'язного графа та відкривають нові можливості для аналізу, візуалізації та ефективної підтримки управління ІК ІТ-компаній.

**Аналіз останніх досліджень і публікацій.** Аналіз сучасних наукових досліджень у галузі управління даними та інженерії знань свідчить про зростаючий інтерес до графоорієнтованих підходів як ефективного інструментарію моделювання складних, слабо структурованих і динамічних систем. У контексті управління ІК ІТ-компаній особливого значення набувають роботи, присвячені порівняльному аналізу моделей даних, подоланню гетерогенності інформаційних джерел, а також інтеграції графових баз даних із сучасними інтелектуальними методами аналізу. Результати цих досліджень формують теоретико-методологічне підґрунтя для застосування графових моделей у задачах представлення та аналізу інтелектуального капіталу як цілісної системи взаємопов'язаних елементів.

Дослідження ефективності різних моделей даних ґрунтовно проведено авторами роботи [2]. У своєму дослідженні вони розглядають базу даних як абстракцію системи, яка складається з сутностей, відношень та атрибутів, і проводять порівняльний аналіз архітектурних підходів. Автори аргументовано доводять, що графові бази даних мають суттєві переваги при роботі з даними, які характеризуються «щільними зв'язками» (dense relationships), оскільки вони спроектовані для репрезентації складних мережевих структур.

Підтвердженням універсальності та високої ефективності застосування графових баз даних для аналізу складних гетерогенних систем є результати досліджень у суміжних галузях, зокрема у сфері кібербезпеки. Так у роботі [3] авторами проведено всебічний аналіз застосування графових баз даних для

моделювання кіберзагроз. Ця проблематика є структурно ізоморфною до задач управління інтелектуальним капіталом: в обох випадках ключовим є виявлення неявних взаємозв'язків між елементами системи (у випадку безпеки між векторами атак, у випадку ІК між його елементами).

Окремий клас досліджень присвячено проблемі гетерогенності даних, що є характерною рисою сучасних корпоративних інформаційних систем. В умовах, коли одна предметна область описується різними моделями даних, виникає потреба у створенні мультимодельних систем. У роботі [4] розглянуто концепцію мультимодельних баз даних як засобу вирішення конфлікту між усталеною реляційною парадигмою, яка домінує в індустрії, та графовими моделями, які набирають популярності для задач аналізу соціальних мереж, рекомендаційних сервісів, транспортних систем і т.д.

У розвитку інтелектуальних інформаційних систем актуальним є і напрям конвергенції технологій баз даних та генеративного штучного інтелекту (ШІ). Зокрема у роботі [5] досліджено фундаментальну проблему оптимізації великих мовних моделей (Large Language Model, LLM), зокрема архітектури LLaMA 3, шляхом інтеграції з графовими базами даних. Автор акцентує увагу на тому, що графова структура дозволяє вирішити критичні для LLM проблеми: покращити пояснюваність результатів генерації та забезпечити ефективне управління контекстом.

Узагальнюючи результати аналізу, можна стверджувати, що технології графових баз даних досягли високого рівня зрілості та демонструють доведену ефективність не лише у структуруванні інформації, а й у синтезі нових знань та ідентифікації нетривіальних залежностей у складних мережах. Проте, у предметній області управління інтелектуальним капіталом, зокрема в ІТ-секторі, спостерігається фрагментарність досліджень. Недостатньо опрацьованим залишається питання системного моделювання інтелектуального капіталу ІТ-компаній як єдиного графа знань, який інтегрує людський, організаційний та капітал відносин і забезпечує підтримку управлінських рішень. Це зумовлює актуальність подальших досліджень у напрямі розроблення формалізованих моделей інтелектуального капіталу на основі графових баз даних та визначає мету даної роботи.

**Мета статті.** Метою роботи є теоретико-методологічне обґрунтування використання графових баз даних для моделювання інтелектуального капіталу ІТ-компаній та розробка формалізованої моделі його подання у вигляді маркованого графу властивостей (Labeled Property Graph, LPG).

**Виклад основного матеріалу дослідження.** Сучасний ландшафт технологій управління знаннями характеризується біфуркацією підходів до моделювання та зберігання даних, що фактично відображає відмінність між теоретично орієнтованими та прикладними вимогами до представлення знань.

