

ГРАФИ ЗНАНЬ В ЦИФРОВІЙ ТРАНСФОРМАЦІЇ АКАДЕМІЧНИХ УСТАНОВ: ТЕХНОЛОГІЧНІ АСПЕКТИ

KNOWLEDGE GRAPHS IN THE DIGITAL TRANSFORMATION OF ACADEMIC INSTITUTIONS: TECHNOLOGICAL ASPECTS

Анотація: Розглянуто графи знань як інструмент цифрової трансформації академічних установ. Окреслено основні виклики, пов'язані з управлінням знаннями, та технологічні аспекти створення і використання графів знань. Визначено ключові характеристики, застосування в різних сферах, а також роль у персоналізації навчання. Підкреслюється важливість інтеграції даних і автоматизації процесів для підвищення якості освіти та досліджень.

Ключові слова: графи знань, цифрова трансформація, академічні установи, управління знаннями, технологічні аспекти.

Abstract: Knowledge graphs are considered as a tool for the digital transformation of academic institutions. The main challenges related to knowledge management and the technological aspects of creating and using knowledge graphs are outlined. Key characteristics, applications in various fields, and the role of personalization in education are identified. The importance of data integration and process automation for enhancing the quality of education and research is emphasized.

Keywords: knowledge graphs, digital transformation, academic institutions, knowledge management, technological aspects.

У сучасному світі цифрова трансформація стає необхідністю для всіх сфер діяльності, включаючи освіту та науку. Академічні установи зіштовхуються з викликами, пов'язаними з обробкою та управлінням величезними обсягами даних, що включають навчальні матеріали, наукові публікації, дослідницькі проекти, та інші ресурси. Ефективне управління знаннями є ключовим фактором для підвищення якості освіти, проведення інноваційних досліджень та підтримки конкурентоспроможності академічних установ.

В контексті забезпечення ефективного управління знаннями перед сучасними академічними установами постають ряд викликів:

а) *інформаційне перевантаження*: зростання кількості доступних інформаційних ресурсів та їхня різноманітність ускладнюють пошук релевантної інформації та знань;

б) *фрагментованість знань*: дані та знання розпорошені по різних системах і платформах, що ускладнює їхню інтеграцію та використання;

в) *необхідність персоналізації*: студенти та науковці потребують індивідуального підходу до навчання та досліджень, що вимагає створення персоналізованих освітніх та дослідницьких траєкторій;

г) *підтримка інновацій*: здатність швидко знаходити нові зв'язки між знаннями та створювати інноваційні рішення залежить від ефективності управління знаннями.

Управління знаннями є одним із ключових елементів цифрової трансформації в академічних установах. Воно включає процеси створення, збирання, збереження, поширення та використання знань з метою покращення освітніх, дослідницьких та адміністративних процесів. Ефективне управління знаннями сприяє підвищенню якості освіти, розвитку наукових досліджень та вдосконаленню внутрішніх процесів установи.

Knowledge Graph (граф знань) — це структура даних, яка дозволяє організувати, представляти та зберігати інформацію у вигляді взаємозв'язаних об'єктів та їх атрибутів [1]. Вона використовується для моделювання знань про різні домени. Основна мета Knowledge Graph — надавати інформацію не просто у вигляді окремих фактів, а як мережу семантично пов'язаних даних, що дозволяє розширити контекст розуміння і забезпечує складнішу інтерпретацію інформації [2].

До переліку основних характеристик Knowledge Graph слід віднести [3]:

1) *семантичний підхід до даних*: граф знань відображає не просто дані, а відносини між ними. Завдяки цьому він здатний показати, як різні елементи знань взаємопов'язані, що дозволяє краще розуміти їх значення та контекст;

2) *гнучка структура*: на відміну від реляційних баз даних, Knowledge Graph має гнучкішу структуру, де можна додавати нові типи об'єктів і зв'язків без необхідності суттєвої реорганізації всієї системи;

3) *модель мережі знань*: Knowledge Graph організований у вигляді мережі знань, де кожен об'єкт або подія пов'язаний з іншими через конкретні відносини. Це дозволяє користувачам отримувати не тільки відповіді на конкретні питання, але й додатковий контекст і нові знання;

4) *інтеграція з іншими системами*: Knowledge Graph може бути використаний для інтеграції різнорідних джерел даних (наприклад, структурованих і неструктурованих даних), що дозволяє організаціям ефективно керувати великими обсягами інформації з різних domenів.

