

АСПЕКТИ ПРОЯВУ НЕВИЗНАЧЕНОСТІ В ПРОЦЕСАХ РОЗРОБЛЕННЯ СИСТЕМ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ

© Верес О. М., 2015

Описано класифікацію і підходи до побудови СППР, які враховують різноманітні аспекти проблеми невизначеності процесу розроблення СППР. Запропоновано та описано ознаки класифікації, удосконалено узагальнену класифікацію СППР. Проаналізовано види архітектури, розглянуто архітектуру інформаційного ресурсу систем підтримки прийняття рішень, яка ґрунтується на принципах побудови сховища даних. Наведено аналітичний опис підходів до побудови концептуальної моделі, а також розглянуто підходи до проектування основних компонентів системи підтримки прийняття рішень. Запропоновано та описано трирівневу архітектуру концептуальної моделі СППР. Проаналізовано можливості багаторазового використання компонентів різних типів структур. Описано принципи побудови користувацької СППР з використанням багаторазових компонент. Побудована онтологія очищення даних для методологічної систематизації методів у реалізації функціональних елементів моделі СППР. Розглянуто особливості застосування “великих даних” як інформаційної технології побудови СППР. Запропоновано та описано складові елементи узагальненої формальної моделі “великих даних”.

Ключові слова: дані, класифікація, об’єктно-орієнтований підхід, архітектура, структура, онтологія, парадигма, прийняття рішення, сховище даних, система підтримки прийняття рішень.

This article describes the classification and approaches to the construction of DSS that take into account various aspects of the uncertainty of the development of DSS. Features of the classification were proposed and described. Generalized classification of DSS was improved. The analysis of the types of architecture, architecture reviewed information resource decision support systems, based on the principles of building a data warehouse. The analytical description of the conceptual approaches to building models and considered approach to the design of the main components of decision support were given. A three-level architecture and describes a conceptual model of DSS. The analysis of possible reusable components for various types of structures was carried out. We describe the principles of building a custom DSS use of multiple components. Ontology data cleaning techniques for methodological systematization of functional elements in the implementation model of DSS were built. Features of the application of Big Data as information technology to build DSS were described. Components of generalized formal model of Big Data technologies were proposed and described.

Key words: data, classification, object-oriented approach, architecture, structure, ontology, paradigm, Data Warehouse, decision making, Decision Support System.

Вступ. Загальна постановка проблеми

Стабільний розвиток економіки є запорукою підвищення конкуренції та збільшення ваги прийняття правильних рішень для ефективної роботи підприємств. Основним фактором успіху управління організаційними процесами із функціонування бізнесу є якість та швидкість прийнятих рішень. Використання комп’ютера як гаранта і підсилювача інтуїції для підтримки управлінських рішень та вирішення системних завдань – це одне з найважливіших його застосувань. Сьогодні в різних країнах світу є значна кількість робіт, які стосуються нового засобу розв’язування завдань

організаційного управління – системи підтримки прийняття рішень (СППР, англ. Decision Support System (DSS)) [1–4]. Система підтримки прийняття рішень поєднує інтелектуальні ресурси менеджера зі здібностями та можливостями комп'ютера для покращення якості рішень. Початок розвитку СППР ґрунтується на особливостях процесу прийняття рішень та зорієнтований на осіб, що їх приймають (децидент). Саме це стало запорукою застосування СППР як одного з основних напрямів використання інформаційних технологій.

Системи підтримки прийняття рішень – інформаційні системи, які якнайкраще пристосовані до виконання завдань з управління діяльністю та є засобами, що надають змогу менеджерам приймати обґрунтовані, якісні та ефективні рішення на кожному з рівнів управління організацією. СППР у режимі реального часу дає змогу автоматично аналізувати достатньо великі обсяги інформації. Засобами СППР можна розв'язувати неструктуровані й слабкоструктуровані багатокритеріальні задачі. СППР – це автоматизована інтерактивна система, яка надає змогу дециденту, використовуючи дані та моделі для виявлення та розв'язання завдань, приймати якісні рішення. Такі корпоративні інформаційні системи використовують роботу з інтерактивними запитами, моделюють ситуації прийняття рішення та формують звіти в онлайн-режимі. Основна мета СППР – підвищення ефективності прийнятих рішень.

Аналіз останніх досліджень і публікацій

СППР – це множина інтелектуальних інформаційних та інструментальних додатків, які застосовуються для маніпулювання даними, їх аналізу і надання його результатів кінцевому користувачеві. Сучасна СППР надає можливість визначати ступінь впливу прийнятих рішень на подальший розвиток бізнесу.

Отже, у СППР поєднуються на загальній основі підходи, характерні для таких напрямів досліджень [2, 4], як прийняття рішень, видобування та подання знань, а також розроблення людино-машинних (діалогових) систем.

Синергетична взаємодія цих напрямів подає СППР як якісно новий засіб процесу прийняття рішення. Більшість дослідників стверджують, що СППР призначені для вирішення слабкоструктурованих проблем. Відповідно до означення Н. Simon [2–4], до слабкоструктурованих належать проблеми, що містять як кількісні, так і якісні ознаки, причому є тенденція домінування якісних аспектів проблемної ситуації. Системи підтримки прийняття рішень – людино-машинні системи, які надають змогу дециденту, використовуючи дані, знання, об'єктивні та суб'єктивні моделі, здійснювати аналіз та вирішувати слабкоструктуровані проблеми. Концептуальна модель СППР, структура якої відповідає поданому означенню, зображена на рис. 1.

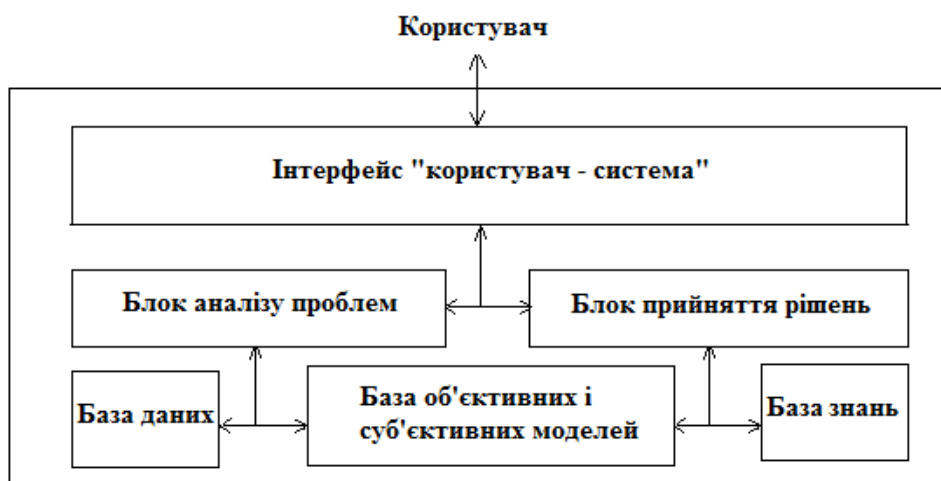


Рис. 1. Концептуальна модель СППР

До інтерфейсу “користувач-система” належать програмні засоби генерування та керування діалогом. Блоки аналізу проблем та прийняття рішень містять процедури і методи, що надають можливість за допомогою баз даних (БД), знань (БЗ) і моделей (БМ) сформулювати проблему, а

також проаналізувати способи її вирішення та отримати результат. СППР містить і засоби для побудови моделей, отримання даних та видобування знань, а також маніпулювання ними.

Аналіз наявних поглядів на розроблення та використання інформаційних систем, на способи отримання, структурування та подання інформації, на переваги інтерфейсу “користувач-система”, на специфічні відмінності СППР від інших типів автоматизованих систем управління, надає можливість вирішити проблему невизначеності типу СППР та виділити як підставу класифікації такі найсуттєвіші ознаки [3]:

- концептуальна модель;
- користувач системи;
- завдання, що вирішуються;
- засоби забезпечення;
- галузі застосування.

Різні автори вирішували проблему таксономії (класифікації) СППР. Сьогодні розроблено тільки узагальнену класифікацію СППР. Найвідоміші класифікації Пауера і Альтера (як одна з найперших [1–8]).

Не вирішені раніше частини загальної проблеми

Розроблення інформаційних систем як великомасштабний проект є надзвичайно активним і має враховувати всі аспекти функціонування такої СППР. Побудова узагальненої класифікації, а також концептуальної моделі СППР, що враховує усі підходи до розроблення таких моделей, сьогодні є актуальним завданням. Також треба дослідити та описати деякі основні аспекти проблеми невизначеності під час розроблення як моделі системи, так і її окремих функціональних компонентів.

Цілі (завдання) статті

Розроблення проекту та впровадження СППР є актуальними, особливо за інтенсивного розвитку бізнесу, вдосконалення структури організації та налагодження міжкорпоративних зв'язків. Метою роботи є дослідження проявів проблеми невизначеності підходів до розроблення концептуальної моделі СППР та компонентів її структури.

Побудова узагальненої класифікації СППР

СППР не має не тільки єдиного загальноприйнятого означення, але і вичерпної класифікації. Такі системи поділяють на декілька рівнів.