У межах академічної парадигми домінує підхід Semantic Web, який ґрунтується на використанні стандартів Resource Description Framework (RDF) та Web Ontology Language (OWL) для формалізованого опису предметної області. OWL забезпечує високий рівень семантичної виразності за рахунок використання класів, об'єктних і атрибутивних властивостей, а також індивідуалів, що дозволяє задавати складні логічні залежності між концептами та виконувати автоматизований вивід знань. RDF, у свою чергу, формує структурний базис для подання інформації у вигляді орієнтованого графа, де будь-які дані декомпозуються на найпростіші логічні твердження-тріплети формату «суб'єкт – предикат – об'єкт». У цій моделі суб'єкт (ресурс) через предикат (властивість) пов'язується з об'єктом (значенням або іншим ресурсом), перетворюючи складні сутності на набір розрізнених фактів. Це забезпечує універсальний механізм зв'язування гетерогенних ресурсів у єдину семантичну мережу.

Однак, попри концептуальну завершеність і формальну строгість, більшість RDF/OWL-орієнтованих рішень демонструють обмежену придатність для високодинамічних корпоративних середовищ через значні накладні витрати на обробку запитів і складність масштабування при зростанні обсягу та щільності зв'язків даних. Через те, що кожен атрибут об'єкта зберігається як окрема трійка, виникає ефект надмірної грануляції даних. Тому більшість RDF-орієнтованих рішень залишаються на рівні метамodelей або громіздких XML-подібних структур, що ускладнює їх імплементацію у системах реального часу. Додатковими стримувальними факторами є висока латентність обробки SPARQL-запитів та проблема експоненційного зростання сховищ трійок, що є критичним для динамічних доменів, зокрема ІТ-компаній.

На відміну від академічної парадигми, у прикладних корпоративних системах дедалі більшого поширення набувають графові бази даних, орієнтовані на високопродуктивну обробку мережових структур. Розвиток парадигми маркованих графів властивостей (LPG) відкриває нові можливості для побудови масштабованих і продуктивних систем управління корпоративними знаннями. LPG-модель дозволяє безпосередньо репрезентувати сутності предметної області у вигляді вершин, їх взаємодію у вигляді типізованих ребер, а контекстну інформацію у вигляді властивостей як вершин, так і зв'язків. Саме така модель є найбільш адекватною для подання ІК як динамічної мережі взаємопов'язаних акторів, знань і артефактів.

Для реалізації запропонованої моделі ІК ІТ-компаній у даній роботі обрано графову систему керування базами даних (СКБД) Neo4j [6]. Вибір цієї графової СКБД обґрунтовано низкою архітектурних і функціональних переваг.

По-перше, Neo4j використовує вбудований механізм зберігання графів, у якому кожен вузол містить прямі посилання на пов'язані з ним елементи. Це

виключає необхідність у ресурсомістких операціях об'єднання таблиць (*JOIN*) або пошуку по індексу при переході між вузлами. Така архітектура забезпечує надзвичайно швидке переміщення по графу (обхід) навіть за наявності мільйонів зв'язків, що є критичним для відстеження розгалужених ланцюжків знань у ІТ-компаніях. Наприклад, при пошуку ланцюгів передачі знань або залежностей між експертами, технологіями та проектами.

По-друге, система гарантує надійність та продуктивність на рівні корпоративних стандартів. Вона підтримує ACID-транзакції (Atomicity (атомарність), Consistency (узгодженість), Isolation (ізолюваність), Durability (довговічність)) для забезпечення цілісності даних, кластерні розгортання для високої доступності, а також використовує протокол Bolt для швидкого та ефективного обміну даними між клієнтським додатком та сервером.

По-третє, сучасна версія графової СКБД підтримує вектори як повноцінний тип даних. На відміну від звичайних списків чисел, спеціалізований тип *Vector* дозволяє системі автоматично контролювати формат та розмірність даних, які зберігаються. Це значно спрощує розробку систем ШІ, дозволяючи зберігати контекст та сутність документів чи описів компетенцій у вигляді числових векторів безпосередньо поруч із графовими зв'язками, що є основою для побудови розумного пошуку. Зокрема, у версії драйвера v6.0+ реалізовано зручний Application Programming Interface (API) для роботи з векторами. Приклад на мові Python виглядає наступним чином:

```
# From a native list → Vector  
x = [1, 2, 3]  
vector = Vector.from_native(x, "i32")  
print(f"Neo4j Vector: {vector}")
```

```
# Back to a native list  
y = vector.to_native()  
print(f"Back to native: {y}")
```

На основі обраної платформи Neo4j та розробленої концептуальної семантичної мережі спроектовано схему даних у нотації Labeled Property Graph. На відміну від жорстких схем реляційних баз даних, графова модель є гнучкою, проте для забезпечення цілісності даних у корпоративному середовищі пропонується використовувати чітку типізацію вузлів (labels) та зв'язків.