Knowledge Graph включають в себе наступні елементи: а) вузли (Nodes) - вузлами графу є об'єкти або концепти, що представляють знання [4]. Це можуть бути люди, місця, події, терміни, об'єкти тощо; б) зв'язки (Edges) - зв'язки між вузлами відображають відносини між об'єктами. Наприклад, «авторство» між вузлом «людина» та вузлом «книга» або «місцезнаходження» між «подією» та «географічним об'єктом»; в) атрибути - вузли та зв'язки можуть мати атрибути, що містять додаткові метадані або властивості об'єктів. Наприклад, вузол «людина» може мати атрибути «ім'я», «дата народження», «місце роботи» та г) запити та пошук - Knowledge Graph підтримує запити, що дозволяють не просто знаходити конкретну інформацію, а й виконувати складніші аналітичні завдання, наприклад, пошук зв'язків між об'єктами або ідентифікацію закономірностей.

Графи знань можуть застосовуватися у різних сферах. В якості прикладів застосування Knowledge Graph можна назвати наступні:

1) *пошукові системи*: Google використовує Knowledge Graph для надання більш інформативних та пов'язаних результатів пошуку, що дозволяє відповідати на запити з урахуванням контексту і зв'язків між об'єктами;

2) *семантичний веб*: Knowledge Graph є основою для семантичного вебу, який дозволяє комп'ютерним системам не просто обробляти інформацію, але й «розуміти» значення даних та зв'язки між ними;

3) *наукові дослідження*: у науці графи знань використовуються для інтеграції та організації наукової інформації, що сприяє пошуку нових зв'язків і висновків;

4) *управління даними в організаціях*: організації використовують Knowledge Graph для побудови єдиного інформаційного простору, який об'єднує внутрішні та зовнішні джерела даних, забезпечуючи централізоване управління знаннями

Роль Knowledge Graph у цифровій трансформації зводиться до вирішення наступних задач:

1) *семантична організація даних*: Knowledge Graph дозволяє структурувати інформацію у вигляді пов'язаних даних, де зв'язки між сутностями відображають реальні відносини між знаннями. Це полегшує пошук, аналіз та повторне використання інформації;

2) *інтеграція різнорідних даних*: завдяки графам знань стає можливим об'єднання даних з різних джерел, що забезпечує повніше представлення знань та їхнє ефективне використання;

3) *автоматизація та аналітика*: Knowledge Graph підтримує автоматизацію процесів обробки даних, створення зв'язків між знаннями, а також їх аналіз, що сприяє прийняттю більш обґрунтованих рішень;

4) *персоналізоване навчання та дослідження*: використання графів знань дозволяє створювати індивідуальні траєкторії навчання та наукових досліджень, що відповідають потребам кожного студента чи науковця.

Технологічні аспекти Knowledge Graph (графів знань) включають сукупність інструментів, методів та алгоритмів, що забезпечують їх створення, обробку, зберігання та використання для різних прикладних цілей. Використання графів знань у сучасних інформаційних системах вимагає впровадження певних технологічних рішень, що дозволяють працювати з великими обсягами даних, інтегрувати різні джерела інформації та забезпечувати семантичний аналіз зв'язків між об'єктами. До переліку технологій створення графів знань слід віднести:

1) *Технології зберігання та обробки даних* - Knowledge Graph потребує спеціальних технологій

зберігання та обробки, які дозволяють працювати з графами, а не традиційними реляційними таблицями. Основні підходи включають:

а) *графові бази даних* - бази даних, спеціально створені для зберігання графових структур, де об'єкти зберігаються як вузли, а їх зв'язки як ребра. Найпоширеніші приклади графових баз даних:

- Neo4j - одна з найпопулярніших графових баз даних з потужними інструментами для роботи з великими знаннями, що підтримує мову запитів Cypher для пошуку та маніпуляції графами [5];

- Amazon Neptune - хмарна графова база даних, що підтримує як Property Graph, так і RDF-технології для зберігання семантичних даних [6];

- ArangoDB - мультимодельна база даних, що підтримує зберігання документів, ключ-значень та графів [7].

