

УДК 004.5

О. С. Горбачик

Інститут проблем реєстрації інформації НАН України
вул. М. Шпака, 2, 03113 Київ, Україна

Організація корпоративної аналітичної діяльності та сучасні технології її підтримки

Визначено етапи аналітичної роботи, розглянуто комп'ютерні технології їхньої підтримки. Запропоновано узагальнену схему організації інформаційно-аналітичної діяльності в корпоративній інформаційній системі.

Ключові слова: інформаційно-аналітична діяльність, корпоративні інформаційні системи, VI-технології.

Авторитет, імідж, конкурентоздатність, живучість суб'єктів ділової активності в сучасних умовах залежить від якості, ефективності та своєчасності прийняття управлінських рішень. Керівники різних рівнів усвідомили, що отримати переваги на ринку в умовах конкурентної боротьби можливо за наявності точної та актуальної інформації щодо розвитку ситуації в їхній професійній сфері, при консолідації та об'єднанні інтелектуальних зусиль і можливостей фахівців різних предметних областей та сфер діяльності, які володіють різними засобами обробки та аналізу інформації. Умови ділового середовища сформували потребу в аналітиці у керівного складу, і, відповідно, визначили вимоги до підготовки фахівців середньої ланки для сприйняття аналітики. У структурах багатьох відомств, корпорацій, окремих організацій, фірм сформовані спеціалізовані служби або підрозділи аналітичної роботи. Аналітичні служби створені в органах державної влади, у бізнесі, при політичних партіях, різних ЗМІ, як самостійні комерційні фірми; відбувається становлення інституту незалежних аналітиків [1, 2].

Завданням аналітичних служб у будь-якій сфері є отримання інформації з величезної кількості різноманітних даних (кількісних, якісних, текстових, відео, аудіо тощо), її систематизація, аналіз для вилучення суттєвих відомостей, знань із метою напрацювання обґрунтованих рекомендацій для прийняття управлінських рішень, а також для прогнозування, виявлення тенденцій (іноді неявних, прихованих) розвитку ситуації у відповідній предметній сфері [3].

Сучасна інформаційно-аналітична діяльність є складною комплексною діяльністю, що спирається на природний інтелект та комп'ютерні технології [1]. Розви-

ток інформаційних технологій дозволив продуктивно опрацювати сьогоднішній практично безперервний потік різнорідних даних. Головна проблема — вилучення дійсно цінних відомостей, знань з інформаційного потоку. Конкретність, зрозумілість, обґрунтованість результатів аналітичної роботи, їхня практична цінність суттєво залежать від кваліфікації аналітика.

Інформаційно-аналітична діяльність має проводитись в режимі життєдіяльності конкретної сфери (політики, економіки, бізнесу тощо) і у відповідності до темпу управлінських реакцій на динаміку подій у цій сфері [2]. Така діяльність поєднує наукові методи (по вилученню чи побудові нового знання) та методи управлінської діяльності (щодо розробки варіантів сценаріїв дій, методик побудови індикаторів загроз, що виникають, схем мінімізації існуючих ризиків тощо). У сфері інформаційної аналітики здійснюється інтеграція формальних наукових знань із неформальним досвідом і знаннями посадових осіб, які приймають рішення.

Розрізняють наступні етапи інформаційно-аналітичної роботи:

- визначення цілі та первинне планування роботи;
- початкова аналітична робота: збір даних та їхнє накопичення;
- обробка даних, вилучення інформації по конкретній проблемі (чи проблемній ситуації);
- аналіз інформації (структурування, класифікація, побудова відношень, взаємозалежностей, пошук регулярностей, генерація гіпотез, їхнє тестування та верифікація, побудова моделі проблемної ситуації тощо);
- синтез знань для реалізації визначених цілей (розробка сценаріїв дій, побудова конкретних методик, схем розв'язання проблемної ситуації, визначення методів та засобів запобігання небажаному розвитку подій та ін.).

Найважливіші вимоги до результату інформаційно-аналітичної роботи — це конкретність, зрозумілість і практична корисність для напрацювання стратегічного чи поточного керівного рішення, схеми або плану дій в існуючій ситуації. Після чи під час реалізації обраного рішення, плану дій доцільне повторення окремих чи всіх етапів інформаційно-аналітичної роботи з метою контролю за виконанням, оцінки, або корегування плану реалізації в залежності від розвитку ситуації.