На **рівні користувача** виділяють СППР:

- *пасивні* – система, яка підтримує процес прийняття рішення, але не надає дециденту пропозиції, яке рішення треба прийняти;
- *активні* – система може запропонувати дециденту, яке рішення необхідно вибрати;
- *кооперативні* – система надає можливість дециденту змінювати, доповнювати або покращувати рішення, які вона пропонує, надсилаючи потім ці зміни в систему для повторної перевірки. Система змінює, покращує або доповнює ці рішення та надсилає їх користувачеві. Триває процес до отримання узгодженого рішення.

На **концептуальному рівні** виділяють СППР:

- *керовані повідомленнями (Communication-Driven DSS)* (групова СППР, *GDSS*) – для підтримки користувачів, які працюють як одна група над виконанням певного завдання;
- *керовані даними (Data-Driven DSS)*, або орієнтовані на роботу з даними (*Data-oriented DSS*) – орієнтуються на доступ та маніпуляції з даними;
- *керовані документами (Document-Driven DSS)* – виконують пошук та маніпулювання неструктурованою інформацією, що подана в різноманітних форматах;
- *керовані знаннями (Knowledge-Driven DSS)* – дають змогу розв'язувати завдання у вигляді процедур, фактів, правил;

- *керовані моделями (Model-Driven DSS)* – здебільшого характеризуються доступом та маніпулюванням математичними моделями (оптимізаційними, імітаційними, статистичними, фінансовими). Для побудови таких систем застосовують OLAP-системи, що надають змогу виконувати складний аналіз даних. Таку СППР класифікують як гібридну систему, яка забезпечує моделювання, пошук і опрацювання даних.

На **рівні даних**, що опрацьовує система:

- *оперативні (Executive Information Systems, інформаційні системи керівництва)* – негайно реагують на зміни в поточній ситуації з управління фінансово-господарськими процесами діяльності компанії;

- *стратегічні* – для аналізу великих обсягів різномірної інформації, яка отримується з різних джерел.

СППР на **технічному рівні** поділяють на:

- *усього підприємства* – використовує велике сховище інформації та обслуговує багато менеджерів підприємства;

- *настільна* – обслуговує тільки комп'ютер одного користувача.

На рівні **вирішуваного завдання і галузі застосування** виділяють СППР:

- *першого класу* – найбільші функціональні можливості; застосовується в органах державного управління найвищого рівня (міністерства, адміністрація президента), органах управління великих компаній (рада директорів) під час планування великих комплексних цільових програм для обґрунтування рішень щодо врахування у програмі різноманітних політичних, економічних або соціальних заходів та розподілу ресурсів між ними на основі оцінювання їхнього впливу на досягнення генеральної мети програми; це системи колективного користування, в яких бази знань формують експерти – фахівці в різноманітних галузях знань;

- *другого класу* – індивідуальні системи, бази знань яких створюють кінцеві користувачі; використовують державні службовці середнього рангу, менеджери середніх та малих фірм для розв'язання задач оперативного управління;

- *третього класу* – персональні інформаційні системи, які здатні адаптуватися до досвіду користувача; призначені для розв'язання прикладних задач управління і системного аналізу; забезпечують розв'язання поточної задачі на підставі інформації про минулий досвід практичного використання результатів розв'язання цієї задачі.

На рівні **архітектури** СППР поділяються на:

- функціональні;
- незалежні вітрини даних;
- сховище даних з двома рівнями;
- сховище даних з трьома рівнями.

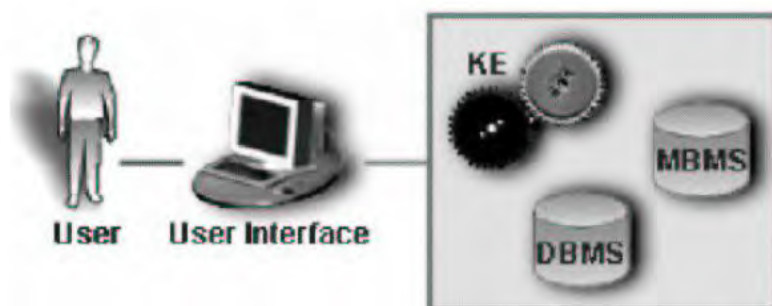


Рис. 2. Узагальнена архітектура СППР

Одним з перших узагальнену архітектуру СППР запропонував Marakas [4]. Вона містить п'ять різних частин: система управління даними (DBMS – data management system); система

управління моделями (MBMS – model management system); машина знань (KE – knowledge engine); інтерфейс користувача (user interface); користувачі (user(s)) (рис. 2).

Функціональні СППР найпростіші за архітектурою (рис. 3). Вони використовуються в організаціях, що мають невисокий рівень розвитку інформаційних технологій і не вирішують глобальних завдань.

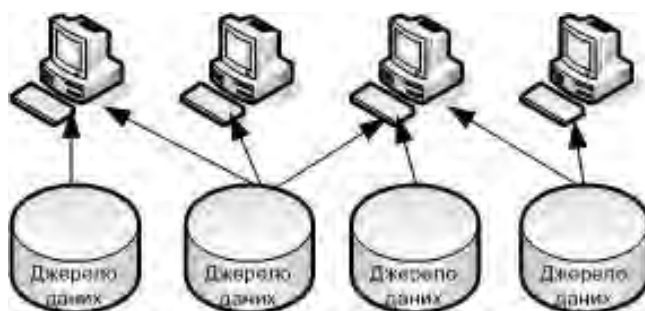


Рис. 3. Функціональна СППР

В організаціях, що мають кілька підрозділів, серед яких і відділи інформаційних технологій, застосовуються СППР з **незалежними вітринами даних** (рис. 4). Вітрина даних створюється для розв'язання певних задач та орієнтована на коло окремих користувачів, що значно підвищує продуктивність інформаційної системи.

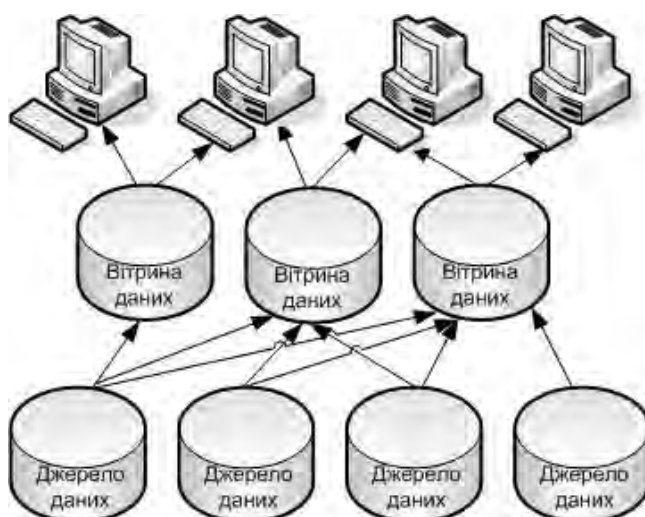


Рис. 4. СППР, що використовують незалежні вітрини даних

У великих компаніях з консолідованими даними використовується **СППР на основі дво-рівневого сховища даних (СД)** (рис. 5). У такому випадку уніфіковано отримання та способи опрацювання даних, які зберігаються в єдиному примірнику. Витрати на зберігання мінімальні. Немає проблеми синхронізації декількох копій даних. Створюється спеціалізована група осіб, яка забезпечує нормальну роботу та обслуговування СППР. Ця архітектура СППР не має недоліків попередньої, але в ній не можна структурувати дані для кожної з груп користувачів, а також виконувати адміністрування доступу до даних. Виникають труднощі з продуктивністю інформаційної системи.

СППР на основі сховища даних з трьома рівнями (рис. 6). Групи користувачів, що розв'язують однотипні задачі, використовують СД для формування вітрини даних.

СППР забезпечує доступ як до структурованих, так і до єдиних консолідованих даних. Використання перевірених та очищених даних спрощує формування вітрин даних. Сховище є єдиним централізованим джерелом корпоративної інформації. Вітрини даних синхронізовані та

сумісні з корпоративним поданням. Продуктивність таких СППР гарантована. Можна легко розширювати сховища та додавати нові вітрини даних. Наявна надлишковість даних потребує підвищених вимог до зберігання. Для кожної предметної області, залежно від вимог, потрібно погоджувати архітектуру.

