



»Гл. ас. д-р Георги Чолаков

»Бази от данни

Въведение >

Защо бази от данни?

- » Защото са в учебния план и нямаше как да се измъкнем.
- » Защото личните ни данни вече изтекоха в Интернет пространството и е забавно да се научим как да се ровим в живота на другите.
- » Защото отвсякъде ни залива информация и е добре да се научим как да я структурираме и съхраняваме (нали уж уча за компютърджия).
- » Защото информацията е най-ценната стока, тя изобилства навсякъде и трябва някак и някъде да бъде съхранявана и управлявана.

>

Ден от живота на студента



>

Данни и информация

» Данни

- > Конкретни факти, стойности, които съхраняваме;
- > Пример: списък на студентите с техните оценки по дадена дисциплина (изпитен протокол).

» Информация

- > Данни, обработени така, че да бъдат смислени и полезни за потребителя;
- > Конкретна част от данните с определено значение – напр. оценката на студент по дисциплината или изчисляване на среден успех на курса;
- > Обработването на данните е полезно за намиране на информация и подпомага вземането на решения.

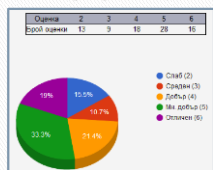
- > Данните са това, което съхраняваме;
- > Информацията е това, което извличаме.



Данни

Виртуелен номер	Студент	Резултат	Оценка
1701321013	Росена Василева	22 / 30	5
1701321010	Ангел Георгиев	25 / 30	5
1701321016	Валентин Димитров	16 / 30	4
1701321001	Светлана Ганчева	25 / 30	5
1701321005	Ивонка Каравацова	25 / 30	5
1701321032	Велик Димитров	27 / 30	6
1701321019	Светослав Петков	9 / 30	2
1701321022	Александра Илиев	30 / 30	6
1701321010	Ангел Василев	28 / 30	5
1701321008	Георги Георгиев	28 / 30	5
1701321006	Владимир Ангелов	18 / 30	4

Информация



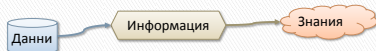
Информация → Знание

» В епохата на информацията:

- > Генерирането на верна, полезна и навременна информация е ключ към вземането на верните решения.

» В епохата на знанията:

- > Вземането на верните решения, от своя страна, е ключ към оцеляването на бизнеса в условията на глобален пазар.



- » Данните са основата на информацията.
- » Информацията е основата на знанията.
- » Нови знания могат да бъдат извлечени от досегашните.



Пирамида на знанието

» Data, Information, Knowledge, Wisdom (DIKW) Pyramid



Пирамида на знанието

» Данни – числа, низове и т.н. в прост вид, без обработка и значения, напр. 1601261054 – последователност от цифри, но ако е в контекст на факултетен номер ще разпознаем година на записване, специалност и др. Така трансформирахме данните в **информация**.



Пирамида на знанието

» Информация – данни в определен контекст или след обработка, притежаващи смисъл/значение, удобни за анализиране – въпросите КОЙ, КОГА, КАКВО могат да извлекат полезна информация от данните. Но въпросът КАК прави стъпката към **знанието**;



Пирамида на знанието



- » Знание – когато разбираме информацията не просто като събрани данни и факти, а така, че да я приложим за постигане на цели, тя се превръща в знание;
- » Откривайки взаимоотношения, които не са били явни в информацията, получаваме по-задълбочено разбиране за контекстната област;
- » Използвайки тези знания за предприемане на стъпки и вземане на решения стигахме до **мъдрост**;



Пирамида на знанието



- » Мъдрост – за да стигнем дотук трябва да сме отговорили на въпросите „Защо правим това?“, „Кой е най-добрият подход?“ – това са знания, преведени в действие.



Да обобщим...

- » Навременната и полезна информация изисква прецизни и верни данни.
- » Тези данни трябва да бъдат съхранени във формат, лесен за достъп и обработка, и да бъдат добре управлявани и поддържани.
- » С оглед на решаващата роля на данните, управлението им е задача, фокусирана върху правилната обработка, съхранение и извличане. Това е основна задача на всеки бизнес, който използва данни.



Какво е база от данни?

» Компютърна система, чиято цел е да съхранява:

- > **Данни** - факти, стойности, представляващи интерес за крайния потребител. Данните могат да имат логически връзки помежду си, които също да се съхраняват;
- > **Метаданни** – описание на данните, както и собствената си структура – имена на обекти, типове данни на атрибутите (integer, char, date...) и др.

» Система за управление на база от данни (СУБД) е множество от програми, които:

- > Управяват структурата на данните;
- > Контролират достъпа до тях;
- > Помагат за лесното администриране на цялостната система.



Видове данни от гледна точка на БД

» Оперативни данни

Данните (фактите), за съхранението на които използваме базата данни.

» Входни данни

Първична информация, която се въвежда в системата с помощта на клавиатура, скенер, светлина писалка и т.н. Те могат да станат оперативни или да предизвикат промяна на оперативни данни.

» Изходни данни

Резултати от обработката на заявки, които се извеждат на печат или на екран. Те се получават като резултат от определени операции с оперативните данни и могат да не присъстват в този си вид като оперативни в самата БД.



Компоненти на една (СУБД)

» Хардуер - този компонент е необходим, за да е налице платформа на която ще работи СУБД.

» Софтуер - за поддръжка на една СУБД са необходими разнообразни програми, напр. операционна система и софтуер, изграждащ самата СУБД.

