

КОНСОЛІДОВАНИЙ ІНФОРМАЦІЙНИЙ РЕСУРС ЕКОНОМІЧНИХ ПОКАЗНИКІВ РЕГІОНУ

© Марковець В.О., Макосій О.О., 2011

Зібрано та проаналізовано інформацію від об'єктів господарювання стосовно процесів консолідації даних, отриманих як із внутрішніх, так і з зовнішніх джерел. Побудовано на її основі сховища даних, розроблено засоби інтеграції та аналізу розрізненої вхідної інформації економічної сфери регіону.

Ключові слова: сховище даних, економічна сфера регіону, бізнес-каталог, інтеграція.

In this article information from the management objects regarding the data consolidation processes obtained from both internal and external sources is collected and analyzed. On this basis data warehouse is built. Integration and analysis means of diverse input data of the region economic sphere are developed.

Key words: data warehouse, economic sphere of region, business catalog, integration.

Вступ

Одним із завдань економіки є побудова комплексної картини розвитку галузей, визначення перспективних об'єктів, аналіз ринків товарів та послуг тощо. Саме з цією метою необхідно розробити засоби централізованого збору інформації від об'єктів господарської діяльності, її аналізу та розповсюдження.

Дослідження інтеграції даних економічної сфери в сьогоденні дає змогу оцінити розвиток технічних засобів, технологій, додатків і продуктів для ефективної інтеграції даних. У фокусі нового дослідження перебуває не тільки роль інтеграції даних у проектах по створенню сховищ, але й технологічна база керування метаданими.

У сучасних ринкових умовах відомо безліч інформації з приводу цін, пропозицій, вигідності тих чи інших товарів та послуг. Але така інформація є доволі хаотичною, з неї повною мірою не можна робити жодних висновків про стан економічної сфери та проблем, які в ній існують. Якщо ж намагатися розібратися детальніше, все ж певні довідкові дані можна віднайти та проаналізувати їх. Саме тому, щоб не припуститися помилок та врахувати індивідуальні потреби кожного клієнта потрібно застосовувати сховища даних.

Постановка проблеми

На етапах розроблення та супроводу сховищ даних часто виникають проблеми, починаючи із самої постановки задачі, утворення об'єктів та, власне, самого сховища, що охоплює всі корпоративні ресурси метаданих.

Варто зазначити, що обсяги ринку будь-якого регіону сьогодні є порівняно сталими, тому, на мій погляд, найважливішим тепер є не тільки збільшення цих обсягів, а передусім спрямування зусилля на вирішення нагальних інформаційних проблем в економічній галузі, про які йтиметься нижче.

Головною проблемою, яка постає під час розроблення сховища даних економічної сфери, є відсутність єдиної системи та політики інтегрування розрізнених даних з метою їх загального використання. Зрозуміло, що неузгоджене створення, накопичення та актуалізація великого обсягу такої інформації потребує витрачання зайвих коштів і не дає очікуваного результату. Враховуючи це, сьогодні немає системи, яка б могла використовуватись для комплексного інформаційного

забезпечення, формування аналітичних та прогнозних даних та підтримки прийняття рішень стосовно гармонійного розвитку господарства в регіонах.

Розгляд проблеми опрацювання невизначеності дозволить вирішити питання підвищення достовірності та надійності інформації, а, отже, підвищить ефективність рішень вироблених системою.

Аналіз останніх досліджень

Для прискорення і спрощення роботи в ринковій галузі сьогодні фірми створили низку програмних засобів автоматизації роботи, серед яких є:

- Fidelio Front Office – «Фіделіо» Система Портъє;
- Gemini Labs – Pollux Hospitality Management System (Pollux HMS);
- «1С: Бухгалтерія»;
- «1С: Підприємство».

Кожний із цих програмних засобів побудований на основі баз даних, що не дає змогу ефективно виконувати операції аналізу. Вони охоплюють вузький спектр задач, що реалізуються в торговельній та інших галузях, в прийманні, розміщенні та відвантаженні товарів, у виконанні робіт, сфері обслуговування тощо.

Жоден з розглянутих програмних продуктів не може одночасно керувати багатьма процесами, а тим більше приймати рішення з управління предметною галуззю. Тому виникла необхідність розроблення інших засобів автоматизації роботи у окремих сферах економіки.

Adelite: ETL. Програмний продукт Adelite: ETL слугує для витягання даних з сімейства програм 1С: Підприємство, перетворення їх (очищення, виключення дублікатів, приведення до єдиного формату ощо) і завантаження в базу даних на MS SQL Server з подальшою можливістю проведення їх багатовимірної аналізу.