б) *RDF-технології (Resource Description Framework)* - RDF є основною технологією для представлення знань у вигляді трійок (суб'єкт, предикат, об'єкт). RDF дозволяє моделювати зв'язки між об'єктами і використовується для зберігання знань у вигляді семантичних мереж. До популярних систем для роботи з RDF відносяться:

- Apache Jena - фреймворк для роботи з RDF і SPARQL, що дозволяє створювати та керувати графами знань;

- Virtuoso - високопродуктивна платформа для зберігання та обробки RDF-графів і виконання SPARQL-запитів.

в) *SPARQL* - мова запитів для графів знань, представлених у форматі RDF. SPARQL дозволяє виконувати складні запити на пошук зв'язків між об'єктами, що зберігаються у графі.

2) *Семантичні технології та онтології* - Knowledge Graph базується на семантичних технологіях, що дозволяють описувати значення і відносини між об'єктами за допомогою формальних моделей:

а) *онтології* - формалізовані моделі певної предметної області, які описують концепти (вузли), їх атрибути та відносини між ними (зв'язки). Онтології використовуються для побудови структури знань у графі:

- OWL (Web Ontology Language) - стандарт мови для опису онтологій. OWL дозволяє створювати складні моделі взаємозв'язаних понять і забезпечує семантичне розуміння даних у Knowledge Graph;

- SKOS (Simple Knowledge Organization System) - стандарт для представлення ієрархій понять у більш простих онтологіях, що використовується для організації таксономій та тезаурусів.

б) *інструменти логічних висновків (Reasoning)* - Knowledge Graph використовує механізми логічного висновку для виведення нових знань на основі наявних зв'язків. Це дозволяє системі автоматично робити висновки про об'єкти або події, що можуть бути неявно виражені в даних. Основні підходи до логічного висновку включають:

- RDFS Reasoning - виведення нових фактів на основі RDF Schema (RDFS), що дозволяє логічно доповнювати граф новими зв'язками;

- OWL Reasoning – виведення нових знань на основі складніших онтологій, створених з використанням OWL.

3) *Інтеграція та обробка різномірних даних* - для побудови Knowledge Graph необхідно інтегрувати різні типи даних з різних джерел, таких як реляційні бази даних, документи, файли, веб-сторінки тощо. Основні технології для інтеграції та обробки різномірних даних:

а) *ETL-процеси (Extract, Transform, Load)* - процеси вилучення, трансформації та завантаження даних дозволяють зібрати дані з різних джерел, очистити їх, перетворити у відповідний формат (RDF, графи), а потім інтегрувати у Knowledge Graph;

б) *API та інтерфейси для інтеграції даних* - використання стандартних протоколів і API дозволяє легко інтегрувати дані з зовнішніх джерел, таких як веб-сервіси або інші інформаційні системи. Наприклад, GraphQL може використовуватися для спрощення запитів до графових баз даних.

4) *Алгоритми аналізу графів* - для роботи з Knowledge Graph використовуються спеціалізовані алгоритми аналізу графів, які дозволяють виявляти приховані закономірності, аналізувати зв'язки між об'єктами та виконувати інші складні операції:

а) *алгоритми пошуку найкоротшого шляху* - використовуються для пошуку найкоротшого зв'язку між двома об'єктами у графі;

б) *алгоритми кластеризації* - допомагають ідентифікувати групи об'єктів, які мають схожі властивості або тісно пов'язані між собою у графі;

в) *PageRank* - алгоритм оцінки важливості вузлів у графі, що використовується для ранжування результатів пошуку або виявлення найвпливовіших елементів у знаннях;

г) *алгоритми виявлення шаблонів* - дозволяють знаходити повторювані патерни у зв'язках між об'єктами, що може бути корисним для виявлення аномалій або тенденцій.

5) *Моделювання та візуалізація графів знань* - одним з важливих аспектів роботи з Knowledge Graph є їх візуалізація для забезпечення зрозумілого інтерфейсу користувачам, які взаємодіють з графом знань:

а) *інструменти візуалізації графів* - існує велика кількість інструментів для візуалізації графів знань, серед яких:

- Gephi - візуалізаційний інструмент для графів, що дозволяє аналізувати великі мережі знань;

- Cytoscape - інструмент для візуалізації складних графових структур та біологічних мереж;

- Graphistry - хмарна платформа для візуалізації великих графів з інтерактивним інтерфейсом.