Об'єктом інформаційно-аналітичної діяльності досить часто стають складні взаємопов'язані проблеми, кожна з яких може характеризуватися великою кількістю параметрів, безперервністю змін, важкістю передбачення динаміки розвитку. Особливістю задач, що виникають при розв'язанні таких проблем, є відсутність чіткого алгоритму, який дозволяв би завжди досягти необхідного рішення. Для цих задач також характерні неповнота, неточність чи суперечність знань про предметну область, велика розмірність простору рішень, що виключає можливість ефективного пошуку рішень звичайним перебором, відсутність чітких критеріїв вибору оптимального рішення, а самі задачі, як правило, важко формалізуються. Основною й часто єдиною інформацією для формування плану рішення (розв'язання проблеми) є саме постановка аналітичної задачі, але ціль і відповідно постановка задачі можуть змінюватися в процесі виконання аналітичної роботи, що вимагає корегування плану рішення та додаткового «здобування» інформації й своєчасного опрацювання її в режимі життєдіяльності конкретної сфери. Зазначе-

ні особливості формують вимоги до засобів та технологій підтримки аналітичної діяльності, формування сценаріїв проведення інформаційно-аналітичної роботи.

Сьогодні процеси збору, збереження, обробки, перетворення, передачі й оно-влення інформації реалізуються з використанням сучасної комп'ютерної техніки та засобів телекомунікації. Впровадження корпоративних інформаційних систем (КІС) та відповідних інформаційно-аналітичних технологій дозволяє в необхідний термін відображати цілісний і максимально об'єктивний стан дій у корпорації, відомстві, фірмі, чи організації, а також підтримувати відповідні організаційно-технологічні моделі управління [4]. Сучасним КІС властива територіальна розподіленість компонентів при інтегрованості інформаційних ресурсів, модульність побудови, масштабованість, інтегрованість з програмними продуктами різних розробників, а також із web-технологіями, відкритість, багатоплатформеність, дво- чи трирівнева архітектура (клієнт-сервер) і наявність інформаційно-аналітичних засобів. Корпоративність вимагає забезпечення гнучкої настройки на специфіку й сферу діяльності в конкретній предметній області, підтримку розподілених баз даних для інформаційної взаємодії підрозділів і територіально віддалених філій, надання користувачам необхідних засобів для підтримки їхніх функціональних обов'язків й інструментальних засобів для подальшого розвитку можливостей КІС [5]. Інформаційно-аналітичні технології КІС дозволяють покращити управління, завдяки тому що [4, 6]:

— на оперативному рівні прискорюється рух інформаційних потоків, забезпечується необхідний темп інформаційної взаємодії учасників ділових процесів, підвищується синхронізація операцій, що виконуються одночасно;

— на тактичному рівні підвищується якість і обґрунтованість управлінських рішень, спрямованих на адаптацію функцій управління до змін зовнішнього середовища;

— на стратегічному рівні своєчасно напрацьовуються рішення щодо розробки нових і модернізації існуючих бізнес-процесів.