Серед основних можливостей Adelite: ETL варто відзначити такі:

- програмний продукт взаємодіє з платформами «1С: Підприємство 7.7» і «1С: Підприємство 8.0»;
- баз даних може бути декілька і необов'язково вони мають бути однаковими;
- інкрементне вивантаження – вивантаження тільки нових або змінених даних; така форма вивантаження дозволяє економити час на оновлення даних у вже побудованому кубі;
- автоматичний режим роботи – дозволяє автоматично оновлювати дані у вивантаженні; всі операції, що виконуються програмою, записуються в log-файл; після виконання всіх запитів або при помилці програма автоматично закривається;
- аналіз структури метаданих 1С – програма сама аналізує структуру метаданих, тому необов'язково добре знатися на структурах баз даних 1С і файлах конфігурацій для впровадження OLAP на підприємстві;
- автоматичний вибір об'єктів, пов'язаних з вибраними, що дозволяє не відстежувати структури початкової бази;
- фільтрація вивантажуваних даних – масштабованість на багатопроесорних машинах (використання потужності декількох процесорів дає вииграш в часі вивантаження);
- збереження наборів налаштувань у файл, що використовуються для автоматичної роботи програми.

EXmile. Інтеграція зі сховищами даних. Програмний комплекс «EXmile. Інтеграція зі сховищами даних» дозволяє трансформувати великі обсяги даних з різноманітних джерел в єдиний стандарт, значно скорочуючи трудовитрати на інтеграцію різних прикладних систем.

Завдяки зручному графічному інтерфейсу, значно спрощується процедура настройки правил експорту даних з довільних інформаційних систем (автоматизованих банківських систем, бек-офісних додатків тощо), а також формату XML-документів для даних, що вивантажуються.

Система «eXmile. Інтеграція зі сховищами даних» включає такі модулі:

- eXmile. Designer, призначений для налаштування алгоритмів генерації XML документів. Модуль дає змогу набудовувати структуру XML документів і відповідні ним SQL-запити до баз даних;
- eXmile. Agent, призначений для пакетного вивантаження даних за розкладом через готові налаштування;
- eXmile.Worksheet, призначений для виконання і відлагодження SQL-запитів через інтерфейс Microsoft ADO. Використання цього модуля дозволяє застосовувати один налагоджувач запитів для різних джерел даних, що істотно спрощує розроблення сценаріїв вивантаження прикладних даних.

Це рішення є оптимальним інструментом для співробітників відділів автоматизації, відповідальних за впровадження сховищ даних, розробку процедур вивантаження даних з різних облікових систем.

Система «eXmile» надає можливість:

- роботи з базами даних через технологію доступу до даних Microsoft ADO (налаштування параметрів з'єднання, виконання певних серверних дій до і після з'єднання);
- експорту даних з декількох різних баз даних натисненням однієї кнопки;
- налаштування вихідних форм у вигляді файлів заданої структури у форматі XML (відповідність вхідних і вихідних даних через запити до СУБД на мові SQL; виконання переліку серверних операцій з підготовки даних; опис параметрів запитів до СУБД, вказання шляху експорту і принципів генерації імен вивантажуваних xml-файлів на підставі параметрів запиту);
- формування в інтерактивному і пакетному режимах вихідних файлів у форматі XML відповідно до налаштувань;
- протоколювання процесу формування вихідних файлів.

Важливою перевагою рішення є універсальність, гнучкість і простота освоєння. Система дозволяє здійснювати інтеграцію різних інформаційних систем шляхом налаштування процедур вивантаження прикладних даних через XML-шлюз, легко змінювати налаштування форматів вихідних XML-документів, застосовувати складні алгоритми підготовки даних перед їх експортом в необхідному форматі, використовувати особливості реалізацій мови SQL, характерних для різних СУБД.

Загальна характеристика. Всі ETL-інструменти, що відомі сьогодні, можна розділити на дві великі групи. З одного боку, окремі постачальники, що працюють з реляційними або багатовимірними серверами СУБД, включають в свої рішення вузькоспеціалізовані засоби завантаження даних, орієнтовані виключно на свої сервери. До таких продуктів можна віднести, наприклад, Oracle Express Relational Access Manager і Hyperion Essbase. З іншого, існують і універсальні засоби, які забезпечують завантаження з альтернативних джерел даних в різних типів цільових приймачів, проводячи при цьому різні перетворення.