Запропонована модель декомпозує інтелектуальний капітал ІТ-компанії на чотири ключові кластери сутностей:

1. Суб'єкти (active agents):

– мітка *:Employee* відображає носіїв неявних знань з ключовими атрибутами *id*, *position*, *grade*;

– мітка *:Team* – це організаційна одиниця, яка дозволяє моделювати групову динаміку.

## 2. Об'єкти компетенцій (competence objects):

– мітка *:Skill* – це атомарна одиниця компетенції (наприклад, «Java Core», «System Architecture»);

– мітка *:Technology* – це конкретний інструментарій (наприклад, «Spring Boot», «Docker»).

## 3. Контекст діяльності та відносин (context & relations):

– мітка *:Project* – це проєктний контекст, у якому відбувається застосування знань;

– мітка *:Domain* – це бізнес-область (наприклад, медичні чи освітні технології, фінтех, тощо), знання якої є критичним активом;

– мітка *:Client* – це суб'єкт зовнішньої взаємодії (замовник), що дозволяє формалізувати капітал відносин ІТ-компанії.

## 4. Артефакти (codified knowledge):

– мітка *:Artifact* – це відчужуваний результат діяльності (документація, програмний код, стаття в базі знань).

Критичною відмінністю розробленої моделі є перенесення семантичного навантаження з вузлів на зв'язки. На відміну від реляційної парадигми, де зв'язок технічно обмежується механізмом зовнішнього ключа (foreign key) без власних атрибутів, в архітектурі LPG кожен зв'язок виступає як повноцінний інформаційний об'єкт.

Для опису інтелектуального капіталу зокрема було виділено наступні типізовані зв'язки:

– асоціація індивідуальних компетенцій, яка реалізується патерном: `(:Employee)-[:HAS_SKILL {level: 'Expert', last_validated: '2024-12-01'}]->(:Skill)`. Цей зв'язок трансформує абстрактну назву навички у конкретну характеристику ІТ-спеціаліста. Наявність подібних атрибутів у ребра дозволяє фільтрувати експертів за рівнем кваліфікації (level) та автоматично відстежувати застарівання знань через дату останньої валідації (last\_validated).

– метрики проєктного внеску, яка реалізується патерном: `(:Employee)-[:CONTRIBUTED_TO {role: 'Lead', code_lines: 540}]->(:Project)`. Така модель фіксує не лише бінарний факт участі особи у проєкті (приймав/не приймав участь), а й додає до нього якісний (роль) та кількісний (інтенсивність внеску, наприклад, кількість строк написаного коду) вимір цієї участі, що є необхідним для оцінки продуктивності персоналу.

– технологічний профіль активів, який реалізується патерном: `(:Project)-[:USES_STACK]->(:Technology)`. Цей тип зв'язку розкриває архітектурний ландшафт ІТ-компанії, визначаючи залежність конкретних проєктів від певних

технологій, що дозволяє аналізувати ризики підтримки проєктів або планувати міграцію на нові стеки.

Для імплементації запропонованої моделі використовується декларативна мова запитів Cypher. Нижче наведено приклад створення сегмента графа, який описує профіль розробника та його компетенції, з використанням векторних властивостей для подальшого семантичного аналізу та пошуку за подібністю або для інтеграції з алгоритмами ШІ.

