

система є простим рішенням для проведення аналізу значень показників міжнародного співробітництва.

Список використаних джерел:

- [1] Козіброда, С.В. (2016). Програмні засоби розробки онтологій у процесі підготовки інженерів-педагогів комп'ютерного профілю. Збірник наукових праць [Херсонського державного університету]. Педагогічні науки, (74), 175-180.
- [2] DuCharme, B. (2013). Learning SPARQL. (2nd Edition). Sebastopol: O'Reilly Media.

DOI 10.36074/26.06.2020.v1.32

СИСТЕМА АНАЛІЗУ ПОКАЗНИКІВ РІВНЯ МІЖНАРОДНОГО СПІВРОБІТНИЦТВА

Петровський Олександр Григорович

здобувач вищої освіти теплоенергетичного факультету
Національний технічний університет України «КПІ ім. Ігоря Сікорського»

НАУКОВИЙ КЕРІВНИК:

Дацюк Оксана Антонівна

старший викладач кафедри АПЕПС
Національний технічний університет України «КПІ ім. Ігоря Сікорського»

УКРАЇНА

Аналіз показників – це статистична обробка даних, яка дає змогу аналітику оцінити правильність вже прийнятих або майбутніх рішень. Більш того, аналітика в деяких випадках може зробити прогнози, що дасть змогу більш правильно розподіляти ресурси для отримання найкращого результату.

Існує багато програмних рішень, що мають загальні методи аналізу. Наприклад, Microsoft Excel дає змогу поверхнево аналізувати дані, проте для конкретних предметних областей показники є більш складними та різноманітними. Тому для аналізу показників міжнародного співробітництва у науково-технічній сфері пропонується створити власну систему, яка давала би можливість швидко отримувати складні показники або генерувати нові за допомогою запитів.

Система має п'ять підсистем (рис. 1): сховище даних, багатовимірна база даних, база даних користувачів, веб-сервер, інтерфейс користувача.

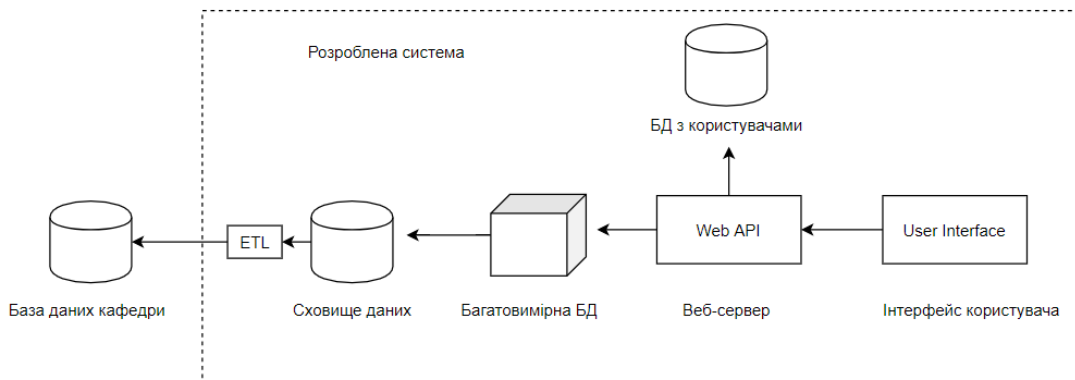


Рис 1. Структура системи

Сховище даних – це реляційна база даних, яка накопичує та зберігає лише необхідні для аналізу дані. Головними відмінностями від звичайної БД є денормалізація зв'язків між таблицями для більш швидкого виконання запитів, та два типи таблиць: таблиці-факти та таблиці-виміри. Перші зберігають корисні для аналізу числові дані, другі – представлятимуть виміри, по якими ці дані розширюватимуться. Сховище зазвичай має структуру зірки або сніжинки. Зірка утворюється, коли в центрі розташована таблиця-факт, а променями відходять таблиці-виміри. Коли зірка ускладнюється та виміри розширюють інші виміри утворюється сніжинка. Модель сховища будується виходячи з показників, які треба аналізувати. Наповнюється даними сховище із БД кафедри, яка зберігає інформацію про участь кафедри у співробітництві. Даний процес називається ETL (extract, transform, load). Під час нього нові дані БД кафедри перетворюються до потрібного вигляду та зберігаються у сховищі. ETL проводиться в автоматичному режимі щоночі, щоб не завантажувати запитами БД кафедри під час робочого дня, а після перерахунку кубів аналітик вранці отримує знову найактуальніші дані.

Багатовимірна БД побудована над сховищем даним та використовує його як джерело інформації. Основною структурною одиницею багатовимірної БД є гіперкуби, які зберігають міри – корисна для аналізу інформація та перетинають виміри. Існують два способи структурувати багатовимірну БД:

1. Створити один куб, який вміщує всі міри та перетинає всі виміри.
2. Створити багато кубів (на кожен сніжинку або зірку).