У межах свого комплексного вирішення Informix Decision Frontier (IDF) компанія Informix пропонує два універсальні ETL-інструменти: DataStage і Formation. Крім того, до IDF належать, зокрема, такі продукти: Extended Parallel Server, Red Brick Decision Server, MetaCube, Seagate Info.

Extended Parallel Server – реляційний сервер СУБД, який в межах Decision Frontier позиціює як засіб підтримки первинного сховища детальних фактів або ODS (Operational Data Store – оперативне сховище даних), що містить сукупну актуальну інформацію операційного середовища.

Red Brick Decision Server – спеціалізований ROLAP-сервер, що володіє додатковими функціями (оптимізація запитів типу star join, матеріалізоване зберігання результатів агрегація, оптимізація масового завантаження тощо), які орієнтують його на підтримку вітрин даних.

MetaCube – засіб аналітичного оброблення даних (побудова звітів, графіків, пошук тенденцій тощо).

Seagate Info – набутий програмний продукт, призначений для побудови звітів за накопиченими даними (функціональність частково перетинається з MetaCube).

Усі розглянуті програмні продукти є дорогими та потребують додаткових налаштувань для роботи у сфері економіки, оскільки тут необхідно інтегрувати не тільки первинну інформацію, але й звітну. Тому виникла потреба у розробленні нового програмного продукту.

Формулювання цілі статті

На етапах роботи з інформацією стосовно економічної галузі регіону постає необхідність застосування сховищ даних, щоб ефективно і оптимально вирішувати проблеми аналітичного складу. За допомогою сховищ даних можна легше реалізувати надскладні задачі, наприклад, визначення сфер покриття регіонів постачальниками та споживачами та відшукати кореляцію між цими сферами.

Щоб опрацювати аналітичне завдання сховищ даних економічної сфери регіону, потрібно звернути увагу на наявність факторів привабливості окремих сегментів:

1. Виробнича інфраструктура – це комплекс галузей, що обслуговують основне виробництво і забезпечують його ефективну економічну діяльність. До них належать транспорт, зв'язок, електроенергетика, кредитно-фінансові заклади, спеціалізовані галузі ділових послуг.

2. Інформаційна інфраструктура – комплекс програмно-технічних засобів, організаційних систем та нормативних баз, який забезпечує організацію взаємодії інформаційних потоків, функціонування та розвиток засобів інформаційної взаємодії та інформаційного простору регіону.

3. Система забезпечення інноваційно-інвестиційної діяльності – весь необхідний спектр державних і приватних установ та організацій, які забезпечують розвиток і підтримку всіх стадій інноваційного циклу.

4. ГІС-технології (геоінформаційні технології) – технологічна основа створення географічних інформаційних систем, що дозволяють реалізувати їхні функціональні можливості. Інформаційно-обчислювальна система, призначена для фіксації, збереження, модифікації, керування, аналізу і відображення усіх форм географічної інформації.

5. Соціальна сфера – комплекс галузей, які безпосередньо пов'язані зі створенням загальних умов для відтворення робочої сили і забезпечення нормальної життєдіяльності людей. Ця сфера охоплює освіту, охорону здоров'я, житлово-комунальне господарство, пасажирський транспорт і зв'язок, культуру, побутове обслуговування.

6. Інфраструктура ринку цінних паперів:

- регулятивна – система регулювання ринку цінних паперів, що включає систему правил поведінки учасників ринку і систему методів впорядкування укладених угод;

- функціональна – система торгівлі цінними паперами (посередники, фондові магазини, фондові біржі, організований та неорганізований позабіржовий ринок, аукціони з продажу цінних паперів), яка забезпечує процес укладання угод з цінними паперами;

- технічна – система розрахунків з укладання угод, збереження і обліку цінних паперів, що включає депозитарну діяльність (збереження цінних паперів, облік прав на цінні папери); діяльність з ведення та збереження реєстру власників цінних паперів; розрахунково-клірингова діяльність;

- інформаційна – включає інформаційну систему (збір, обробку інформації про емітентів, новий випуск, ціни і місця торгівлі цінними паперами) і систему зв'язку (об'єднує засоби зв'язку учасників ринку цінних паперів).

Діаграма (рис. 1) розрахована залежно від існуючих факторів привабливості сегментів чи регіонів. Враховуючи очевидність необхідності кожного з даних факторів, можна змодельовати закономірність, за якою інвестор спрямовуватиме нові фінансові вливання у визначені господарські об'єкти відповідно до його очікувань.