Важливим аспектом є застосування обмежень (constraints) для гарантії унікальності сутностей, що функціонально відповідає концепції первинного ключа в реляційних СКБД. Приклад запитів на мові запитів Cypher наведено нижче.

```
// 1. Обмеження унікальності (аналог первинного ключа)
CREATE CONSTRAINT employee_id IF NOT EXISTS
FOR (e:Employee)
REQUIRE e.id IS UNIQUE;
```

```
CREATE CONSTRAINT skill_name IF NOT EXISTS
FOR (s:Skill)
REQUIRE s.name IS UNIQUE;
```

```
CREATE CONSTRAINT project_code IF NOT EXISTS
FOR (p:Project)
REQUIRE p.code IS UNIQUE;
```

```
CREATE CONSTRAINT client_name IF NOT EXISTS
FOR (c:Client)
REQUIRE c.name IS UNIQUE;
```

```
// 2. Створення або оновлення профілю співробітника (людський капітал)
// Вектор подано у скороченому вигляді для ілюстрації. У реальній
системі властивість profile_vector містить список чисел (LIST<FLOAT>)
фіксованої довжини, наприклад 1024 елементи, який використовується для
побудови векторного індексу типу VECTOR<FLOAT32>(1024).
```

```
MERGE (emp:Employee {id: 'E-1024'})
SET emp.name = 'Олексій П.',
emp.position = 'Senior Data Scientist',
emp.profile_vector = [0.12, -0.45, 0.88, 0.31, 0.54, -0.22, 0.17, 0.05],
emp.updated_at = datetime();
```

```
// 3. Ідемпотентне створення пов'язаних сутностей предметної області
MERGE (s:Skill {name: 'Graph Neural Networks'})
MERGE (p:Project {code: 'PRJ-AI-2025'})
SET p.domain = 'Artificial Intelligence',
    p.start_year = 2025
```

```
// Додавання клієнта як окремого вузла (капітал відносин)
MERGE (c:Client {name: 'FutureFinTech Inc.})
SET c.industry = 'Banking',
    c.country = 'USA'
```

```
// Встановлення зв'язку між проектом та замовником
MERGE (p)-[:DELIVERED_FOR]->(c);
```

// 4. Формування семантично збагачених зв'язків. Зв'язок компетенції з контекстними атрибутами. Використовуємо змінні (*emp*, *s*, *p*), визначені вище.

```
MERGE (emp)-[hs:HAS_SKILL]->(s)
SET hs.level = 5,
    hs.experience_years = 4,
    hs.confirmed_by = 'Certification Center',
    hs.last_used = date('2025-02-01');
```

```
// Фіксація участі ІТ-фахівця в проекті
MERGE (emp)-[wp:WORKED_ON]->(p)
SET wp.role = 'Tech Lead',
    wp.start_date = date('2025-01-10'),
    wp.status = 'Active';
```

Запропонована схема імплементації графової моделі ІК має низку принципових переваг, які забезпечують її придатність для використання у промислових інформаційних системах та аналітичних платформах.

По-перше, використання ідемпотентних операцій створення даних (*MERGE*) забезпечує коректну поведінку системи при повторному виконанні транзакцій. Це унеможливує дублювання вузлів і зв'язків у графі та є критично важливим для сценаріїв інкрементального наповнення бази даних, синхронізації із зовнішніми джерелами та безперервної інтеграції корпоративних знань.

По-друге, модель реалізує принцип семантично насичених зв'язків, коли контекстна інформація (рівень володіння компетенцією, роль у проекті, часові характеристики) зберігається безпосередньо у відношеннях між вузлами. Такий

підхід повністю відповідає парадигмі Labeled Property Graph і відображає специфіку інтелектуального капіталу, де саме зв'язки між агентами, артефактами та діяльністю несуть ключове змістовне навантаження.

По-третє, цілісність даних (data integrity) забезпечується на рівні схеми, шляхом явного задання обмежень унікальності (constraints) для базових концептів предметної області. Це функціонально відповідає механізму первинних ключів у реляційних СКБД, але при цьому не порушує гнучкості графової моделі та не накладає обмежень на еволюцію схеми.

Запропонована модель є концептуально готовою до використання методів Graph Artificial Intelligence через поєднання графової топології, атрибутивно збагачених зв'язків та векторних представлень, які дозволяють реалізувати:

- семантичний пошук на основі векторної подібності;
- графову аналітику (розрахунок показників центральності, впливовості, структурної ролі);
- механізми логічного та статистичного виводу для виявлення прихованих експертів і латентних залежностей у структурі ІК.

У сукупності ці властивості забезпечують не лише коректне зберігання знань, але й перехід від пасивних репозиторіїв до інтелектуальних систем підтримки прийняття рішень.

Експериментальна верифікація запропонованого підходу була проведена шляхом розгортання прототипу графової бази даних на платформі Neo4j та наповнення її тестовим набором даних, які емулюють структуру середньостатистичної ІТ-компанії в частині інформації про співробітників, технологічні стеки та історію реалізованих проєктів. Побудова графа здійснювалася відповідно до запропонованої онтологічної моделі та правил відображення гетерогенних знань у структуру Labeled Property Graph. Результат візуалізації фрагмента побудованого графа ІК наведено на рис. 1.

