

## **ОРГАНИЗАЦИЯ НА ИНФОРМАЦИОННАТА БАЗА НА БИС**

### **Същност и изисквания към информационната база**

Информационната база е основен компонент на информационните системи и представлява съвкупност от данни, организирани и съхранявани по определени способи, и с помощта на които се удовлетворяват потребностите на изпълняваните в информационните системи процеси.

Информационната база трябва да отговаря на следните изисквания:

- пълнота и достоверност в отразяването на особеностите на предметната област;
- да поддържа набор от данни, които да удовлетворяват изискванията на приложенията;
- да бъде организирана така, че да подлежи на развитие и усъвършенстване и обратно – развитието на приложенията да не налага преструктуриране на информационната база;
- защита на данните;
- да има минимално дублиране на информацията;
- лесно да се актуализира;
- да се осигурява поддържането на непротиворечиви данни;
- да се осигурява бързодействие в работата на системата.

Информационната база на системата може да се реализира най-често като:

- конвенционална (файлова) организация на данните;
- бази от данни
- складове от данни;
- хипертекст и др.

### **Конвенционална (файлова) организация на информационната база**

Този вид организация на информационната база предшества появата на базите данни и се използва широко при най-ранните информационни системи. При конвенционалната организация за всяко приложение се разработва собствен

набор от файлове. Дефинициите на файловете, тяхната структура и форматите на данните се поддържат от приложенията, които създават и използват файловете. Разширяването и добавянето на елементи към съществуващите приложения като правило добавя нови файлове в информационната база.

#### **Предимства на конвенционалната организация:**

- наборът от данни е ориентиран пряко към конкретното приложение;
- приложението осъществява бърз достъп до данните за четене и запис;
- файловете се създават лесно, без да са необходими допълнителни специализирани средства освен програмната среда за създаване на приложението;
- файловата информационна база е с по-малък обем в сравнение с базите от данни.

#### **Недостатъци на конвенционалната организация:**

- разделение и изолация на данните – данните са в независими файлове, не е осигурена връзка между файловете на различни приложения; обновяването на данните в едно приложение не предизвиква автоматично обновяване на същите данни в друго приложение
- дублиране на данните. Тъй като файловете обслужват конкретно приложение е възможно едни и същи данни (например за продажби, клиенти и др.) да се обработват от различни приложения и следователно да се съхраняват в различни файлове;
- противоречивост на данните, както по отношение на техния формат, така и на значенията им. Този недостатък произтича от дублирането на данните, което означава, че един и същи факт е съхранен на няколко места (в няколко файла). Дублирането на данни дава предпоставки за съхраняване на различни версии на фактите, а оттам и до получаване на противоречиви отговори на запитвания от потребителите. Противоречивостта на данните при файловата организация е свързана и с нарушаване на целостта на данните.

Данните са цялостни, ако са логически съвместими, непротиворечиви и изпълняват определени ограничения<sup>1</sup>;

- зависимост на данните от приложенията и обратно

Описанието на данните (структура на файловете, имена и тип на полетата) както и управлението на файловете е възложено на приложенията т.е. промените във физическата структура на файловете изискват (в общия случай) съответни промени в приложенията, които да са съобразени с промяната в структурата на файловете. Заедно стова промяната в програмите може да предизвика необходимост от промяна във файловете.

- ограничени средства за защита на данните;
- трудности при реализиране на многопотребителски достъп до файловете;
- не съществуват други инструменти за достъп до данните, освен чрез приложенията

### **Проектиране на файловете.**

Според начина на организация файловете биват:

- последователно организирани – за да се прочете даден запис е необходимо прочитането на всички предходни записи;
- индексно-последователно организирани – при тях записите се организират по определени полета (индекси), а в рамките на съответния индекс – последователно, в този смисъл времето за търсене по индекси е доста намалено;
- директно организирани – осигуряват възможност за директен достъп до запис;

С файлове се изпълняват следните основни операции:

- 1) Създаване – при създаване на файла се дефинира името и структурата му.
- 2) Актуализация – това са операции, свързани с промяна във файловете, като добавяне на записи, изтриване на записи и промяна в съдържанието на полетата.