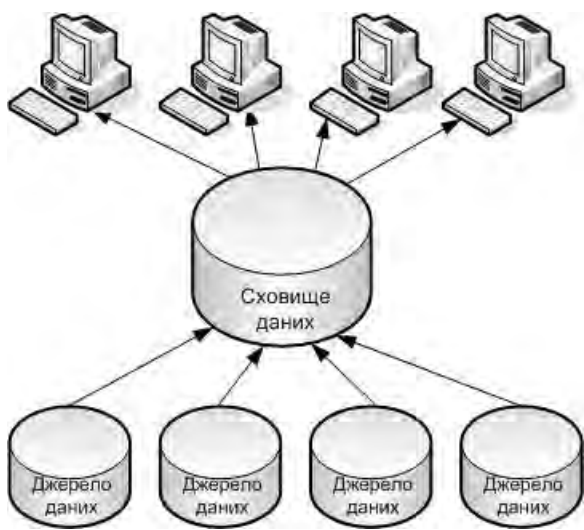


Рис. 5. СППР на основі дворівневого сховища даних

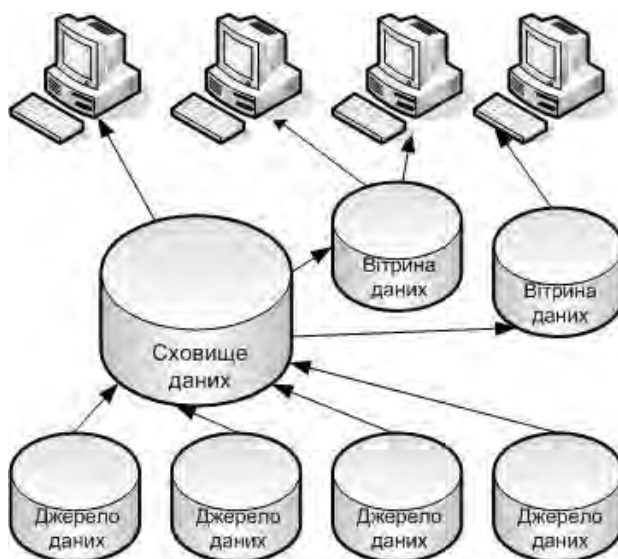


Рис. 6. СППР на основі тривірневого сховища даних

Найпростішою архітектурою СППР на основі сховища, з погляду розміщення та організації доступу до інформаційного ресурсу, є архітектура “клієнт–сервер”. На сервері (або на серверах) традиційно розміщується сховище. На клієнтах виконується аналіз даних. Особливим у цій схемі є використання вітрини даних, які також розміщуються на серверах. Здійснюється обмін даними між вітринами через перехідники (Hub Servers).

Тривірнева архітектура “клієнт–агент–сервер” – напрям розвитку архітектури систем “клієнт–сервер”. Традиційна дворівнева архітектура СППР передбачає, що на сервері розміщуються сховище або вітрина даних, а клієнт підтримує аналітичне опрацювання та інтерфейси користувача.

Особливістю тривірневої архітектури є сервер додатків, який розміщується між клієнтом і сервером (корпоративним сервером). До обов’язків корпоративного сервера належить опрацювання корпоративних даних, наприклад, зі сховища даних, а саме: організація доступу, розподіл ресурсів між клієнтами тощо. Клієнт реалізує інтерфейс користувача, а також виконує певні операції над даними та містить локальні дані. Сервер додатків як посередник між клієнтом і корпоративним сервером зменшує навантаження на нього.

Логічний поділ сховища СППР на три рівні не передбачає наявності трьох фізичних рівнів обробки. Теоретично на одній машині можна реалізувати всі три рівні. Три логічні рівні передбачають жорсткий поділ обов’язків та регламентацію зв’язків. Наприклад, до корпоративного сервера клієнт безпосередньо не може звернутися.

На рівні **ідеології логічного моделювання** СППР виділяють такі підходи:

- з використанням ідеології інформаційних систем;
- із застосуванням ідеології штучного інтелекту;
- інструментальний підхід;
- з використанням об’єктно-орієнтованої парадигми;
- інтегрований підхід.

Побудова відповідної концептуальної моделі є підґрунтям логічного моделювання інформаційних систем.

Побудова СППР відповідно до інформаційного підходу – це створення автоматизованих інформаційних систем. Застосовуючи інформаційні технології, вони здатні підвищити ефективність розумової діяльності працівників у організації (knowledge workers). Концептуальна модель СППР

R. Sprague [6] втілює особливості інформаційного підходу. Основними компонентами цієї концептуальної моделі є база даних (БД), база моделей (БМ) та інтерфейс “користувач–система” (рис. 7).

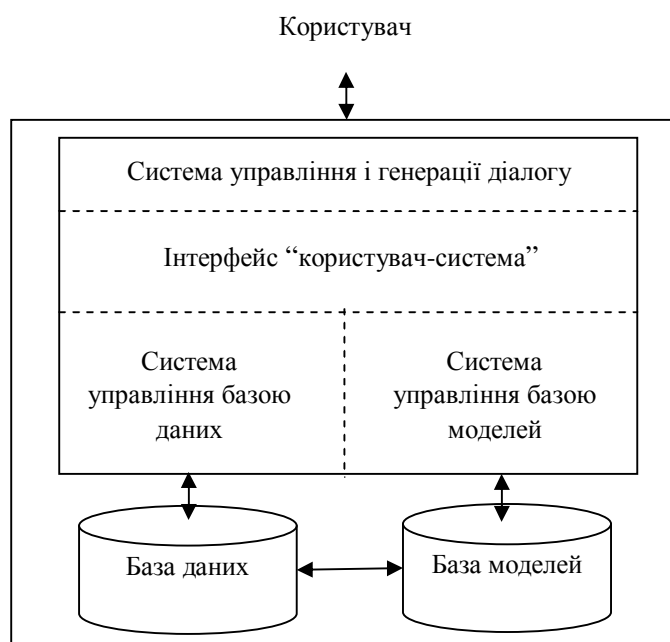


Рис. 7. Концептуальна модель СППР (інформаційний підхід)

Ефективність СППР залежить від різноманіття використовуваних даних. База даних СППР містить як кількісні, так і якісні дані з найрізноманітніших джерел. Під час реалізації СППР особливо актуальною є проблема розроблення процедур “видобування” даних з цих джерел інформації.

Важливою для СППР особливістю є здатність генерувати моделі для прийняття рішень. Передбачається, що база моделей має містити не локальні моделі, а моделі, сумісні з базою даних. Процедури моделювання повинні забезпечувати гнучкість побудови та легкість керування моделями, а саме використовувати готові блоки, підпрограми тощо. Система управління базою моделей надає можливість каталогізувати та обслуговувати найрізноманітніші типи моделей, що підтримують усі рівні управління, легко, швидко, продукувати нові моделі. Вона пов’язує ці моделі з відповідними БД та керує БМ засобами функцій управління.

“Еволюціонуюча” СППР [6], як узагальнення моделі R. Sprague, окрім інтерфейсу “користувач–система”, баз даних і моделей, містить бази текстів і правил. Додаткові компоненти доповнюють інформаційну базу СППР засобами використання слабоструктурованих видів інформації (індексовані документи, тексти природною мовою) та структурованих форм інформації (евристичні процедури, правила подання декларативних знань).

Перехід від бази текстів через бази даних і моделей до бази правил уможливило еволюційний розвиток структури СППР як у разі зміни інформаційних потреб користувача та когнітивного стилю, так і за зміни проблемної області. Сьогодні така еволюція здійснюється у процесі розроблення СППР “вручну”. Така еволюція, вважає R. Bellew [6], відбудеться тоді, коли система буде автоматично розпізнавати та утворювати “еволюціонуючі” зв’язки між компонентами структури СППР, які забезпечують динамічний обмін інформаційним ресурсом у процесі інформаційної підтримки.

Автори огляду [6] вважають, що СППР-генератори мають містити такі основні компоненти: лінгвістичний та користувацький інтерфейси; системи управління моделями і даними; система видобування даних; системна директорія. Перші чотири компоненти функціонально аналогічні до інтерфейсу “користувач-система” у моделі R. Sprague (див. рис. 7), а також мовної системи та, частково, функцій проблемного процесора в моделі R. Wozczek (застосування ідеології штучного

інтелекту) (див. рис. 8). Ціллю поділу на користувацький та лінгвістичний інтерфейс є відокремлення забезпечення користувача і системи від задач перетворення та управління повідомленнями й операціями в системі засобами різних інформаційних каналів зв'язку (різноманітне введення–виведення інформації, графіка, введення–виведення природною мовою, світлове перо тощо). Призначення системи видобування даних – формування бази даних СППР. У разі використання різних джерел інформації, дані в яких зазвичай мають різну структуру, конвертація даних виявляється достатньо складною проблемою. Системна директорія містить метазнання про систему, які надають можливість користувачеві ознайомитися із системою, зменшити потребу в повній деталізації проблеми, запобігти хибному застосуванню системи.

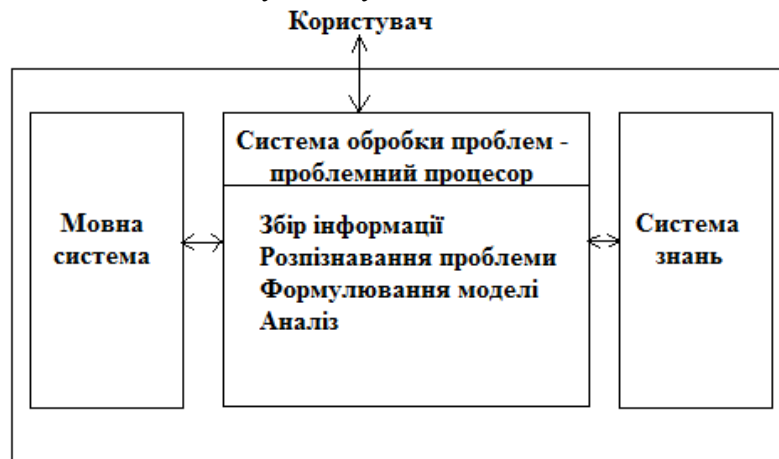


Рис. 8. Концептуальна модель СППР (підхід, побудований на знаннях)

Основними недоліками ранніх проектів СППР є велика споживча вартість та складність адаптації до швидких змін щодо дотримання вимог підтримки організації. Стрімкий прогрес у галузі інформаційних технологій сприяв дослідженню ступеня *універсальності* як основи для гнучкого та швидкого розвитку СППР. У ранніх проектах СППР підґрунтям універсальності для розвитку СППР є *функціонально-орієнтовані* моделі.