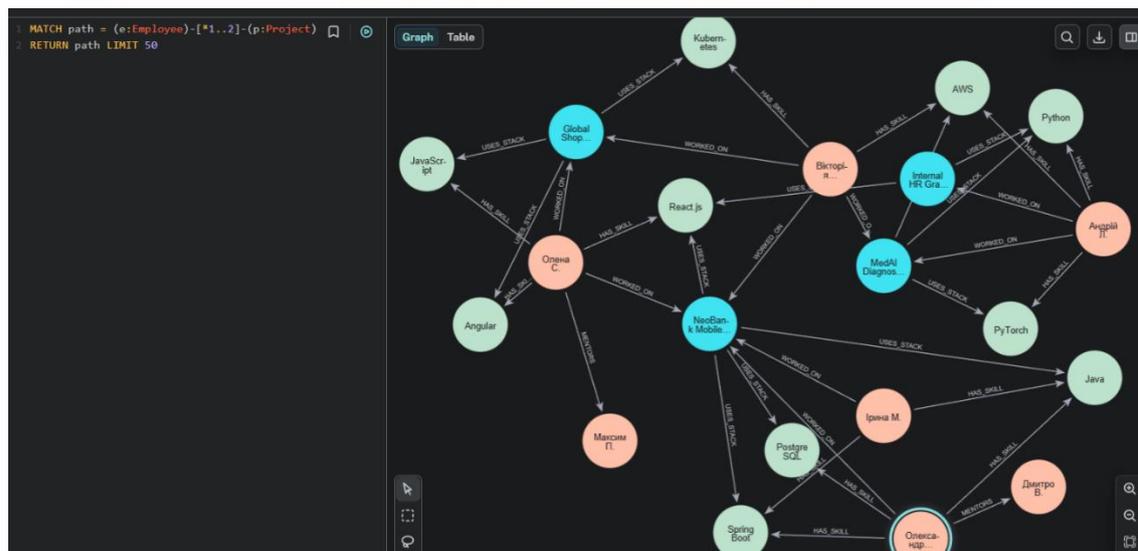


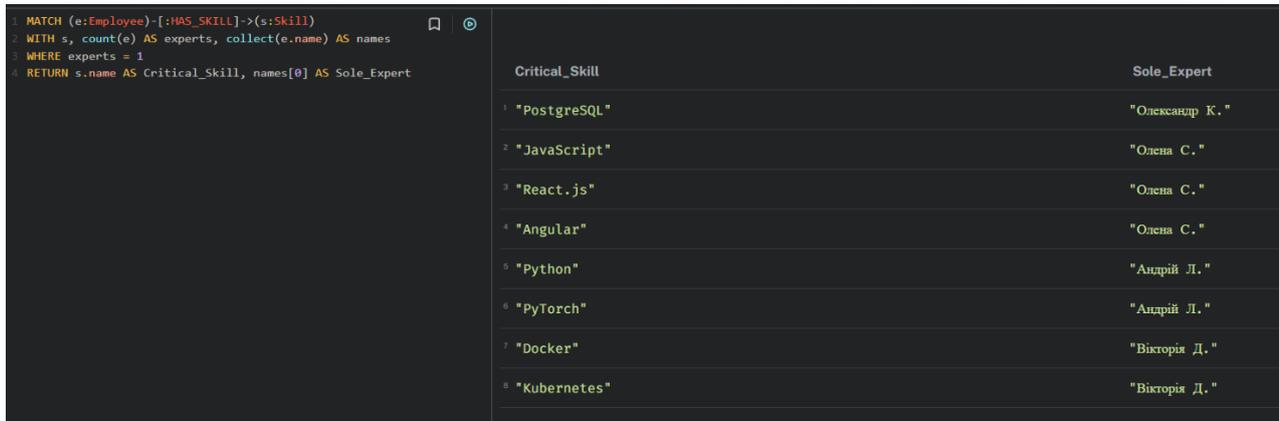
Рис. 1. Візуалізація фрагмента графа ІК ІТ-компанії в середовищі Neo4j (кластер взаємозв'язків сутностей: Employee, Skill, Project).