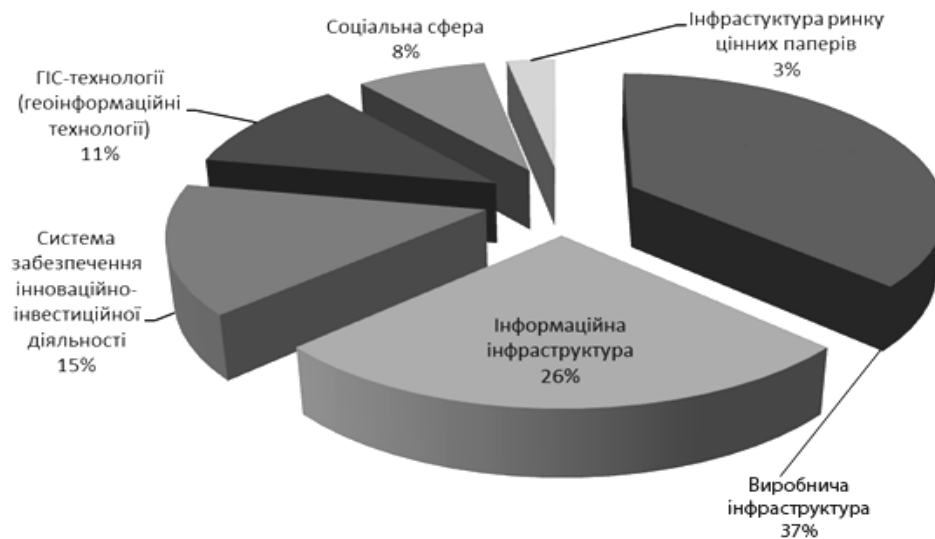


Рис. 1. Фактори привабливості регіонів та сегментів економіки

Для точнішої оцінки цієї задачі треба враховувати географічне положення, особливості транспортної інфраструктури, соціальні та політичні умови кожного конкретного регіону. Нині потрібно не оминати досить вагому проблему: існує неповноцінне використання інформаційних ресурсів. Ситуація ускладнюється тим, що інформаційний попит найбільш високий в щільно населених урбанізованих районах, оскільки ця територія використовується доволі інтенсивно усіма сферами господарства.

Розгляд проблеми опрацювання дасть змогу вирішити питання підвищення достовірності та надійності інформації, а, отже, підвищить ефективність рішень, вироблених системою.

Задачі, що виникають у предметній області

Виділимо необхідні умови та задачі, що має вирішувати пакет програм для комплексного інформаційного забезпечення, формування аналітичних та прогнозних даних та підтримки прийняття рішень стосовно гармонійного розвитку галузі:

1. Облік:
 - нормативної та законодавчої бази;
 - структури адміністративно-територіальних одиниць;
 - природних умов;
 - структури населення;
 - культурно-освітнього забезпечення населення;
 - транспортної мережі та сполучення.
2. Аналіз:
 - структури населення;
 - зайнятості та структури зайнятості;
 - впливу господарської діяльності на навколишнє середовище, прогнозування та попередження надзвичайних ситуацій;
 - розвитку господарської діяльності;
 - забезпеченості житлово-комунальними послугами.
3. Планування та прогнозування:
 - обсягів забезпечення та розвитку господарської діяльності;
 - забезпеченості житлово-комунальними послугами;
 - культурно-освітнього забезпечення населення;
 - в галузі капітальних ремонтів та будівництва;
 - природних умов проживання населення;
 - демографічного стану території;

- споживання мінеральних, паливно-енергетичних, водних, біологічних, земельних ресурсів у господарській діяльності;
- розвитку рекреаційної сфери;
- аналіз та планування функціональної структури території;
- аналіз та планування умов життєдіяльності;
- аналіз ефективності та планування економічної діяльності;
- типове проектування.

Залежно від типу об'єкта інформація може зберігатися у різних моделях та надходити з різних джерел:

- виробниче підприємство – база даних, динамічний Web-сайт з базою даних, розміщеною на Web-сервері;
- адміністративна одиниця – сховище даних;
- особа – Web-сайт, база даних, текстові дані тощо;
- суб'єкт торгівлі – база даних, Web-сайт.

Оскільки специфікою багатьох галузей є подання інформації в Інтернет у вигляді реклами, замовлень тощо, то на найвищому рівні ієрархії моделей даних знаходяться колекції іменованих ресурсів з базовими властивостями – розмір, дата створення і тип (наприклад, зображення JPEG, база даних MYSQL).

Одним з основних завдань збору даних економічної сфери регіону є каталогізація елементів даних учасників.