Для вирішення проблеми універсальності перспективне застосування **об'єктно-орієнтованої парадигми** до побудови концептуальної моделі СППР. Можливість багаторазового використання вважається ключем для досягнення продуктивності та якості програмного забезпечення. Процес розроблення СППР можна зробити ефективнішим, застосовуючи предметно-орієнтовані компоненти багаторазового використання, що позитивно вплине на якість такої СППР. Для досягнення цих цілей особливо важливий внесок об'єктно-орієнтованої (ОО) парадигми.

Об'єктно-орієнтований підхід застосовують професіонали (програмісти, фахівці зі знанням ОО механізмів) для створення СППР через багаторазове використання структур. Основна ідея використання багаторазових компонент полягає в тому, що за допомогою додатків компонентно-орієнтованого моделювання СППР розроблення моделі для децидента стає простим процесом вибору та застосування комплексу спеціальних, предметно-орієнтованих понять для опису вирішення проблемної ситуації [9–19]. Можна виокремити особливості таких користувацьких СППР (*DSS users*), а саме: парадигма моделювання; візуальне подання; рекомендації.

Враховуючи цілі СППР та особливості застосування об'єктно-орієнтованої парадигми для проектування, концептуальну модель СППР можна подати у вигляді архітектури з трьома рівнями, а саме: рівень семантики, презентації та рекомендації (рис. 9). Ця архітектура відповідає вимогам до інтерактивних додатків, які сприяють незалежності між застосуванням семантики й інтерфейсу [14]. Відмінність між рівнями в архітектурі СППР дає змогу ідентифікувати різні обов'язки, які мають набори класів, що представляють кожну частину, а саме:

- **рівень семантики:** його класи використовуються для опису спільного рішення, спрямованого на клас вирішення проблеми. Пропонуються методи моделювання проблеми прийняття рішення та її вирішення за допомогою запитань і процедур рішення;

- **рівень презентації:** його класи застосовують для опису структури загального представлення (таблиці, графіки тощо); може використовуватися для візуальної презентації концепцій, які подані на рівні семантики. Пропонуються методи для встановлення концептуального значення структури презентації, її графічного зображення;
- **рівень рекомендації:** його класи охоплюють всі аспекти розмовного інтерфейсу, що визначає взаємодію сценаріїв для моделювання та виконання заходів.

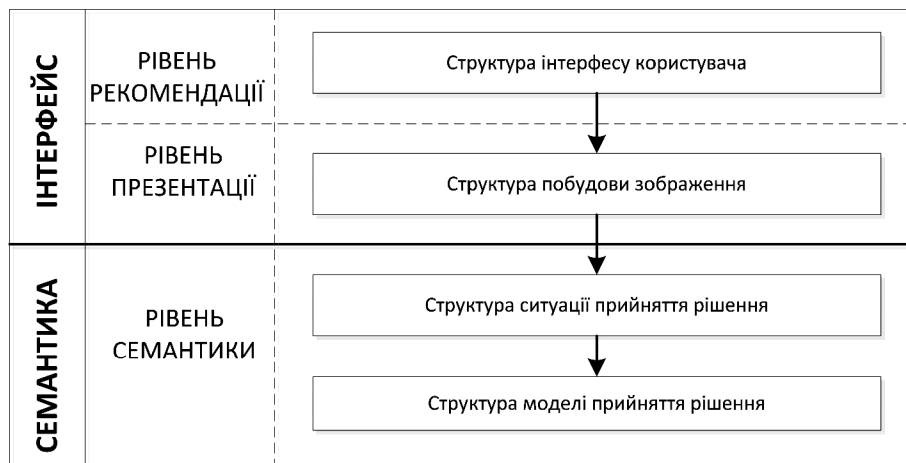


Рис. 9. Концептуальна модель СППР (об'єктно-орієнтований підхід)

Архітектура може розглядатися як структура вищого рівня для побудови цілей СППР, які складаються з нижчого рівня типів структури і принципів їхнього поєднання. Структура діє як “спеціалізована методологія”, полегшуючи діяльність із розроблення пропозицій, детально описуючи характер і особливості різних блоків та їхній взаємозв’язок [20].

Компоненти багаторазового використання рівня семантики. Моделі рішення – це представлення елементів, які характеризують проблеми прийняття рішення, для того, щоб шукати, аналізувати та оцінювати можливі рішення. Для розроблення рівня семантики СППР визначено два типи структур, а саме: структура ситуації прийняття рішення і структура моделі прийняття рішення.

Структура ситуації прийняття рішення забезпечує загальну парадигму розроблення моделі, яка орієнтована на певний клас вирішення проблем. Структура ситуації прийняття рішення – набір взаємопов’язаних класів, що представляють предметно-орієнтовані, цільові концепції моделювання, отримані з домена теорії. *Структура моделі прийняття рішення* містить загальну парадигму моделювання, покладену в основу методу прийняття рішення, і алгоритмічні деталі цього методу.

Багаторазове використання компонентів рівня презентації. Основна мета рівня презентації – отримати з рівня семантики поняття, важливі для опису та виконання вирішення проблем цього класу, надаючи користувачам СППР узгоджений і однорідний вигляд моделей, сумісних з їхнім когнітивним світом.

Багаторазове використання компонентів рівня рекомендації. Рівень рекомендації охоплює всі розмовні аспекти інтерфейсу “людина–машина”, що визначають взаємодію сценаріїв для розроблення моделі та виконання.

Компонування структур для побудови СППР. Досліджений процес декомпозиції, який дає змогу ізолювати структури на різних рівнях абстракції і з різними ролями на основі чіткого визначення цільової СППР.

Розроблення додатків є процесом складання, у якому компоненти багаторазового застосування будуть удосконалюватися і поєднуватися, щоб розробити конкретні СППР, згідно з багаторівневою архітектурою, поданою вище.

Сучасні СППР є результатом **інтегрованого** поєднання технології зберігання та накопичення даних, що ґрунтується на інформаційних сховищах, з технологією інтелектуального аналізу даних [21, 22]. Концептуальна модель такої інтелектуальної системи підтримки прийняття рішень зображена на рис. 10.

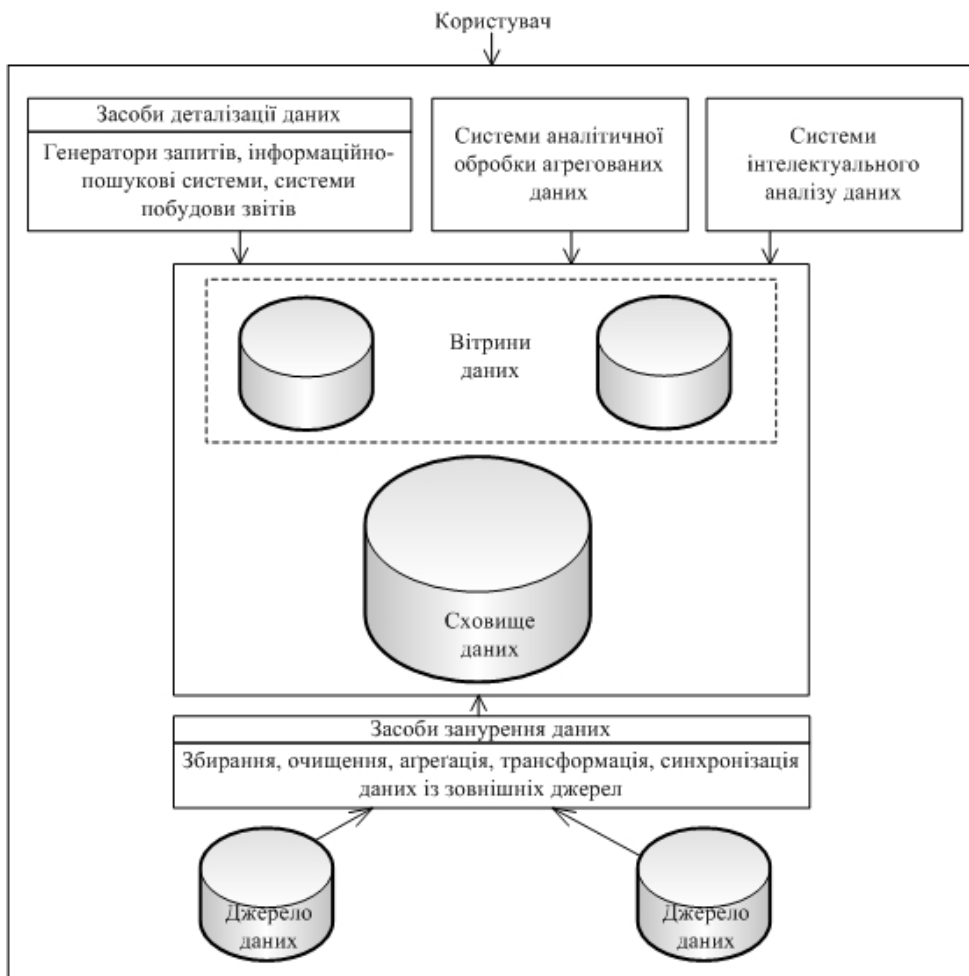


Рис. 10. Концептуальна модель СППР (інтегрований підхід)

Структура корпоративної СППР

Інтегрована концептуальна модель (див. рис. 10), основою якої є сховище даних, стала підґрунтям структури компонентів корпоративної СППР (рис. 11). Подана структура відтворює шлях даних від різноманітних джерел до столу аналітика як основу для прийняття рішення.