---

<sup>1</sup>Кашева, М., О.Тулешкова, Ив.Куюмджиев, Бази от данни, Наука и икономика, Варна, 2009

3) Обръщение към файловете. Обръщението (търсене) е извличане на един или няколко записа от файла. Търсенето обикновено се осъществява по предварително дефинирани ключови полета (идентификатори).

4) Сортиране – операция на подреждане на записите. Тя винаги предшества извеждането на справки и се използва при обработка по нива. Записите се подреждат по един или няколко признака. В резултат на това се формират групи от записи с общо значение на признака.

5) Сливане – това е операция, при която два и повече файла с различна структура и поне един общ идентификатор се обединяват в нов файл, чиято структура е базирана на структурата на участващите в сливането файлове. Например при сливането на файл F1 (A1,A2,A3) и файл F2 (A1, A4, A5, A6), където A1-An са полета във файловете, се получава резултатен файл F3 (A1, A2, A3, A4, A5, A6). За да може да се осъществи сливането двата файла F1 и F2 трябва да имат общ идентификатор (в примера полето A1). Типична команда за сливане е JOIN.

6) Смесване – при тази операция смесването и резултатният файл са с еднаква структура, т.е. сливането представлява обединяване на файлове без промяна на структурата им. По същество се обединяват записи от участващите в операцията файлове (най-често на базата на някакво условие, критерий). Резултатният файл може да включва всички или част от записите от смесването файлове. Типична команда за смесване е APPEND.

7) Разделяне – операция, при която от един входен файл се формират два или повече резултатни файла. Разделянето може да стане по два начина:

- разделяне без промяна на структурата – формират се два или повече файла със структурата на входящия файл. Операцията е обратна на смесването на файлове;
- разделяне с промяна на структурата – резултатните файлове са с различна структура от тази на входния файл. Операцията е обратна на сливането на файлове.

При конвенционалните системи се използват следните **типове файлове**, в зависимост от характера на съдържащите се в тях данни:

- **Постоянни файлове.** Това са файлове, които съдържат относително постоянни данни, които се обновяват рядко. Типичен пример за постоянни файлове са номенклатурните файлове. Основно с номенклатурните файлове се извършват операции по редактирането (добавяне на записи, изтриване, актуализация значенията на данните).

- **Работни (променливи) файлове.** Работните файлове съдържат текуща, оперативна информация, която отразява особеностите на процесите в организациите. В тези файлове текущо (при възникване и реализиране на бизнес процес) се добавят нови записи. С други думи в тях се съдържа оперативно възникналата информация за дейността. Използват се при обработките за търсене, извличане подреждане на записи, преобразуване, агрегиране и извеждане на резултатна информация;

- **Временни файлове** – те се създават от програмата на основата извличане на записи, сливане, смесване или разделяне с цел формиране на по-ефективен набор от данни (напр. за да не се обработват всички, а само част от данните) с което се ускорява или рационализира обработката. След приключване на обработката тези файлове могат да бъдат унищожени.

### **Проектиране на файлове**

При проектиране на файловете е необходимо първо да се определи наборът от данни, нужен за изпълнение на функциите на приложението. Наборът от данни се структурира в отделни файлове в зависимост от типа на данните, от възникването им и в зависимост от тяхното използване. В резултат се формира набор от различни файлове за всяко приложение.

За всеки файл се определя:

- името на файла;
- структура на файла (набор от полета, участващи във файла);
- характеристики на полетата във файловете. Характеристиките включват: тип на полето; разрядност (дължина); значение по подразбиране и ограничения. Последните две характеристики не са задължителни, за разлика от първите две;

- първични и други ключове на всеки файл. Първичните ключове са уникални за всеки запис във файла и служат за идентифициране на записите. В процеса на обработка на файловете е необходимо дефинирането и на други ключове, необходими за реализацията на обработките.