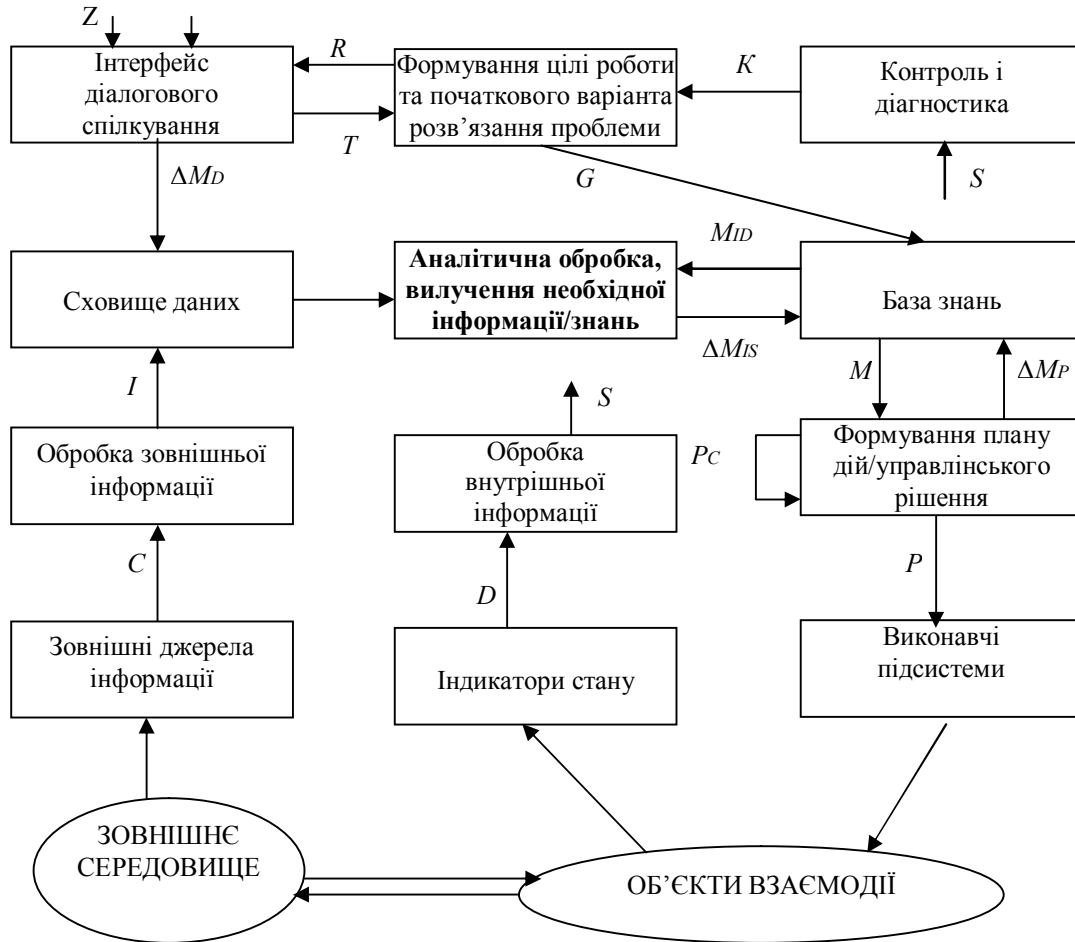
Якість і достовірність результатів інформаційно-аналітичної діяльності залежить від даних, з яких вилучається інформація та формуються знання, інструментарію доступу та обробки даних, обраних фахівцями моделей аналізу й синтезу, сценарію проведення аналітичної роботи.

Узагальнену схему організації інформаційно-аналітичної діяльності в КІС наведено на рисунку.

Визначена ціль аналітичної роботи Z у діалоговому режимі перетворюється (інтерпретується) у відповідне формалізоване завдання T . Виходячи з формалізованого завдання T та контрольної-діагностичної інформації K про стан дій у корпорації та наявні ресурси (як матеріальні, так і інформаційні) формується початковий варіант G рішення проблеми (розв'язання проблемної ситуації) або формуються вимоги R щодо корекції чи уточнення завдання T і відповідно цілі Z .

База знань, за допомогою якої вирішується проблема, має містити опис прецедентів (рішень певного набору задач управління), опис (модель) середовища (предметної сфери) M , який має змінитися після реалізації управлінського рішення. Поточний стан моделі середовища M_D визначається з урахуванням корегуючої інформації ΔM_D й інтегрованої зовнішньої інформації I . Додаткова інформація

щодо проблемної ситуації ΔM_{IS} напрацьовується при обробці даних зі сховища даних і після відповідної інтерпретації передається у базу знань.



Узагальнена схема організації інформаційно-аналітичної діяльності в КІС

Обробка початкового варіанта G рішення проблеми та узагальнених у моделі M знань про середовище та проблему виконується засобами «Формування плану дій/управлінського рішення». Ця обробка є ітераційним процесом, на кожному кроці якого обирається певний план дій P_c , який призводить до змін моделі середовища ΔM_p . Якщо зміни ΔM_p корегують модель M таким чином, що виконується завдання T (досягається ціль Z) за критерієм мінімальної неузгодженості, то відповідний план дій (управління) P передається до виконавчих систем, які інтерпретують його для об'єктів взаємодії. Через об'єкти взаємодії відбувається керований вплив на зовнішнє середовище для досягнення цілі Z .