СППР середнього або великого підприємства чи організації має надавати доступ кінцевим користувачам до аналітичної інформації. Такі дані захищені від несанкціонованого доступу, але відкриті користувачам внутрішньої мережі організації або через Інтернет чи інтранет. Сучасна СППР має такі рівні структури:

- 1) збір та первинне опрацювання даних;
- 2) видобування, редагування та завантаження даних;
- 3) накопичення та збереження даних у сховищі даних;
- 4) подання даних у вітринах даних;
- 5) аналіз даних;
- 6) Web-портал.

Між компонентами концептуальної моделі та рівнями структури наявні тісні взаємозв'язки.

Збір та первинне опрацювання даних. Перший рівень структури СППР формують транзакційні (операційні) джерела (бази) даних. Вони є елементом OLTP-систем. У будь-якій організації цей рівень структури СППР є джерелом множини транзакційних даних.

Видобування, редагування та завантаження даних. Процес видобування, перетворення та завантаження даних підтримується ETL-засобами (Extraction, Transformation, Loading). Вони призначені для видобування даних з різноманітних джерел, їхнього перетворення та консолідації, а також завантаження до спеціальної аналітичної бази даних – сховища і вітрини даних.

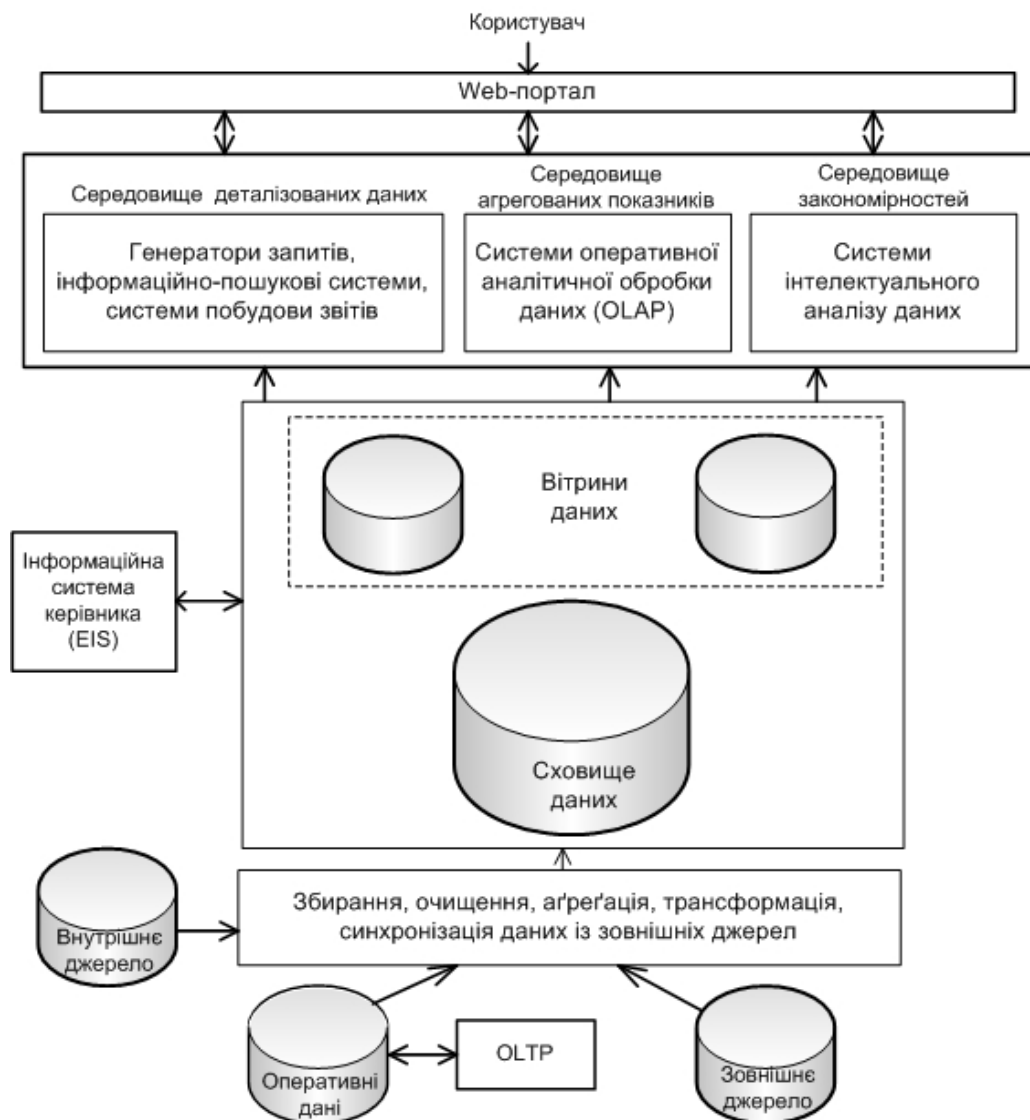


Рис. 11. Повна структура корпоративної СІПР

Для покращення якості даних застосовується комплекс методів і алгоритмів, що отримали назву “очищення даних” (*cleaning, refinement*). Щоб правильно підготувати дані до аналізу, необхідно мати стратегію їхнього очищення, яка розробляється на основі знання структури і особливостей джерел, з яких отримують дані, характеру самих даних, методики і мети їхнього аналізу. **Очищення даних** (*data scrubbing*) – виправлення успадкованих даних підприємства шляхом виявлення неузгодженостей, дублювання і помилок введення [23–26]. Очищення даних (*data cleaning, data cleansing* або *scrubbing*) – виявлення та видалення помилок і невідповідностей в даних з метою покращення їхньої якості. Очищення даних – це процес аналізу якості даних у джерелі даних з виконуваним вручну затвердженням або відхиленням рекомендацій, що дає система, і внесенням змін до даних.

Сьогодні є величезна кількість методів очищення даних від неточностей і помилок. Виявити найефективніший складно. Кожен з методів абсолютно по-різному підходить до вирішення цієї проблеми. Для вирішення проблеми невизначеності у виборі методів очищення даних розроблено онтологію.

Онтологія – це точна специфікація деякої предметної області [28–32]. Вона містить словник для подання та обміну знаннями про вибрану предметну область, а також множину зв'язків, які встановлено між термінами в словнику.

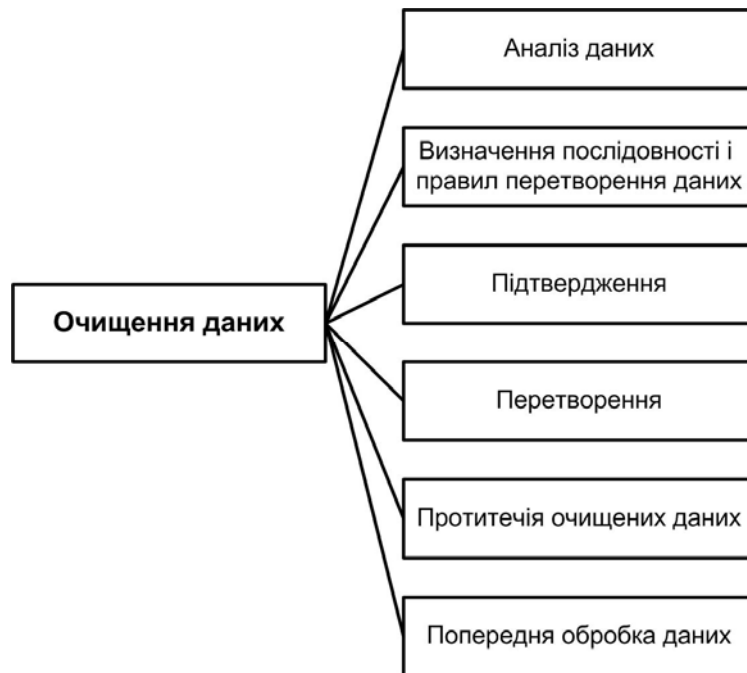


Рис. 12. Ієрархія процесу очищення даних

Пропонована онтологія побудована відповідно до підходу METHONTOLOGY [27], який відображає процес ітеративного проектування. У життєвому циклі створення онтології виділяють такі процедури: управління проектом (планування, контроль, гарантії якості), власне розроблення (специфікація, концептуалізація, формалізація, реалізація) та підтримка (супровід). Побудова онтологій – послідовність підпроцесів створення проміжних подань.

