## Життєвий цикл даних

Дані представляють собою спосіб представлення, збереження та елементарних операцій обробки інформації.

Дані - це основа інформації. Поняття "дані" - відносно новий термін. Зазвичай дані є вхідною інформацією для інформаційного процесу.

Дані можна збирати, обробляти, зберігати, змінювати форму їх представлення. Вони можуть створюватись, знищуватись, багаторазово використовуватись. Головною особливістю даних сьогодні є те, що їх стає надзвичайно багато. При масовому застосуванні комп'ютерів виникла гігантська кількість джерел даних. Для прикладу можна взяти обсяг даних у всесвітній мережі Інтернет, що збільшується щохвилини.

Ключовим поняттям маніпулювання даними є структура типу "файл", що представляє собою множину однотипних елементів (записів). Також файл займає певну ділянку на носії пам'яті і характеризується ім'ям, типом та іншими атрибутами. В свою чергу запис - це структура, яка складається з полів (мінімальної структури даних).

Основними етапами життєвого циклу даних є

* виникнення,
* збереження,
* застосування
* знищення. (Знищення, з точки зору життєвого циклу даних, не представляє інтересу, оскільки причиною видалення є втрата інформативності даних.

Фаза використання даних включає три етапи:

* пошук;
* обробку;
* аналіз.

Результатом використання даних є інформація.

Існує декілька методів збору, необхідних для аналізу даних:

1. Облікові системи. Як правило, в облікових системах є механізми побудови звітів і експорту даних, тому отримання потрібної інформації є відносно нескладною операцією.
2. Непрямі дані. Про чинники можна скласти оцінку і за непрямими ознаками.
3. Відкриті джерела. Велика кількість даних наявна у відкритих джерелах, таких як статистичні вибірки, звіти корпорацій, опубліковані результати маркетингових досліджень тощо.
4. Проведення незалежних маркетингових досліджень і аналогічних заходів щодо збору даних. Це може бути достатньо дорогим способом, проте, такий варіант збору даних не виключений.
5. Внутрішні дані. Інформація заноситься в базу за різного роду експертними оцінками працівниками організації. Трудомісткий метод.

Зібрані дані перетворюються до єдиного формату:

* таблиць Excel,
* текстових файлів
* компонентів довільної бази даних.

Однією із важливих дій при цьому є визначення способу представлення даних. Як правило, вибирають один з наступних видів - число, рядок, дата, логічна змінна (так/ні).

Також дані повинні бути уніфікованими - одні й ті ж самі дані скрізь повинні описуватись однаково. Часто при видобуванні знань основну увагу приділяють механізмам аналізу даних, не враховуючи важливість попередньої обробки та очищення даних.

Очевидно, що некоректні початкові дані призводять до некоректних висновків. Зазначимо, що в більшості випадків джерелом інформації для аналітичних систем є сховище даних, в якому акумулюються відомості з різнорідних джерел, тому гострота проблеми суттєво зростає.

Для дослідження процесів різної природи дані повинні бути підготовлені спеціальним чином.

Зупинимося докладніше на двох типах даних:

* впорядкованих
* невпорядкованих.

Впорядковані дані потрібні для вирішення задач прогнозування - коли визначають хід того чи іншого процесу в майбутньому на основі наявних хронологічних даних. Як правило, одним з параметрів виступає дата або час, проте можуть використовуватись довільні відліки, наприклад, покази лічильників, зняті з певною періодичністю.

Невпорядковані дані потрібні для задач, де часовий чинник не має значення, наприклад, оцінка кредитоспроможності, діагностика, сегментація споживачів. В таких випадках ситуація вважається статичною і тому інформація про те, що одна подія відбулась раніше, буде неістотною.

Життєвий цикл об'єктів технології передбачає їх наступну трансформацію: інформація - дані, дані - знання та знання - знання, що складає інформаційний процес.

Інформаційні процеси поділяють на процеси циркуляції та переробки інформації, які, в свою чергу поділяються на процеси збору, передачі, обробки, збереження та спрямування до користувача.

[https://westudents.com.ua/glavy/27359-14-jittviy-tsikl-danih-zbr-sistematizatsya-danih.html]