

УДК 681.3

ІНТЕЛЕКТУАЛІЗАЦІЯ ІНФОРМАЦІЙНИХ КОМП'ЮТЕРНИХ СИСТЕМ В ОХОРОНІ ЗДОРОВ'Я УКРАЇНИ

Ю.А.Міхненко, В.І. В'юн

Інститут проблем математичних машин і систем НАН України

e-mail: myna1941@mail.ru

Спільною проблемою більшості інформаційних комп'ютерних систем (далі ІС) в охороні здоров'я України є те, що дані, які в них накопичуються, (а це десятки, інколи, сотні гігабайт) недостатньо використовуються при прийнятті рішень.

Пояснити цю “дивну” ситуацію можна лише тим, що для осіб, зайнятих підготовкою, обґрунтуванням та прийняттям рішень в системі охорони здоров'я, **ці дані не є корисними**, або тим, що **в існуючих ІС недостатньо розвинута “культура даних”**.

На нашу думку, покращити цю негативну ситуацію можна застосовуючи в інформаційно-аналітичних блоках ІС охорони здоров'я сучасні засоби інтелектуалізації.

В якості інструментарію інтелектуалізації ІС пропонується використовувати [1,2]:

- Сховища даних (Сх.,Д.), певним чином інтегровані в Єдине інформаційне середовище системи МОЗ України;
- Функціонально-орієнтовані сховища – “Вітрини даних”;
- Традиційні статистичні та аналітичні методи аналізу даних;
- Онлайнові аналітичні (OLAP) та інтелектуального аналізу даних Date mining (ДМ) технології;
- Web та ГІС технології;
- Сучасні засоби інтелектуального пошуку та перегляду даних та ін.

Для розгортання зазначених засобів інтелектуалізації потрібно мати відповідну програмно-технічну платформу, яка була б здатна забезпечити інтерактивний доступ широкому колу користувачів до даних, що генеруються, імпортуються або зберігаються в корпоративних ІС МОЗ України [3], тобто забезпечувати повноцінну підтримку прийняття рішень на різних ієрархічних рівнях управління галуззю.

Єдину корпоративну інформаційну систему доцільно створювати на базі вже існуючих в галузі ІС шляхом їх реінжинірингу.

Для цього потрібно провести інвентаризацію існуючих в галузі ІС та систематизувати їх по наступним сферам:

- **“Сфера деталізації даних”**. Сюди віднести інформаційні системи, створені в охороні здоров'я, які за своїм призначенням націлені на пошук інформації. В більшості з цих систем використовуються реляційні СУБД. Язык маніпулювання даними переважно SQL. В системах, за звичай, реалізований стандартний інтерфейс взаємодії з користувачем;

- **“Сфера агрегованих показників”**. До цієї сфери слід віднести інформаційні системи / переважно аналітичні/, які надають можливість користувачу комплексно розглянути інформацію, що зберігається в різних за призначенням Сховищах, у тому числі і розподілених по різних ієрархічним рівням управління. За правило, в таких системах інформація, яку отримує користувач, вже узагальнена та агрегована у вигляді показників, або представлена у вигляді гіперкуба, який дозволяє реалізувати технологію оперативної аналітичної обробки даних (OLAP). В сховищах систем цього типу можна орієнтуватись і отримувати необхідну інформацію за допомогою спеціалізованої

багатомірної СУБД, або працювати з заздалегідь агрегованими показниками (переважно статистичними або аналітичними), залишаючись в рамках реляційних технологій.

- **«Сфера статистичних закономірностей»:** Сюди слід віднести інформаційні системи, в яких виконується інтелектуальна обробка даних з використанням засобів та методів ДМ. Головною задачею систем цього типу є - пошук функціональних і логічних закономірностей в даних, що зберігаються в БД. В системах цієї групи на основі отриманих знань будуються моделі і алгоритми, які пояснюють знайдені закономірності, або аномалії, а також прогнозуються процеси в самій системі охорони здоров'я, які ймовірно можуть мати зв'язок з виявленими закономірностями чи аномаліями.

Систематизація ІС по зазначеним сферам дасть можливість обґрунтовано вибрати засоби інтелектуального аналізу даних та технології їх застосування.

В залежності від вибраного засобу інтелектуалізації та технології його застосування, визначатимуться рівні знань та типи закономірностей, які застосований метод дозволяє виявити. Наприклад, прості SQL запити дозволяють провести поверховий аналіз даних. За допомогою OLAP технології можна провести не глибоку оперативну обробку даних. Виявити “скриті” закономірності або аномалії можна за допомогою технології ДМ (“розкопки даних”). Методи та засоби ДМ можуть зіграти провідну роль при встановленні закономірностей, які з тих або інших причин не можуть бути точно описані звичайними статистичними методами, або за допомогою класичних математичних моделей. Особливо це стосується тих предметних областей охорони здоров'я, в яких похідні дані, за правило, неоднорідні, гетерогенні, або нестационарні, до того ж мають високу розмірність.