Бізнес-каталог – це реєстр суб'єктів господарювання, що містить найбільш базову інформацію про кожний з них: ринкова одиниця, ім'я, місцезнаходження в економічному та географічному розрізі, розмір, дата створення, власник і т.д. Каталог є інфраструктурою для більшості інших сервісів простору економічних даних, але він також може підтримувати базовий, призначений для конкретного користувача інтерфейс перегляду даних.

Бізнес-каталог містить не тільки описову інформацію (тобто виконує роль метаданих), але й зберігає для кожного учасника схему бізнес-одиниці, статистичні дані, швидкість зміни, точність, можливості відповідей на запити, інформацію про власника і дані про політику доступу і підтримку конфіденційності. Оскільки господарські одиниці простору даних фізично не переносять у нього інформацію та можуть обмінюватись між собою інформацією, то у бізнес-каталозі необхідно зберігати дані і про зв'язки між окремими суб'єктами.

Подання джерел даних до каталогу схематично подані на рис. 2. Джерелами даних сховища даних є бази даних (реляційні, об'єктно-реляційні або багатовимірні) та оперативні сховища даних, а джерелами простору даних є об'єкти, подані у довільній моделі. Тому у каталозі необхідно також вказувати тип джерела та засоби його опрацювання (програмні продукти, стандарти передачі тощо).



Рис. 2. Надходження даних до бізнес-каталогу

На рис. 3 подана схема застосування інформаційних технологій в економічній сфері.



Рис 3. Сфери застосування інформаційних технологій в економічній сфері

Концептуальна модель зображена на рис. 4.

Створення і функціонування бізнес-одиниці опосередковане інформаційними відносинами із засновником та трудовим колективом, постачальниками і покупцями, бюджетом та позабюджетними фондами, страховими компаніями і банками, а також відносини, пов'язані з отриманням і розподілом власних доходів і накопичень, формуванням і використанням відповідних фондів грошових коштів.



Рис. 4. Концептуальна модель інформаційної діяльності бізнес-одиниці

Інформаційні відносини між суб'єктами господарювання відбуваються:

- із засновниками – з приводу формування статутного капіталу;
- із акціонерами – з приводу викупу власних акцій для їх дострокового погашення;
- із страховальниками – з приводу внесків до страхового полісу; виплата страхових сум; оплати гарантійних послуг;
- із державою – з приводу сплати податків та зборів до бюджету; внесків до позабюджетних фондів.

Всередині суб'єкта існують інформаційні відносини як з працівниками, так і між відділами та підрозділами.

Кожна бізнес-одиниця також вступає у інформаційні відносини з грошово-кредитною системою:

- з банками – для відкриття розрахункових рахунків, розміщення депозитних вкладів;
- з страховими компаніями – для страхування транспортних засобів, майна, персоналу;
- з позабюджетними фондами – стосовно сплати платежів.

Основна робота аналізу фінансово-господарських операцій включає в себе чотири процеси: аналіз ринку ресурсів, аналіз потреб споживачів, аналіз конкурентоспроможності та формування ринкової пропозиції (рис. 5).