б) *Semantic Web Browsers* - інтерфейси для візуалізації графів знань у середовищі семантичного вебу, такі як LodLive або OntoGraf, дозволяють відображати онтології та зв'язки між даними у зрозумілому вигляді.

б) **Штучний інтелект та машинне навчання** - штучний інтелект (ШІ) і машинне навчання відіграють важливу роль у розвитку Knowledge Graph, особливо для автоматичної побудови, оновлення і доповнення графів знань:

а) **автоматичне вилучення знань** - ШІ може автоматично вилучати нові знання з текстів, веб-сторінок або документів і додавати їх до графу знань. Це включає процеси NLP (обробка природної мови), такі як семантичний аналіз тексту, виявлення сутностей та відносин між ними;

б) **машинне навчання для класифікації та рекомендації** - графи знань можуть використовуватися для побудови моделей машинного навчання, які допомагають у прийнятті рішень, рекомендацій або класифікації даних.

Графи знань є важливим інструментом для цифрової трансформації академічних установ, оскільки вони сприяють ефективному управлінню знаннями та персоналізації навчання. Їх використання дозволяє інтегрувати різноманітні дані, що покращує якість освітніх та дослідницьких процесів. Подальший розвиток технологій графів знань відкриває нові можливості для інновацій у сфері освіти та науки.

Література

1. Schneider, E. W. (1972, April). Course modularization applied: The interface system and its implications for sequence control and data analysis. Association for the Development of Instructional Systems (ADIS), Chicago, Illinois. <https://files.eric.ed.gov/fulltext/ED088424.pdf>
2. Singhal, A. (2012, May 16). Introducing the Knowledge Graph: Things, not strings. Google. <https://blog.google/products/search/introducing-knowledge-graph-things-not/>
3. Ehrlinger, L., & Wöß, W. (2016). Towards a definition of knowledge graphs. https://www.researchgate.net/publication/323316736_Towards_a_Definition_of_Knowledge_Graphs
4. Ji, S., Pan, S., Cambria, E., Marttinen, P., & Yu, P. S. (2022). A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Transactions on Neural Networks and Learning Systems*, 33(2), 494-514. <https://doi.org/10.1109/TNNLS.2021.3070843>
5. Neo4j. (n.d.). Graph database & analytics: Build what others can't. Retrieved March 5, 2025, from <https://neo4j.com/product/auradb/>
6. Amazon Web Services, Inc. (n.d.). Managed graph database – Amazon Neptune features. Retrieved March 5, 2025, from <https://aws.amazon.com/neptune/features/>
7. ArangoDB. (n.d.). The rise of graph. Retrieved March 5, 2025, from <https://arangodb.com/the-rise-of-graph/>

References

1. Schneider, E. W. (1972, April). Course modularization applied: The interface system and its implications for sequence control and data analysis. Association for the Development of Instructional Systems (ADIS), Chicago, Illinois. <https://files.eric.ed.gov/fulltext/ED088424.pdf>
2. Singhal, A. (2012, May 16). Introducing the Knowledge Graph: Things, not strings. Google. <https://blog.google/products/search/introducing-knowledge-graph-things-not/>
3. Ehrlinger, L., & Wöß, W. (2016). Towards a definition of knowledge graphs. https://www.researchgate.net/publication/323316736_Towards_a_Definition_of_Knowledge_Graphs
4. Ji, S., Pan, S., Cambria, E., Marttinen, P., & Yu, P. S. (2022). A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Transactions on Neural Networks and Learning Systems*, 33(2), 494-514. <https://doi.org/10.1109/TNNLS.2021.3070843>
5. Neo4j. (n.d.). Graph database & analytics: Build what others can't. Retrieved March 5, 2025, from <https://neo4j.com/product/auradb/>
6. Amazon Web Services, Inc. (n.d.). Managed graph database – Amazon Neptune features. Retrieved March 5, 2025, from <https://aws.amazon.com/neptune/features/>
7. ArangoDB. (n.d.). The rise of graph. Retrieved March 5, 2025, from <https://arangodb.com/the-rise-of-graph/>