Джерело: створено авторами

Для демонстрації аналітичних можливостей розробленої графової системи було змодельовано сценарій виявлення критичних точок вразливості ІК з так званим аналізом «фактору автобуса» («bus factor»). Спираючись на топологію графа, де навички (Skill) представлені як вузли-концентратори, до яких сходяться ребра від різних співробітників, реалізовано алгоритм пошуку унікальних експертів. Наведений нижче Cypher-запит виконує агрегацію вхідних зв'язків (*HAS\_SKILL*) для кожної технології та відфільтровує лише ті компетенції, які мають ступінь вершини, рівний одиниці (тобто відомі лише одному співробітнику в межах ІТ-компанії):

```
MATCH (e:Employee)-[:HAS_SKILL]->(s:Skill)
WITH s, count(e) AS experts, collect(e.name) AS names
WHERE experts = 1
RETURN s.name AS Critical_Skill, names[0] AS Sole_Expert
```

Виконання цього запиту дозволяє менеджменту ІТ-компаній оперативно ідентифікувати критичні точки вразливості інтелектуального капіталу, зокрема технології або компетенції, знання про які зосереджені в одного фахівця. Отримані результати можуть бути використані для обґрунтування управлінських рішень щодо планування навчання, організації передачі знань, формування резерву експертизи та зниження операційних ризиків. Це емпірично підтверджує практичну цінність запропонованої графової архітектури для задач стратегічного управління інтелектуальним капіталом.



```
1 MATCH (e:Employee)-[:HAS_SKILL]->(s:Skill)
2 WITH s, count(e) AS experts, collect(e.name) AS names
3 WHERE experts = 1
4 RETURN s.name AS Critical_Skill, names[0] AS Sole_Expert
```

Critical_Skill	Sole_Expert
"PostgreSQL"	"Олександр К."
"JavaScript"	"Олена С."
"React.js"	"Олена С."
"Angular"	"Олена С."
"Python"	"Андрій Л."
"PyTorch"	"Андрій Л."
"Docker"	"Вікторія Д."
"Kubernetes"	"Вікторія Д."

Рис. 2. Результат виконання Cypher-запиту  
для виявлення унікальних носіїв знань (аналіз «bus factor»)  
Джерело: створено авторами

Отримана топологія характеризується високою щільністю зв'язків, що наочно підтверджує складну мережеву природу інтелектуального капіталу. Це, у свою чергу, емпірично обґрунтовує обмеженість реляційної моделі для опису такої предметної області та підтверджує доцільність використання графового підходу для її формалізації та аналізу.

**Перспективи подальшого розвитку.** Подальший розвиток запропонованого підходу вбачається у таких стратегічних напрямках:

1. Автоматизація наповнення графа: розробка інтеграційних модулів для автоматичного вилучення метаданих із виробничих систем (Jira, GitHub) та їх трансформації у графові структури в режимі, наближеному до реального часу.

2. Імплементация архітектури Graph Retrieval-Augmented Generation (GraphRAG), яка передбачає інтеграцію розробленої графової бази знань інтелектуального капіталу з великими мовними моделями (LLM) з метою забезпечення контекстно-орієнтованої генерації відповідей. У межах такого підходу граф виступає джерелом структурованого контексту, що використовується на етапі пошуку релевантних підграфів перед генерацією відповіді мовною моделлю. Це створює можливість побудови інтелектуальних корпоративних асистентів, здатних виконувати глибокий семантичний пошук експертизи та формувати обґрунтовані аналітичні висновки на основі взаємопов'язаних знань. Зокрема, можлива реалізація діалогового інтерфейсу, здатного відповідати на складні запитання, наприклад: «Хто в компанії має досвід роботи з Kubernetes у фінтех-проєктах понад 2 роки?».

3. Темпоральний аналіз: розширення моделі шляхом підтримки версійності та часових зрізів графа з метою аналізу динаміки розвитку компетенцій співробітників і накопичення інтелектуального капіталу в часі.

**Висновки.** У роботі вирішено актуальне науково-прикладне завдання моделювання ІК ІТ-компаній шляхом розробки та обґрунтування архітектури системи управління корпоративними знаннями на основі графових баз даних.

1. Теоретичне обґрунтування. Показано, що реляційна парадигма та класичні підходи Semantic Web (RDF/OWL) мають суттєві обмеження щодо продуктивності, масштабованості та гнучкості в умовах динамічного ІТ-середовища. Натомість модель Labeled Property Graph визначена як більш адекватна для репрезентації гетерогенних знань завдяки можливості зберігання атрибутів безпосередньо у зв'язках.