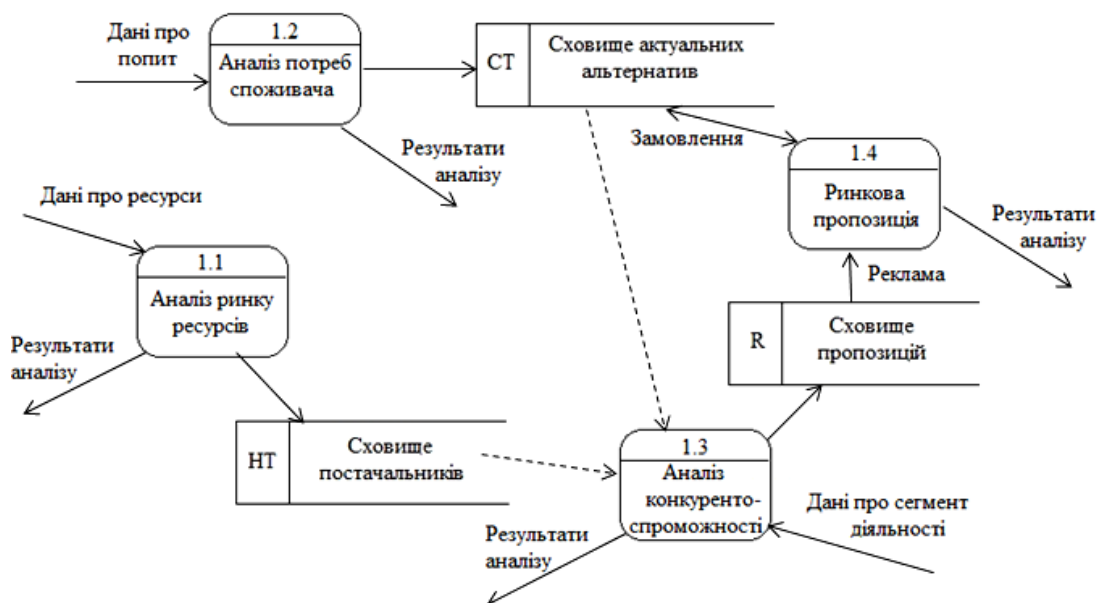


Рис. 5. Діаграма потоків даних аналізу та обліку фінансово-господарських процесів

Кожен з цих процесів передбачає виконання таких основних кроків (рис. 6):

- отримання даних від об'єкта;
- перевірка даних на несуперечність;
- очищення даних від повторень;
- завантаження даних;
- власне сам аналіз.

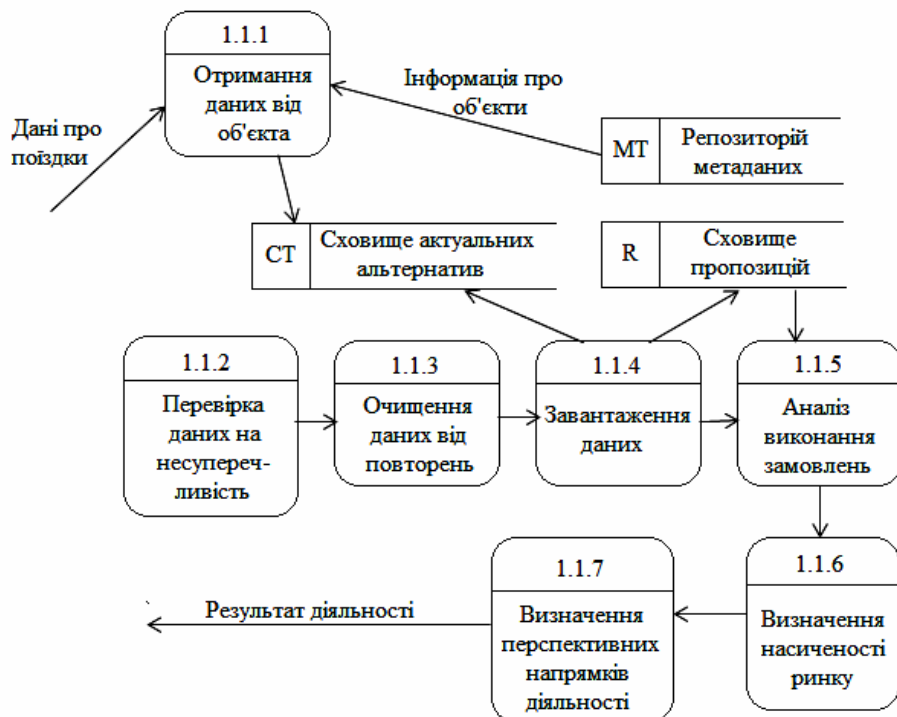


Рис. 6. Діаграма потоків даних аналізу потреб споживачів

У діаграмі використовуються такі сховища даних: сховище постачальників, сховище актуальних альтернатив та сховище пропозицій, а також сховище даних, яке використовується для опису об'єктів економічної сфери регіону. Це є репозиторій метаданих, який дозволяє ще й здійснювати автоматичне завантаження та узгодження даних з різних джерел. Побудуємо основні об'єкти та їх характеристики, що використовуються для подання економічної інформації.

Вхідною інформацією для комплексної інформаційної системи є:

- результати економічного аналізу окремого ринкового сегменту;
- дані соціальної сфери;
- інформація стосовно транспортної інфраструктури;
- історичні та культурні особливості даного регіону;
- результати аналізу суспільно-політичної складової господарської діяльності.

Розв'язання задачі

Отже, ставимо таку задачу: побудувати сховище даних та розробити засоби інтеграції та аналізу розрізної вхідної інформації економічної сфери.

В інформаційному сховищі використовується методологія метаданих, оскільки завдяки цьому блоки бази перетворюються в єдиний працездатний об'єкт. Всі формати подання метаданих жорстко пов'язані між собою. Для опису даних та методів їх отримання з різних джерел використовують таблиці метаданих. Метадані систем сховищ даних іноді розділяють на два типи:

1. Службові метадані, що використовуються для функцій витягання, перетворення і завантаження, для перенесення інформації з транзакційних систем в сховищі.