Загалом очищення даних у СППР містить декілька етапів (рис. 12) [33].

Аналіз даних – виявлення видів помилок і невідповідностей, що підлягають видаленню.

Визначення послідовності та правил перетворення даних. Залежно від кількості джерел даних, ступеня їхньої неоднорідності та забрудненості даних, вони потребуватимуть достатньо глибокого перетворення та очищення. Іноді для відображення джерел для загальної моделі даних використовується трансляція схеми; для сховищ даних зазвичай застосовують реляційне зображення. Перші кроки з очищення даних можуть скоригувати проблеми окремих джерел даних і підготувати дані для інтеграції. Подальші кроки спрямовані на інтеграцію схеми/даних та усунення проблем множинності елементів, наприклад, дублікатів.

Підтвердження – правильність і ефективність процесу і визначення перетворення, що здійснюється за допомогою тестування та оцінювання. Під час аналізу, проектування та підтвердження може знадобитися безліч ітерацій, наприклад, треба врахувати, що деякі помилки стають помітні тільки після певних перетворень.

Перетворення – виконання перетворень або в процесі ETL для завантаження і оновлення сховища даних, або під час відповіді на запити з множини джерел. Процес перетворення потребує великих обсягів метаданих – наприклад, схем, характеристик даних рівня схеми, означень технологічного процесу тощо. Для узгодженості, гнучкості та спрощення використання в інших випадках ці метадані повинні зберігатися у репозиторії на основі СУБД. Для підтримки якості даних ґрунтовна інформація про процес перетворення має записуватися як у репозиторій, так і в трансформовані елементи даних, особливо інформація про повноту та актуальність початкових даних і походження інформації про першоджерело трансформованих об'єктів та проведені з ними зміни.

Протитечія очищених даних – заміна забруднених даних у першоджерелах на очищені. Після того, як помилки (окремого джерела) видалені, очищені дані мають замінити забруднені дані в початкових джерелах, щоб покращені дані потрапили і в успадковані застосування, й надалі під час

видобування не потребували додаткового очищення. Для сховищ даних очищені дані містяться в області зберігання даних.

Попередня обробка даних – комплекс методів і алгоритмів, які використовуються в аналітичному додатку для підготовки даних до розв’язання конкретного завдання та увідповіднення їх до вимог, зумовлених специфікою завдання і способами його вирішення.

Очищення даних в аналітичному додатку – лише один з аспектів попередньої обробки даних, хоча ці два процеси часто ототожнюються. Однак очищення не є синонімом попередньої обробки даних. У випадку, якщо в даних, завантажених в аналітичний додаток, відсутні проблеми, які потребують очищення, або їхній вплив на якість рішення оцінюється як мінімальний, то очищення даних у процесі їхньої попередньої обробки може взагалі не проводитися. Попередня обробка здійснюється завжди. Потреба у цьому етапі полягає в тому, що проблеми виникають і можуть бути виявлені тільки на певному етапі збирання та консолідації даних.

Глосарій термінів онтології очищення даних містить означені вище терміни, які можна семантично розділити на три групи: структура завдання (етапи очищення, зв’язки), дані, що наповнюють задачу (методи, що застосовують на кожному з етапів), і результати обчислень (очищені дані).

Інформаційне сховище даних (Data Warehouse). Сховище даних містить джерела даних, що орієнтовані на збереження та аналіз інформації [21]. Такі джерела об’єднують інформацію з декількох транзакційних систем і надають можливість її аналізу в комплексі із використанням сучасних програмних інструментів ділового аналізу даних.

Представлення даних у вітринах даних. Метою вітрини даних (*Data Marts*) є проведення спеціального ділового аналізу. Вітрина даних – невелике інформаційне сховище зі спрощеною архітектурою, зберігає невелику підмножину даних. Вони також зменшують навантаження на основне СД організації. Вітрини даних формуються на основі інформації зі сховища даних, а у випадку, коли сховище даних в організації ще не реалізовано, створюються з даних транзакційних систем.

Аналіз даних. Наступний рівень структури СППР організації формують сучасні програмні засоби, які називаються інструментами інтелектуального, або бізнес-аналізу даних (*Business Intelligence Tools*), або *BI-додатки*. Основні складові процесу підтримки ухвалення рішень – оперативна аналітична обробка та інтелектуальний аналіз даних. Більшість систем *OLAP* сьогодні забезпечують доступ до багатовимірних даних, а засоби *ІАД*, які застосовуються у галузі закономірностей, працюють з одновимірними даних. Обидва види аналізу треба об’єднати, тобто система *OLAP* має здійснювати не тільки доступ до даних, але і пошук закономірностей. Для означення такого об’єднання К. Parsaye [34] запровадив термін “*OLAP Data Mining*” (багатовимірний інтелектуальний аналіз). Ще простішу назву запропонував J. Han – “*OLAP Mining*” і декілька варіантів інтеграції цих технологій, а саме: “*Cubing then mining*”; “*Mining then cubing*”; “*Cubing while mining*”.

Web-портал. Для надання традиційній архітектурі СППР дедалі вагомішої ролі її доповнюють *Web-порталом* [21]. Надання можливості доступу до інформації через звичайний *Web-браузер* економить витрати організації, що пов’язані зі придбанням та підтримкою настільних аналітичних програм для значної кількості клієнтських місць. Реалізація *Web-порталу* надає можливість забезпечувати аналітичною інформацією через Інтернет як користувачів всередині офісу, так і мобільних користувачів-аналітиків у будь-якій точці світу.

Застосування інформаційної технології Big Data в проєкті СППР

Дані нині оновлюються все частіше і частіше і складається ситуація, в якій традиційні методи аналізу інформації не можуть наздогнати величезні обсяги постійно оновлюваних даних, що в підсумку і відкриває дорогу технологіям великих даних. Виникає проблема вибору засобів реалізації відповідних інформаційних технологій. Її формалізація спростить вирішення проблеми невизначеності даних у СППР.

Поняття “**великі дані**” передбачає роботу з інформацією величезного обсягу і різноманітного складу, що вельми часто оновлюється та міститься у різних джерелах з метою збільшення

ефективності роботи, створення нових продуктів і підвищення конкурентоспроможності. Консалтингова компанія Forrester дає коротке означення: “Великі дані об’єднують техніки і технології, які витягують сенс з даних на екстремальній межі практичності” [35–41].

eWEEK’s подає таке, запропоноване дослідницькою компанією Gartner, означення: “Великі дані характеризуються обсягом, різноманітністю та швидкістю заливки через мережі структурованих і неструктурованих даних, в процесори і пристрої зберігання, а також перетворення цих даних на бізнес-консалтинг для підприємств” [38]. Ці елементи поділено на три категорії як визначальні характеристики великих даних (“три V”): обсяг (*volume*), різноманітність (*variety*) і швидкості (*velocity*).

На додаток до обсягу, різноманітності й швидкості, є ще один “v”, який вписується в загальну картину даних: *значення*. Точний аналіз великих обсягів даних забезпечує значення, які допомагають бізнесменам прийняти правильне рішення у потрібний час.

Великі дані (Big Data) в інформаційних технологіях – набір методів та засобів опрацювання структурованих і неструктурованих різнотипних динамічних даних великих обсягів з метою їх аналізу та використання для підтримки ухвалення рішень [41].

Bill Inmon розглядає концепцію “великих даних” як нову інформаційну технологію [42]. Великі дані – це технологія, що має такі властивості:

- можна опрацьовувати дуже великі обсяги даних;
- носії даних є недорогими;
- дані керуються методами “Римського перепису” (“*Roman Census*” *method*);
- дані, що керуються зв допомогою великих даних, є неструктурованими.

Узагальнюючи наведені означення терміна “великі дані” як інформаційної технології, можна побудувати таку формальну модель у вигляді четвірки:

$$\mathbf{BD} = \langle \mathbf{Vol}_{BD}, \mathbf{Ip}, \mathbf{A}_{BD}, \mathbf{T}_{BD} \rangle,$$

де \mathbf{Vol}_{BD} – множина типів обсягів; \mathbf{Ip} – множина типів джерел даних (інформаційних продуктів); \mathbf{A}_{BD} – множина методик аналізу “великих даних”; \mathbf{T}_{BD} – множина технологій обробки “великих даних”.

Хінчкліф ділить підходи до “великих даних” на три групи залежно від обсягу [38]:

$$\mathbf{Vol}_{BD} = \{ \mathbf{Vol}_{FD}, \mathbf{Vol}_{BA}, \mathbf{Vol}_{DI} \},$$

де \mathbf{Vol}_{FD} – “швидкі дані” (*Fast Data*): їхній обсяг вимірюється терабайтами; \mathbf{Vol}_{BA} – “велика аналітика” (*Big Analytics*): петабайтні дані; \mathbf{Vol}_{DI} – “глибоке проникнення” (*Deep Insight*): екзабайти, зетабайти.