2. Розробка моделі. Побудовано формалізовану схему представлення інтелектуального капіталу, яка інтегрує в єдиний граф людський капітал (співробітники, ролі, команди), структурний капітал (артефакти, технології) та капітал відносин (клієнти). Запропонований підхід дозволяє долати фрагментацію корпоративних знань та усувати ефект «силосної» архітектури.

3. Технологічна реалізація. Запропоновано практичний механізм імплементації моделі засобами ГСКБД Neo4j із використанням ідемпотентних операцій Cypher та обмежень цілісності. Продемонстровано, що застосування вбудованих векторних типів даних створює основу для побудови гібридних графово-семантичних пошукових систем придатних для інтеграції з ШІ.

Практична цінність отриманих результатів полягає у формуванні масштабованої архітектурної основи для корпоративних систем управління знаннями, здатних виконувати пошук експертів, аналіз ризиків втрати знань (зокрема «bus factor») та підтримувати прийняття управлінських рішень на основі аналізу топології графа.

#### *Література:*

1. Starkova, O.V. & Andreichikov, O.O. (2025) Risk-based management of intellectual capital as a strategic resource of an IT company. Science and Information Technologies in the Modern World: Collection of Scientific Papers with Proceedings of the 2nd International Scientific and Practical Conference. International Scientific Unity. May 21-23, 2025. Athens, Greece. 230-233 p., <https://doi.org/10.70286/isu-21.05.2025>

2. Đukić, Marija; Pantelić, Ognjen; Pajić Simović, Ana; Krstović, Stefan; and Jejić, Olga "A Systematic Approach for Converting Relational to Graph Databases", IPSI Transactions on Internet Research, vol. 20(1), pp. 17-28, 2024. <https://doi.org/10.58245/ipsi.tir.2401.03>

3. Семенюк, А., & Юхимчук, М. (2025). Використання графових баз даних у кібербезпеці. Measuring And Computing Devices In Technological Processes, (1), 70–78. <https://doi.org/10.31891/2219-9365-2025-81-9>

4. Palagashvili, A.M., Stupnikov, S.A. Reversible Mapping of Relational and Graph Databases. Pattern Recognit. Image Anal. 33, 113–121 (2023). <https://doi.org/10.1134/S1054661823020098>

5. Цап, В. (2025). Застосування графових баз даних для оптимізації мовної моделі. Herald of Khmelnytskyi National University. Technical Sciences, 349(2), 447-450. <https://doi.org/10.31891/2307-5732-2025-349-65>

6. Neo4j Documentation [Електронний ресурс]. URL: <https://neo4j.com/docs> (дата звернення: 15.12.2025).

**References:**

1. Starkova, O.V. & Andreichikov, O.O. (2025) Risk-based management of intellectual capital as a strategic resource of an IT company. Science and Information Technologies in the Modern World: Collection of Scientific Papers with Proceedings of the 2nd International Scientific and Practical Conference. International Scientific Unity. May 21-23, 2025. Athens, Greece. 230-233 p., <https://doi.org/10.70286/isu-21.05.2025>

2. Đukić, Marija; Pantelić, Ognjen; Pajić Simović, Ana; Krstović, Stefan; and Jejić, Olga "A Systematic Approach for Converting Relational to Graph Databases", IPSI Transactions on Internet Research, vol. 20(1), pp. 17-28, 2024. <https://doi.org/10.58245/ipsi.tir.2401.03>

3. Semeniuk A., & Yukhymchuk M. (2025). Using Graph Databases In Cybersecurity. Measuring And Computing Devices In Technological Processes, (1), 70–78. <https://doi.org/10.31891/2219-9365-2025-81-9>

4. Palagashvili, A.M., Stupnikov, S.A. Reversible Mapping of Relational and Graph Databases. Pattern Recognit. Image Anal. 33, 113–121 (2023). <https://doi.org/10.1134/S1054661823020098>

5. Tsap, V. (2025). Application Of Graph Databases For Language Model Optimization. Herald of Khmelnytskyi National University. Technical Sciences, 349(2), 447-450. <https://doi.org/10.31891/2307-5732-2025-349-65>

6. Neo4j Documentation [Electronic resource]. URL: <https://neo4j.com/docs> (date of access: 15.12.2025)