2. Інтерфейсні метадані, що використовуються для опису екранів і створення звітів.

Оскільки сховищем даних користуватимуться різні групи користувачів, до сховища доцільно застосовувати шестивимірну класифікаційну схему Захмана.

При цьому використовується така формалізація:

- сутність або зміст сховища даних;
- люди, які використовують сховище даних;
- місце розташування даних, важливі з погляду сховища даних;
- моменти завантаження даних і обчислення підсумкових таблиць;
- рушійні сили створення і розвитку сховища даних;
- дії, які виконуються з даними;
- повчальні метадані (як новий чинник, що використовуватиметься для підтримки розвитку моделі сховища даних).

Оптимальний вихід для економічної сфери регіону та його ефективного розвитку – необхідність враховувати кожен деталь у реалізації запланованої задачі. Очевидно, що необхідність опрацювання невизначеності зустрічається майже для кожного об'єкта.

Інтеграція даних – це об'єднання даних, які спочатку вводяться в різні системи. Самі ці системи можуть розташовуватися в одній локальній мережі, але мати різні платформи і внутрішню архітектуру.

Розглянемо деякі проблеми реалізації сховища даних, які приводять до виникнення задачі інтеграції даних. Серед них можна виділити:

- неоднорідність програмного середовища;
- розподілений характер організації;
- підвищені вимоги до безпеки даних;
- необхідність наявності багаторівневих довідників метаданих;
- потреба в ефективному зберіганні і обробці дуже великих обсягів інформації.

Приклад метаданих суб'єктів наведено у таблиці.

ID	Таблиця БД	Таблиця СД	Атрибути	Джерело
1	Виробничий ресурс	Виробничий ресурс	ID, Код територіальної одиниці, код бізнес-одиниці	C:\agent1.mdb
2	Ресурс	Виробничий ресурс	ID, Код територіальної одиниці, код бізнес-одиниці	C:\agent2.mdb
3	Виконавець	Виробничий ресурс	ID, Код територіальної одиниці, код бізнес-одиниці	C:\agent.dbf

Технологія в інтеграції даних є найбільш корисною в тих випадках, коли необхідно створити сховище даних, що містить добре документовані і надійні дані для історичного аналізу, наприклад, для аналізу часових рядів або багатовимірних запитів. Її також використовують для інтеграції ключових довідкових даних. Технологія ETL незамінна для таких завдань, як видалення даних, що дублюються, здійснення процесів перевірки якості даних та інше. Ці інструменти також використовують для створення окремих вітрин даних, що обслуговують конкретний відділ або бізнес-процес, або призначених для яких-небудь довгострокових цілей. Інструменти ETL дають користувачу можливість запустити процеси, що повторюються, для більшої злагодженості дій і можливості їх багатократного використання. Такі процеси включають створення точних технічних метаданих, що підтримують загальну цілісність середовища.

Стандартизація послідовності операцій при завантаженні сховищ даних, з огляду на важливість і вартість багатьох рішень ETL, дозволить уникнути повторення помилок, допущених у попередніх розробках. Крім цього, досвід розробки ETL допоміг виявити загальні частини ETL-процесів при завантаженні різнорідних джерел, що дозволяє говорити про однаковість підходу до розробки ETL для джерел даних довільного походження.

Загалом застосування ETL витягують інформацію з початкової бази даних, перетворюють її у формат, що підтримується базою даних призначення, а потім завантажують в неї перетворену інформацію. Для того, щоб ініціювати процес ETL, застосовуються програми витягання даних для читання записів в початковій базі даних і для підготовки інформації, що зберігається в цих записах, до процесу перетворення.

Здебільшого існуючий інструментарій ETL здатний задовольнити велику частину вимог до перенесення даних. Проте він неефективний для потокових даних, а також даних, що надсилаються з певною періодичністю, оскільки між періодами оновлення існує ймовірність появи неактуальності, а, може, і суперечливості інформації. Тому одним з методів зменшення невизначеності інформації, яка виникає внаслідок ETL, є вказання часу вивантаження, перетворення та завантаження даних і врахування цих часових термінів при формуванні звіту. Для кінцевого звіту включатиметься інформація з найбільшим часом актуальності.

Вирішуючи питання про створення федеративного сховища в господарській сфері, потрібно врахувати такі ключові фактори (рис. 7):

1) внутрішня організація компанії та вимоги бізнесу. Прикладом цього може бути ситуація, при якій користувачі фірми на своїх місцях звикли контролювати й здійснювати зміни на локальному рівні, тоді федералізація є підходящим рішенням.