Групи різняться між собою не тільки оперованими обсягами даних, але і якістю рішень їхньої обробки.

Для технології “великі дані” необхідне опрацювання інформації з різних за виразною потужністю типів джерел інформації – структурованих, напівструктурованих, неструктурованих. Множину інформаційних продуктів поділяють на три блоки:

$$\mathbf{Ip} = \langle \mathbf{St}, \mathbf{SemS}, \mathbf{UnS} \rangle,$$

де $\mathbf{St} = \langle \mathbf{DB}, \mathbf{DW} \rangle$ – структуровані дані (бази, сховища даних); $\mathbf{SemS} = \langle \mathbf{Wb}, \mathbf{Tb} \rangle$ – напівструктуровані дані (XML, електронні таблиці); $\mathbf{UnS} = \langle \mathbf{Nd} \rangle$ – неструктуровані дані (текст).

Над цим вектором та його окремими елементами визначено операції та предикати, які забезпечують: перетворення різних елементів вектора один на одного; об’єднання елементів однакового типу; пошук в елементах за ключовим словом.

Сьогодні наявна множина $\mathbf{A}_{BD} = \{ \mathbf{A}_i \}$ різноманітних методик аналізу масивів даних, в основі яких – інструментарій, запозичений зі статистики та інформатики (наприклад, машинне навчання).

Методики і методи аналізу, які застосовують до великих даних, описано в звіті McKinsey [43, С. 27–31]: методи Data Mining; краудсорсинг; консолідація та інтеграція даних; машинне навчання;

нейронні мережі, мережевий аналіз, оптимізація, зокрема, генетичні алгоритми; розпізнавання образів; аналітика прогнозування; імітаційне моделювання; просторовий аналіз; статистичний аналіз; візуалізація аналітичних даних.

Нові засоби для аналізу потрібні тому, що даних стає не просто більше, ніж раніше, а більше їхніх зовнішніх і внутрішніх джерел, тепер вони складніші й різноманітніші (структуровані, неструктуровані та квазіструктуровані), використовуються різні схеми індексації (реляційні, багатовимірні, noSQL). Колишніми способами впоратися з даними вже неможливо – *Big Data Analytics* поширюється на великі та складні масиви, тому ще вживають терміни *Discovery Analytics* (що відкриває аналітика) і *Exploratory Analytics* (що пояснює аналітика).

На фазах обробки “великих даних” застосовують такі технології:

$$T_{BD} = \langle T_{NoSQL}, T_{SQL}, T_{Hadoop}, T_V \rangle,$$

де T_{NoSQL} – технології NoSQL баз даних; T_{Hadoop} – технології забезпечення масивно-паралельної обробки; T_{SQL} – технології обробки структурованих даних (бази даних SQL); T_V – технології візуалізації “великих даних”.

Найчастіше як базовий принцип опрацювання великих даних використовують SN-архітектуру (*Shared Nothing Architecture*), яка забезпечує масивно-паралельну обробку, що масштабується без деградації на сотні та тисячі вузлів обробки [43, с. 31–33]. McKinsey, на відміну від пропонуєваних більшістю аналітиків інформаційних технологій **NoSQL**, **MapReduce**, **Hadoop**, **R**, також розглядає використання для опрацювання великих даних технології Business Intelligence і системи управління реляційними базами даних, які підтримують мову SQL.

Одна з найперспективніших інформаційних технологій, яку доцільно застосовувати для роботи з “великими даними”, – це простір даних. Простір даних – блоковий вектор, структурними елементами якого є множини інформаційних продуктів для подання особливостей предметної області [44, 45].

Висновки та перспективи подальших наукових розвідок

Якість персональних даних є проблемою, що значно знижує результативність аналізу. Приймати обґрунтовані рішення можна, тільки ґрунтуючись на повних і достовірних відомостях.

Застосування спеціалізованих інструментів і методів дає змогу перетворити зібрані в облікових системах дані на цінну інформацію, що використовується у процесі прийняття рішень.

Системи підтримки прийняття рішень – це доволі складний і серйозний проект. Актуальність розроблення та впровадження проекту СППР зростає з розвитком бізнесу, вдосконаленням структури організації та розвитком міжкорпоративних взаємозв'язків. Сучасним підходом до побудови таких систем є використання технології сховищ даних. Створення СППР на основі СД – складний, але доступний процес, який потребує знання не тільки бізнесу, але і програмно-технічних засобів за наявності досвіду виконання великих проектів. Побудова таких СППР вимагає сучасних технологічних рішень від основних виробників промислових СУБД і систем аналізу даних.

Запропоновано повну структуру корпоративної СППР, підґрунтям якої є узагальнена концептуальна модель СППР, що відповідає архітектурі на основі інформаційного сховища даних з трьома рівнями. Найкращим варіантом запровадження технології OLAP у проект СППР, яка пов'язана з технологіями побудови СД і методами Data Mining, є комплексний підхід.

Для вирішення проблеми універсальності СППР проаналізовано застосування об'єктно-орієнтованого підходу для швидкого та легкого проектування СППР, яке ґрунтується на багаторазовому використанні структур різних типів. Запропоновано трирівневу архітектуру концептуальної моделі СППР, яка відображає розвиток компонентів багаторазового використання, що мають різний характер, а саме: структури ситуації прийняття рішень, структури моделі прийняття рішення, структури побудови зображення і структури розроблення інтерфейсу користувача. Описано типи компонентів багаторазового використання та послідовність компонування користувацької СППР. Цей підхід можна розглядати як проектування структур вищого рівня, які

складаються зі структур нижчого рівня і рекомендацій з використання цих структур для розроблення конкретних СППР.

Застосування онтології як інструменту пізнання забезпечує системний підхід до вивчення предметної області. При цьому досягаються:

- систематичність – онтологія відображає цілісний погляд на предметну область;
- однотипність (стандартність) – матеріал, який подано в однаковому форматі, набагато краще сприймається та відтворюється;
- науковість – побудова онтології уможливує повноту відновлення потрібних, але відсутніх логічних взаємозв'язків.

Для досягнення поставленої мети здійснено аналіз етапів очищення даних у корпоративних СППР. Розглянуто особливості очищення даних на рівні аналітичних додатків як структурних елементів корпоративної СППР. Можливість безпосереднього управління підготовкою даних важлива для аналітика ще й тому, що в деяких випадках характер його втручання може відрізнитись від формальних процедур.

Описані засоби є достатньо сучасними, але вони не вирішують всіх проблем і все ще потребують додаткової обробки вручну або додаткового програмування. Очищення даних потрібне не тільки для сховищ даних, але і для обробки запитів за неоднорідними джерелами даних, наприклад, в інформаційних WEB-системах. Це середовище має істотніші обмеження для очищення даних, які потрібно враховувати, вибираючи відповідні методи. Також актуальним для вивчення сьогодні є моделювання очищення даних у Big Data. Подальші дослідження також будуть спрямовані на дослідження методів очищення частково структурованих даних, оскільки структурні обмеження постійно знижуються, а обсяги XML-даних стрімко зростають.

Big Data – це не спекулятивні міркування, а символ прийдешньої технічної революції. Необхідність аналітичної роботи з “великими даними” помітно змінить ІТ-індустрію та стимулюватиме появу нових програмних і апаратних платформ.

Для вирішення проблеми невизначеності застосування інформаційної технології Big Data побудовано її формальну модель та описано основні структурні елементи.

Сьогодні для аналізу великих обсягів даних застосовують найпередовіші методи: штучні нейронні мережі – моделі, побудовані за принципом організації та функціонування біологічних нейронних мереж; методи предиктивної аналітики, статистики та Natural Language Processing (напрями штучного інтелекту та математичної лінгвістики, які вивчають проблеми комп'ютерного аналізу та синтезу природних мов). Використовуються також і методи, які залучають людей-експертів, чи краудсорсинг, А/В тестування, сентимент-аналіз тощо. Для візуалізації результатів застосовуються відомі методи, наприклад, хмари тегів і зовсім нові Clustergram, History Flow і Spatial Information Flow.

“Великі дані” підтримуються технологією розподілених файлових систем Google File System, Cassandra, HBase, Lustre і ZFS, програмними конструкціями MapReduce і Hadoop і безліччю інших рішень. За оцінками експертів, наприклад, McKinsey Institute, під впливом “великих даних” найбільшої трансформації зазнає сфера виробництва, охорони здоров'я, торгівлі, адміністративного управління та контролю над індивідуальними переміщеннями. У подальших роботах досліджуватимуться методи, моделі та інструменти для ефективнішої підтримки загальної діяльності з розроблення моделі системи підтримки прийняття рішень.