Підсистемами обробки зовнішньої C та внутрішньої D інформації виконується підготовка даних для аналітичної роботи: структурування, узгодження, узагальнення, необхідне впорядкування з подальшим збереженням у сховищі даних.

Спеціальним чином організована інтегрована зовнішня *I* та внутрішня *S* інформація зі сховища даних використовується для проведення аналітичної роботи.

Узагальнена схема організації інформаційно-аналітичної діяльності розрахована на вирішення різноманітних аналітичних задач, але реалізація підсистем, вибір інформаційних технологій залежать від проблемної орієнтації КІС.

Сучасні інформаційно-аналітичні технології базуються на результатах досліджень у сфері теорії інформації та штучного інтелекту й належать до так званих ВІ (Business Intelligence) технологій і засобів. Поняття ВІ-технологій об'єднує різні засоби та технології аналізу й обробки даних, серед них розрізняють [7]: OLAP-продукти, інструменти здобування даних, засоби побудови сховищ та вітрин даних, управлінські інформаційні системи й додатки, інструменти кінцевого користувача для виконання запитів і побудови звітів. Ці технології забезпечують навігацію в постійно зростаючому інформаційному середовищі, прискорюють пошук необхідної інформації з відповідної предметної області, її узагальнення, установлення причинно-наслідкових зв'язків впливу одних факторів на інші, дозволяють спрогнозувати розвиток ситуації в майбутньому на основі отриманих результатів аналізу й наявного досвіду розвитку ситуації з аналогічними параметрами моделі. Комп'ютерні технології забезпечують своєчасне надання користувачу (керівнику, аналітику, експерту, консультанту) усієї необхідної інформації для прийняття рішення, але яка інформація буде замовлена, і яке рішення буде прийнято на її основі, залежить виключно від конкретного користувача, від чіткого розуміння ним цілі аналітичної роботи, організації процесів збору релевантних даних, вибору методів аналізу, вибору програмних засобів, якості виконання аналізу та прийняття рішення про використання результатів.

У КІС, як правило, накопичуються великі масиви даних у форматах офісних застосувань та систем документообігу. Корпоративне інформаційне середовище складають різноманітні за форматами й структурами корпоративні бази даних, бази даних із доступом через внутрішні або зовнішні корпоративні сайти, чи портали. Технології пошуку, які підтримують аналітичну діяльність у КІС, як правило, базуються на методах та інструментах подібних тим, що використовуються в Internet. Традиційний критерій для вибору пошукового інструментарію — співвідношення ефективності до витрат.

Пошукові системи універсальні щодо застосування, інтуїтивно зрозумілі користувачам, проблемно-незалежні, але вибагливі до професійного рівня фахівця-аналітика. Саме аналітик формує пошуковий запит (множину ключових слів із зафіксованими відносинами між ними) і робить висновок щодо важливості, точності та значущості отриманої інформації. Важливість інформації визначається через наявність зв'язку з вирішуваною проблемою та необхідність інформації для досягнення поставленої цілі аналітичної роботи. Точність — через ступінь узгодженості з інформацією з інших джерел, рівень довіри до джерела інформації, наявність підтверджень інформації з інших джерел. Значущість інформації оцінюється її повнотою, своєчасністю та достатністю для прийняття рішення.

Сьогодні базою для підтримки аналітичної діяльності в КІС стали такі системи пошуку, як mпоGoSearch (універсальна пошукова система призначена для Internet чи intranet-серверів; опрацьовує документи у форматах html, txt, doc, pdf; у запитах сприймаються різні формати слів і логічних операторів), «Ищейка» (ро-

бота з документами у форматах txt, doc, rtf, html; підтримується морфологічний пошук; автоматично розпізнаються основні типи кодування тексту — ASCII, ANSI, Unicode), «Следопыт» (пошук на окремому web-сайті або intranet-сервері; працює з форматами doc, rtf, html, xls, pdf, zip, pst, а також з папками Microsoft Outlook; реалізується морфологічний пошук), CROS 4.01 (накопичення та обробка документів у різних форматах; автоматичне розпізнавання форматів Microsoft Word версій 6.0, 7.0, 97, 2000, форматів rtf та html; навігація по знайдених документах; підтримка захисту від несанкціонованого доступу), програмно-апаратний комплекс Google Search Application (пошуковий механізм забезпечує роботу більш як з двомастами типами файлів; підтримується робота з захищеною і секретною інформацією), інформаційно-пошукова система RetrievalWare (див. www.convera.com) та інші [8, 9].