2) ступінь розбіжності моделей бізнес-даних. Це один з найважливіших критеріїв для вибору рішення про використання федералізації для інтеграції інформації. Наприклад, якщо двом частинам суб'єкта господарювання потрібні різні класифікації послуг, тоді обидві ці класифікації можуть бути включені в одну модель. Глобальні структурні розбіжності – це ще одна перевага на користь федералізації. Наприклад, якщо один підрозділ бізнес-одиниці працює прямо зі споживачами, а інший – тільки з посередниками, то контекст і структура транзакцій будуть дуже сильно відрізнятися, що ускладнить їхнє сполучення в одній моделі даних.

3) ступінь розбіжності вмісту даних. Часто один бізнес-підрозділ не використовує дані іншого підрозділу при прийнятті рішень щодо свого бізнесу. Але при цьому можуть існувати деякі загальні довідкові дані, наприклад, загальна корпоративна класифікація клієнтів.

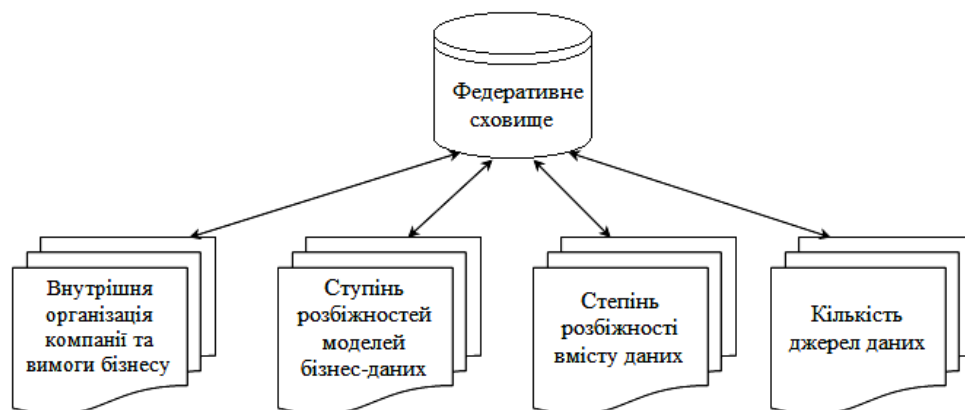


Рис. 7. Фактори існування федеративного сховища

Кількість джерел даних. Якщо існує одне основне джерело даних, то їхня фільтрація для завантаження в окремі сховища збільшить обсяг роботи. Але якщо дані розміщені в локальних джерелах, то федералізація виявляється більше підходящим рішенням.

Існує три основні архітектурні моделі федералізації:

- ✓ за географічною ознакою;
- ✓ за областями бізнесу;
- ✓ функціональна федералізація.

У першому випадку члени федералізації виражають вимоги бізнесу кожного географічного регіону. Федералізація за областями бізнесу означає, що члени федерації представляють окремі види бізнесу. Функціональний підхід до федералізації будується на основі різних бізнесів-функцій, таких як продаж, фінанси, маркетинг і ланцюги послуг.

Висновки

Описано предметну область та обгрунтовано необхідність інтеграції даних для прийняття рішень в економічній сфері, загалом. Проаналізовано предметну область на наявність факторів привабливості окремих сегментів економічної діяльності регіону, виділено необхідні умови та задачі, які розв'язуватиме пакет програм для комплексного інформаційного забезпечення, формування аналітичних та прогнозних даних та підтримки прийняття рішень щодо розвитку галузі, побудовано концептуальну модель предметної області.

1. Уайт К. Интеграция данных: использование технологий ETL, EAI и EII для создания интегрированной корпорации / К. Уайт. – 2005. 2. Особенности построения хранилищ данных. – Режим доступа: <http://citforum.uar.net/seminars/cis99/sch.shtml>. 3. Огляд технологій інтеграції інформаційних систем, 2006. – Режим доступа: <http://www.microsoft.com/Ukraine/Government/Analytics/IntegrationTechnologies/Overview.msp>. 4. Дрюэк К. Хранилища данных: сходство и различия подходов Билла Инмона и Ральфа Кимболла. – Режим доступа: <http://www.tdan.com/i021hy01.htm>. 5. Franklin M. From Databases to Dataspace: A New Abstraction for Information Management // M. Franklin, A. Halevy and D. Maier. ACM SIGMOD Record 34 – 2006 – No.4. 6. Інтеграція даних у сховищах даних. – Режим доступа: <http://it.ridne.net/node/311>.