1. Power D. J. What is a DSS? // *The On-Line Executive Journal for Data-Intensive Decision Support*, 1997. – V. 1. – N3. 2. Ситник В. Ф. Системи підтримки прийняття рішень: навч. посіб. / В. Ф. Ситник. – К. : КНЕУ, 2004. – 614 с. 3. Ларичев О. И. Системы поддержки принятия решений. Современное состояние и перспективы их развития / О. И. Ларичев, А. В. Петровский // *Итоги науки и техники. Сер. Техническая кибернетика*. М.: ВИНТИ, 1987. – Т. 21. – С. 131–164. 4. Ларичев О. И. Теория и методы принятия решений: учебник / Ларичев О. И. – М. : Логос, 2000. –

296 с. 5. Alter S. L. *Decision support systems: current practice and continuing challenges* // Reading, Mass.: Addison-Wesley Pub., 1980. 6. Scott Morton M. S. *Management Decision Systems: Computer-based Support for Decision Making*. – Boston: Harvard University, 1971. 7. Трахтенгерц Э. А. *Компьютерная поддержка принятия решений : научн.-практ. изд. / Э. А. Трахтенгерц. – М. : СИНТЕГ, 1998. – 376 с. – (Серия “Информатизация России на пороге XXI века”)*. 8. Интеллектуальный анализ данных в системах поддержки принятия решений / М. Шанот // *Открытые системы*. 1998. – № 1. – С. 30–35. 9. Johnson P. *Object-oriented technology: the competitive advantage* / P. Johnson // *GEC*, 9(1): 28–41, 1993. 10. Nierstrasz O. *Component-oriented software development* / O. Nierstrasz, S. Gibbs, D. Tsichritzis // *Communications of ACM*, 35(9): 160–165, Sept. 1992. 11. Becker K. *Reusable object-oriented specifications for decision support systems* / K. Becker, F. Bodart // In: *IFIP WG 8.1 Working Conference on the Object Oriented Approach in Information Systems – Quebec City, Oct. 28–31, 1991. Proceedings*. North-Holland, 1991. – P. 137–155. 12. Hartson H. *Human-computer interface development : concepts and systems for its management* / H. Hartson, D. Hix // *ACM Computing Surveys*, 21(1):5–93, March 1989. 13. Dolk D. R. *Knowledge representation for model management systems* / D. R. Dolk & B. Konsynski // *IEEE TSE*, 10(6): 619–628, Nov. 1984. 14. Blanning R. W. *A relational theory of model management* / R. W. Blanning // In: *Holsapple C.W. & Whinston A.B. (eds.). Data base management: theory and applications*. Springer-Verlag, 1987. – P. 15–53. 15. Ma J. *An object-oriented framework for model management* / J. Ma // *Decision Support Systems*, 13(2): 133–149, Feb. 1995. 16. Muhana W. A. *An object-oriented framework for model management and DSS development* / W. A. MUHANA // *Decision Support Systems*, 9(2): 217–229, Feb. 1993. 17. Muhana W. A. *Symms: a model management systems that supports model reuse, sharing and integration* / W.A. Muhana // *European Journal of Operations Research*, 72(2): 214–242, Jan. 1995. 18. Верес О. М. *Застосування об’єктно-орієнтованого підходу до побудови моделі СППР* / О. М. Верес // *Вісн. Нац. ун-ту “Львівська політехніка”. Серія : Інформаційні системи та мережі. – 2013. – № 770. – С. 30–36*. 19. Верес О. М. *Компоненти концептуальної моделі системи підтримки прийняття рішень* / О. М. Верес // *Вісн. Нац. ун-ту “Львівська політехніка”. Серія : Комп’ютерні науки та інформаційні технології. – 2010. – № 686. – С. 103–112*. 20. Спирли Э. *Корпоративные хранилища данных. Планирование, разработка, развитие* / Э. Спирли. – М. : Вильямс, 2001. – 400 с. 21. Арустамов А. *Предобработка и очистка данных перед загрузкой в хранилище [Электронный ресурс]* / Алексей Арустамов. – Режим доступа: http://www.basegroup.ru/library/dw_olap/data_clearing/. 22. Фоурино Р. *Электронное качество данных: скрытая перспектива очистки данных [Электронный ресурс]* / Роналд Фоурино. – Режим доступа: http://www.olap.ru/basic/el_data_quality.asp/. 23. Рам Э. *Очистка данных: проблемы и актуальные подходы [Электронный ресурс]* / Э. Рам, Хонг Хай До. – Режим доступа: http://www.olap.ru/basic/data_clean.asp/. 24. Некипелов Н. *Онтология анализа данных [Электронный ресурс]* / Н. Некипелов, А. Шахиди. – Режим доступа : <http://www.basegroup.ru/library/methodology/ontology/>. 25. Гаврилова Т. А. *Базы знаний интеллектуальных систем* / Т. А. Гаврилова, В. Ф. Хорошевский. – СПб. : Питер, 2000. – 384 с.: ил. 26. Гаврилова Т. А. *Онтология для изучения инженерии знаний* / Т. А. Гаврилова // *Труды Международной научн.-практ. конференции KDS-2001. – 2001*. 27. Гаврилова Т. А. *Онтологический подход к управлению знаниями при разработке корпоративных информационных систем* / Т. А. Гаврилова // *Новости искусственного интеллекта. – 2003. – № 2. – С. 24–30*. 28. Буров С. В. *Формальна модель подання знань у системі онтологічного моделювання задач* / С. В. Буров // *Вісн. Нац. ун-ту “Львівська політехніка”. Серія : Інформаційні системи та мережі. – 2013. – № 770. – С. 21–30*. 29. Литвин В. В. *Метод використання онтологій у петлі OODA* / В. В. Литвин // *Вісн. Нац. ун-ту “Львівська політехніка”. Серія : Інформаційні системи та мережі. – 2014. – № 783. – С. 137–145*. 30. Литвин В. В. *Проблема автоматизованої розбудови базової онтології* / В. В. Литвин, Т. М. Черна // *Вісн. Нац. ун-ту “Львівська політехніка”. Серія : Інформаційні системи та мережі. – 2014. – № 805. – С. 306–315*. 31. Верес О. М. *Онтологія очищення даних* / О. М. Верес // *Вісн. Нац. ун-ту “Львівська*

політехніка”. Серія : Інформаційні системи та мережі. – 2015. – № 814. – С.237–245. 34. Parsaye K. OLAP and Data Mining: Bridging the Gap / Parsaye K. // *Database Programming and Design*. 1997. – № 2. 35. Большие данные (Big Data) / [Электронный ресурс] / – Режим доступа: http://www.tadviser.ru/index.php/Статья:Большие_данные_%28Big_Data%29#cite_note-g-6. 36. Олхорст Ф. 2012–й: “облачный” год для “больших данных” / [Электронный ресурс] / Ф. Олхорст. – Режим доступа: <http://www.pcweek.ua/themes/detail.php?ID=136094>. 37. Черняк Л. Большие Данные – новая теория и практика / [Электронный ресурс] / Л. Черняк // *Открытые системы. СУБД*. – М.: Открытые системы, 2011. – № 10. – Режим доступа: <http://www.osp.ru/os/2011/10/13010990>. 38. Preimesberger C. Hadoop, Yahoo, 'Big Data' Brighten BI Future практика / [Электронный ресурс] / Chris Preimesberger // *EWeek* (15 August 2011). – Режим доступа: <http://www.eweek.com/c/a/Data-Storage/TBA-Hadoop-Yahoo-Big-Data-Brightens-BI-Future-254079>. 39. Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data. Gartner (27 June 2011). <http://www.gartner.com/newsroom/id/1731916>. 40. Моррисон А. Большие Данные: как извлечь из них информацию / А. Моррисон и др. // *Технологический прогноз. Ежекварт. журнал, рос. изд.* – 2010. – Вып. 3. PricewaterhouseCoopers (17 декабря 2010). 41. Шаховська Н. Б. Організація Великих даних у розподіленому середовищі / Н. Б.Шаховська, Болюбаш Ю. Я., Верес О. М. // *Збірник наукових праць ДонНТУ “Обчислювальна техніка та автоматизація”*. – 2014. – № 2. – С. 147–155. 42. Inmon W. H. Big Data – getting it right: A checklist to evaluate your environment / [Электронный ресурс] / W. H. Inmon // *DSSResources.COM*, – 2014. – Режим доступа: <http://dssresources.com/papers/features/inmon/inmon01162014.htm>. 43. Manyika J. Big data: The next frontier for innovation, competition, and productivity / J. Manyika et al. // *McKinsey Global Institute*, June, 2011. 156 с. 44. Шаховська Н. Б. Формальне подання простору даних у вигляді алгебраїчної системи / Шаховська Н. Б. // *Системні дослідження та інформаційні технології / Національна академія наук України, Інститут прикладного системного аналізу*. – К., 2011. – № 2. – С. 128–140. 45. Shakhovska N. B. Big Data Federated Repository Model / N. B. Shakhovska, Yu. Ja. Bolubash, O. M. Veres // *The Experience of Designing and Application of CAD Systems in Microelectronics (CADMS'2015) Proc. of the XIII Int. Conf., (Polyana-Svalyava (Zakarpattia), Ukraine, 24–27 February, 2015)*. – Lviv: Publishing Lviv Polytechnic, 2015. – P. 382–384.