Системи, що мають невелику кількість мір, або які розробляються під спеціальне програмне забезпечення, що працює лише з одним кубом, розробляють першим способом. Проте у таких систем є кілька важливих недоліків. По-перше, єдиний куб у системі має перетинати всі виміри, проте не всі міри можуть перетинати ці виміри. На практиці це дає користувачу можливість робити перерізи по тим вимірам, які паралельні даній мірі і ніяк не впливають на результат. Тому ефективною користування такою БД буде лише за умови чіткого розуміння як цей куб був створений, що накладає поріг входу для програмного забезпечення. Псевдо перевага такої структури також розвіюється під час розширення системи. Так як усі міри та виміри знаходяться в одному кубі, при великій кількості мір, вони логічно групуються у каталоги для зручності. Пошук потрібної міри у списку каталогів абсолютно аналогічний пошуку кубу у списку кубів, проте маючи багато невеликих кубів, вони починають представляти логічно-виокремлену структуру. Користувач зможе оперувати виключно тими вимірами, які доступні для саме цього кубу. Проте якщо ще є засоби для групування мір у каталогах, то для роботи з вимірами аналогічного функціоналу немає, тому з єдиним великим кубом крім мір користувач буде вимушений шукати потрібні йому виміри у величезному списку, що також не сприятиме покращенню користувацького досвіду. Виходячи з усіх переваг структури з багатьма кубами для системи була обрана саме вона.

Головною особливістю багатовимірних БД є реалізація технології OLAP (Online Analytical Processing, аналітична обробка в реальному часі), суть якої полягає прорахунку кубу наперед – усі можливі міри агрегуються та зберігаються. Тому наступні запити виконуються набагато швидше, ніж якби розрахунок вівся під час запиту.

База даних користувачів зберігає інформацію про користувачів, ролі в системі та їхнє відношення.

Веб-сервер реалізований у вигляді трьохрівневої системи. Найнижчим є рівень доступу до даних, з якого здійснюється робота з багатовимірною базою даних. Наступним є рівень бізнес логіки, який вміщуватиме основну логіку

додатку та опрацьовує результати роботи рівню доступу до даних та приводить їх до вигляду, в якому їх необхідно повернути користувачу. Найвищим рівнем на сервері є презентаційний рівень, на якому знаходиться Web API. Його задача заключається в отриманні HTTP-запитів та генерування HTTP-відповідей. Крім цього тут реалізована авторизація користувачів через JSON Web Token.

Розроблена система є веб-додатком (рис. 2), що дозволяє використовувати її на будь-якій платформі, а система авторизації дозволить встановлювати її на робочі комп'ютери.

Кількість міжнародних наукових програм для КПІ за останні чотири роки

	КПІ
2016	77
2017	90
2018	66
2019	63

Download Excel

Show/Hide Chart

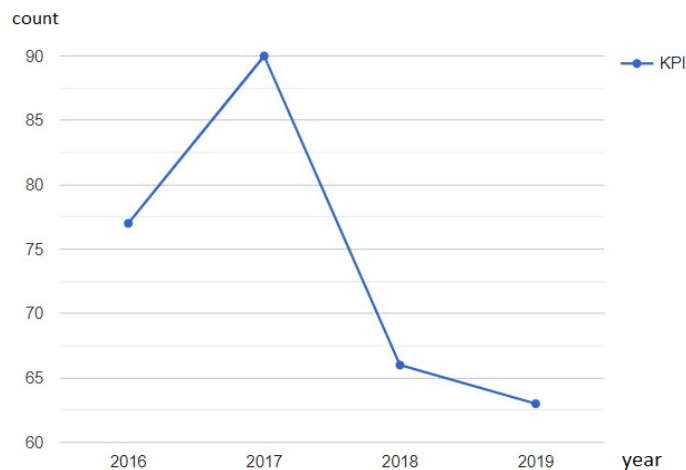


Рис 2. Приклад роботи системи

На рисунку зображено показник із кількістю міжнародних наукових програм, в який брав участь КПІ за останні чотири роки та лінійний графік, який відображає цей результат.

Запропонована система дозволяє аналізувати показники міжнародного співробітництва та реалізуючи OLAP дає можливість аналітику швидко отримувати показники, будувати нові запити за допомогою візуального конструктора, виконувати MDX-запити, зберігати результати виконання запитів у форматі Excel та будувати лінійні графіки для результату виконання запиту.

Список використаних джерел:

[1] What is Data Warehouse? Types, Definition & Example. (2020). Вилучено з: <https://www.guru99.com/data-warehousing.html>.

[2] Виртуальный куб — вместо OLAP. (2019). Вилучено з: <https://habr.com/ru/post/452154/>.