Відаючи перевагу тій чи іншій інформаційно-пошуковій системі, або визнаючи вимоги щодо розробки пошукових засобів для КІС, необхідно враховувати ряд факторів, від яких залежить оцінка результатів пошуку, а саме, повноти, точності, актуальності та швидкості отримання результату [10]. Так, повнота пошуку залежить від інформаційно-пошукового простору, що визначається сервісами, у рамках яких функціонує пошукова система, від особливостей власне інформаційно-пошукової системи, а також якості побудови запиту. Точність залежить від функціональних особливостей інформаційно-пошукової системи і якості запиту. Актуальність — від особливостей пошукової системи й частоти оновлення інформації в предметній сфері. Швидкість отримання результату — від технічних характеристик каналів зв'язку й завантаження цих каналів та сервісів пошукових служб, особливостей алгоритмів пошуку, які реалізовані в інформаційно-пошуковій системі, а також від побудови запиту користувача.

Згідно професійних обов'язків більшість користувачів КІС повинні не тільки виконувати пошук необхідної інформації, але й вирішувати задачі групування тематично близьких документів, реферування, інтеграції даних із різних джерел, готувати різноманітні звіти, як екранні, так і друквані. Впровадження інструментів кінцевого користувача для виконання запитів і побудови звітів (систем класу Query & Reporting) дозволяє формулювати пошукові запити в термінах відповідної професійної сфери, а візуалізація результатів пошуку може бути наведена в традиційному для користувача вигляді — таблиць, графіків, діаграм, різних спеціалізованих інтерфейсів.

Для пошуку інформації та виявлення знань у великих масивах неструктурованої або слабо структурованої інформації використовуються технології та інструменти здобування знань (Knowledge Mining, Text Mining, Data Mining). Ці технології базуються на математичних засадах (методах фільтрації, асоціативних правилах, деревах рішень, генетичних алгоритмах, нейронних мережах) і являють універсальний інструмент для аналітики. Програмні засоби підтримки технологій здобування знань дозволяють створювати корпоративні портали знань — інтегровані інформаційні депозитарії, які доступні для оперативного узагальнення та аналізу.

Задачі прогнозування, виявлення тенденцій, які надзвичайно важливі для напруження стратегічних рішень, потребують відповідної впорядкованості даних (узагальнення та узгодженості, структурованості та доступності, хронологічної

впорядкованості). Головна ідея технології оперативної аналітичної обробки (On-Line Analytical Processing, OLAP) полягає в створенні багатовимірних таблиць (кубів), що будуються на основі даних, які зберігаються як у реляційних, так і багатовимірних базах даних, і орієнтовані на продуктивну оперативну обробку.

Для більш складної, комплексної аналітичної роботи, що вимагає єдиного логічного подання корпоративних даних (єдиної моделі даних), просувається й впроваджується сьогодні технологія сховищ даних. До складу реляційних і багатовимірних СУБД таких виробників, як Microsoft, Oracle, IBM, Sybase, включені засоби проектування сховищ даних — середовища для ефективного здобування знань. Сховище даних являє собою сукупність даних із різних джерел, наявних у корпорації, відомстві, фірмі чи організації, у якому забезпечується швидкий доступ до даних і усунення їхньої неузгодженості. У КІС, що розробляються на засадах концепції сховища даних, підтримується єдиний логічний погляд на дані певної предметної області, забезпечується утворення єдиного інтегрованого джерела даних і формування корпоративного інформаційного середовища, у якому реалізуються єдиний інтерфейс, уніфіковані структури зберігання, загальні довідники тощо, а також будуються всі довідково-аналітичні застосування відповідної прикладної сфери для підтримки інформаційно-аналітичної діяльності. Дані для опрацювання фахівцями окремого підрозділу утворюють, у термінах концепції сховищ даних, вітрини даних. Вітрини даних можуть бути реалізовані як бази даних клієнт-сервер, електронні таблиці, бази даних на робочому місці, OLAP-куби тощо. Вітрини даних можуть утворюватися на основі інформації з корпоративного сховища даних, або безпосередньо із транзакційних систем (якщо сховище даних у корпорації не реалізується). Інформація у вітринах — це інформація в найбільш доступному для фахівців підрозділу вигляді, у термінах відповідної предметної сфери, підготовлена для проведення аналітичної роботи.