В основу технології роботи саме з такими даними покладена концепція шаблонів (паттернов), які відображають фрагменти багатоаспектних взаємовідношень в масивах даних. Ці шаблони представляють собою закономірності, які властиві певним підвибіркам даних, які можна компактно представити в зрозумілій користувачу формі. Пошук потрібних шаблонів проводиться методами, не обмеженими рамками апріорно сформульованих гіпотез щодо структури вибірки та уявлень щодо показників, які розглядаються в процесі інтелектуального аналізу даних.

В якості прикладу, на рис.1, показані в узагальненому вигляді структура та елементи платформи, яка дозволяє застосовувати методи ДМ. На цій платформі може функціонувати ІС МОЗ, орієнтована на підтримку прийняття рішень.

Ядром цієї ІС є аналітичний програмний комплекс /АПК/, який фактично є універсальним інструментальним засобом збору, консолідації, обробки та аналізу медико-статистичних і інших даних, що зберігаються в Сховищах системи.

В основу роботи АПК покладена технологія, що передбачає створення Єдиного корпоративного та інтегрованого Сховища даних.

Лише після попередньої обробки дані, що поступають в Єдине корпоративне Сховище, стають доступними користувачам для послідуочого маніпулювання.

Поряд з центральним Сховищем в системі функціонують і відносно невеликі функціонально орієнтовані сховища /так звані “Вітрини даних”/, які призначені для вирішення окремих аналітичних задач.

В ІС є сервер, який забезпечує прийом мета даних з усіх джерел інформації, що існують в інформаційному просторі МОЗ України. До центрального серверу підключені засоби вводу даних, які дозволяють вводити похідні дані в АПК з таблиць різних форматів (Dbase, Excel, Oracle, MS-SQL, InterBase і т.ін. При цьому дані, що вводяться, можуть бути перетворені в один з підтримуваних Єдиним Сховищем видів даних (одномірні,

багатомірні або неструктуровані дані). Адаптація системи до того, або іншого джерела інформації відбувається в залежності від задачі, що вирішується. Для цього в системі використовуються спеціальні шаблони завантаження похідних даних.

До серверу АПК підключено локальну мережу з периферійним обладнанням робочих місць кінцевих користувачів інформаційно-аналітичної системи МОЗ України.

Наведену вище структуру пропонується доповнити Сучасними інструментальними засобами: Data mining (“розкопки даних”) та OLAP (“оперативний аналіз даних”). На думку авторів це суттєво підвищить ефективність процесів підтримки прийняття управлінських рішень в Системі МОЗ України.

Резюме:

1. Сучасні технології інтелектуального аналізу даних тісно пов’язані між собою та з іншими технологіями статистичного та/або логічного (аналітичного) аналізу даних. Ці технології можуть органічно взаємодоповнювати одна-одну. По суті – це дві складові одного і того ж процесу підтримки прийняття рішень;

2. Автори на цьому етапі не ставили перед собою за мету дати читачам рекомендацій щодо вибору і застосування конкретних інструментальних засобів інтелектуального аналізу. Їх задача – лише загострити увагу читачів на ключових моментах сучасних інформаційних технологій, шляхах та засобах їх застосування в ІС МОЗ України.

3. Поза увагою в цій роботі лишились також питання:

- реінжинірингу існуючих в МОЗ України ІС;
- взаємозв’язку засобів інтелектуального аналізу з ГІС технологіями;
- механізми експорту результатів аналізу в стандартні форми статистичних звітів;
- обмеження, які слід враховувати при використанні нових технологій та інструментальних засобів їх реалізації і т. ін.

Ці питання стануть предметом розгляду в іншій роботі.

Наприкінці остання ремарка - багаторічний досвід роботи авторів з різними технологіями аналізу даних свідчить про те, що навіть найкращі інформаційні технології, сучасні програмно-технічні засоби та інструментарій інтелектуального аналізу даних нічого не варті без створення методологічної основи використання інформаційного субстрату, який генерується в ІС МОЗ України.

Без проведення валідазації цього субстрату /перевірки на істинність отриманих в ході інтелектуального аналізу результатів та закономірностей/ не можна використовувати його в системі Управління охороною здоров’я України.