Засоби підтримки інформаційно-аналітичної діяльності, традиційної для корпорації, відомства, фірми чи організації (наприклад, фінансової звітності та управління, інвестиційного аналізу, маркетинг-аналізу, управління проектами тощо), що впроваджуються в КІС, належать до так званих управлінських інформаційних систем і застосувань, які реалізуються як на засадах базових інформаційних технологій, так і на сучасному аналітичному інструментарії.

Можливості для розробки і реалізації унікальних спеціалізованих корпоративних рішень щодо організації ефективної комплексної аналітичної роботи надають пропозиції від провідних компаній розробників ВІ-технологій таких, як Humingbird Inc., Documentum (Nasdaq: DCTM), Cognitive Technologies Ltd., Convera Technologies International Ltd., NeurOk та інших [9].

Кожний проект із організації й технологічної підтримки інформаційно-аналітичної діяльності в КІС має суто індивідуальний характер, як правило, потребує розробки й впровадження спеціальних програмних засобів, що реалізують відповідні методи, методики, технології вирішення важкоформалізуємих задач, притаманних діяльності корпорації, відомству, фірмі чи організації, і надають можливість опрацювання різних евристичних стратегій, тактик, методів конкретної професійної діяльності; можливість використання методів виявлення та вирішення протиріч різної природи та походження, методів систематизації функцій і властивостей технологій, інформації тощо. Розробка таких засобів складна, оскі-

льки вимагає взаємодії спеціалістів, які мають досвід впровадження й використання інформаційних технологій і спеціалістів корпорації, які спроможні сформулювати задачі аналітичного характеру, мають досвід їхнього вирішення, і можуть інтегрувати наявні знання фахівців та унікальний досвід функціонування корпорації. Цінність засобів підтримки аналітичної діяльності в КІС, якщо вони належним чином створюються й керуються, із часом зростає, оскільки накопичуються знання, які складають найбільш важливий досвід корпорації, її інтелектуальний капітал. Знання, які інтегруються в корпоративні інформаційно-аналітичні ресурси, є знаннями-зразками, збереженням унікального управлінського досвіду фахівців, технологій дій у залежності від розвитку ситуації у відповідній сфері діяльності.

1. Сладнева Н.А. Информационная аналитика — эзотерическое искусство или современная профессия? // www.bpi.consultinfo.ru/publication/sladneva1.php
2. Сладнева Н.А. Информационно-аналитическая деятельность: проблемы и перспективы // Информационно-аналитический журнал. — 2001. — № 6.
3. Дюк В., Самойленко А. Data Mining. — СПб: Питер, 2001. — 368 с.
4. Мишенин А.И. Теория экономических информационных систем. — М.: Финансы и статистика, 1999. — 240 с.
5. Горбачик Е.С. Корпоративные информационные системы и обеспечение их безопасности: Сб. науч. тр. «Информационные технологии и безопасность». — К.: ИПРИ НАНУ, 2004. — В. 7. — С. 3–6.
6. Горбачик О.С. Корпоративні інформаційні системи: підтримка аналітичної діяльності: Сб. науч. тр. «Информационные технологии и безопасность». — К.: ИПРИ НАНУ, 2005. — В. 8. — С. 9–11.
7. Чаусов В., Амириди Ю. Классификация аналитических систем. Три года спустя // Банки и технологии, — 2002. — № 6.
8. Ландэ Д.В. Поиск знаний в Internet. Профессиональная работа: Пер. с англ. — М.: Издательский дом «Вильямс», 2005. — 272с.
9. Печенкин И.А. Информационные технологии на службе разведки. Обзор современных программных средств обеспечения принятия управленческих решений // Защита информации. Конфидент. — 2004. — № 4. — С. 2–15.
10. Тюрин Ю.Н., Макаров А.А. Статистический анализ данных на компьютере. — М.: Инфра-М, 1998. — 528 с.
11. Нежданов И.Ю. Анализ неструктурированной информации // Защита информации. Конфидент. — 2004. — № 4. — С. 53–55.

Надійшла до редакції 15.08.2006