Список літератури

1. Морозов А.О., В’юн В.І., Кузьменко Г.Є. Інтелектуалізація АСУ: проблеми, напрямки досліджень//Математичні машини і системи.-2001,-№1-2 –С.21-25.
2. В’юн В.І. Інтелектуалізація інформаційних систем - механізми та інструментарій інтерактивного ситуативного аналізу//Математичні машини і системи.-2004,-№3-С.125-131.
3. Голубчиков М.В., Кузьменко Г.Є., Міхненко Ю.А. Стан та перспективи розвитку галузевої системи підтримки прийняття управлінських рішень //ІІІ конференція Інформаційні технології в охороні здоров’я та практичній медицині /наукові праці /м.Київ, 2003р.

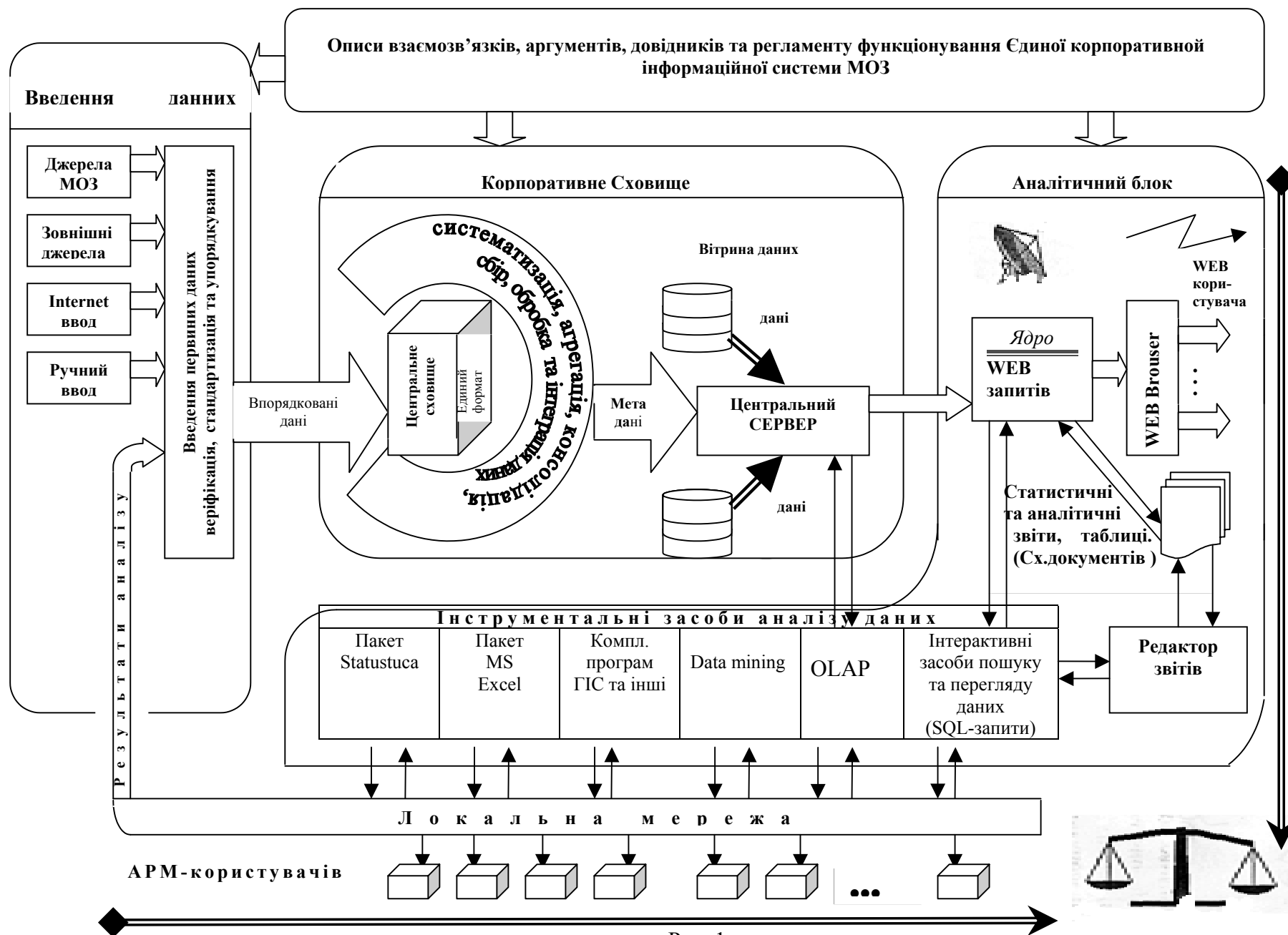


Рис. 1

Керівництво МОЗ України