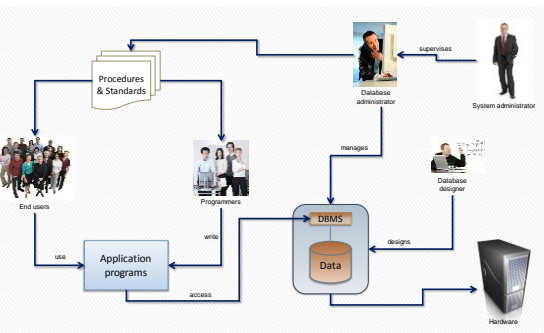
» Данни - съхраняваните в базата факти.



Компоненти на една (СУБД)

- » **Процедури** - действията, които потребителите могат да извършват, например:
 - > включване в системата;
 - > промяна на дизайна ѝ;
 - > въвеждане, редактиране, изтриване на данни;
 - > стартиране, спиране, възстановяване и управление на производителността.
- » **Потребители** - биват различни видове:
 - > Крайни потребители - използват системата за извършване на различни информационни дейности;
 - > Приложни програмисти – разработчиците на приложения за работа с базата данни;
 - > Администратори - поддържат цялостното ѝ функциониране.





Ползи от СУБД

- » По-лесно споделяне на данни – крайните потребители имат лесен достъп до организирани и добре структурирани данни;
- » Сигурност на данните – налице са различни механизми за управление на достъпа;
- » Интегритет на данните – възможност за по-широк поглед върху голямата картина;
- » Минимизиране на неконсистентността – неконсистентност означава съхранение на едни и същи данни на различни места, но имащи различни стойности - напр. оценката в книгката и в протокола да се различават;
- » И др.



Класификация според брой потребители

- » **Single-user** – поддържа само един потребител в даден момент, обикновено работи на desktop машина;
- » **Multi-user** – осигурява достъп до много потребители в един и същ момент.



Класификация според място на съхранение на данните

- » **Централизирана** – данните се съхраняват в една база от данни, на едно място;
- » **Разпределена** – съхранява данните на различни места в мрежата, обикновено.



Класификация според модела на данните

- » Йерархичен модел;
- » Мрежов модел;
- » Релационен модел;
- » Обектно-ориентиран модел;
- » Обектно-релационен модел;
- » Други – NoSQL.



Класификация според начин на употреба

Операционни (транзакционни) – Online Transaction Processing (OLTP)

- » Предназначени за отразяване на ежедневните бизнес дейности – продажби, плащания и т.н. – т.е. обработка на транзакции в реално време, ситуации, където е налице ежедневно събиране, променяне и поддържане на данни;
- » Типът данни, който се съхранява в такава база данни, е динамичен, което означава, че той се променя постоянно и винаги отразява най-новата информация;
- » Такъв тип бази данни използват магазини и складове за продажба на стоки, банки, телекомуникационни компании и др. ➤

Класификация според начин на употреба

Аналитични – Online Analytical Processing (OLAP, Decision Support Systems)

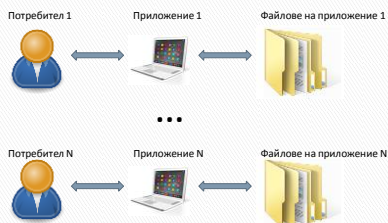
- » Аналитичните бази данни се използват главно при аналитична обработка в реално време, където има нужда да се съхраняват и проследяват минали и зависещи от времето данни;
- » Една такава база данни е ценно предимство, когато има нужда да се следят тенденциите, да се извлича статистическа информация за определен период и да се правят тактически и стратегически бизнес прогнози;
- » Този вид бази данни съхраняват статични данни, което означава, че данните не се променят или се променят много рядко. Намират приложение в компаниите за маркетингови анализи, изследователски лаборатории, състезателни отбори от Формула 1 и др.;
- » Обикновено тези бази данни се състоят от хранилище за данни (data warehouse) и приложения за аналитична обработка (front-end). Данните в тях са във формат, оптимизиран за вземане на решения. Data warehouse системите съдържат исторически данни, получени от операционни бази от данни и други външни източници, претърпявайки процес на подходяща трансформация. ➤

Съвременни течения

Развитие на WWW и социални медии и мрежи

- » Системи като Google, Facebook, Twitter, LinkedIn и др. събират огромни масиви от данни за потребителите;
- » През последните години се появи нова разновидност бази от данни – т.нар. NoSQL (Not only SQL) – използвани за съхранение на безпрецедентни обеми от данни, различни типове данни и структури. ➤

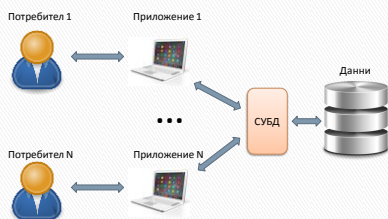
Системи за работа с файлове



Системи за работа с файлове - недостатъци

- » Зависимост на данните – приложенията, работещи с файловете, са зависими от формата на данните в тях;
- » Данните често се дублират – възможност за загуба на интегритет;
- » Данните са разделени и съответно изолирани – това затруднява комбинирането на данните. В приложенията трябва да е добре направена синхронизацията при обработката на данни от два файла, за да е сигурно, че извлечените данни са коректни;
- » Трудно е да се представят сложни обекти като се използват системи за работа с файлове.

Системи за управление на бази от данни



Цели на СУБД

- » Сигурност и защита на данните от неправомерен достъп;
- » Надеждност на обработка и съхранение;
- » Обслужване на множество потребители във всеки момент;
- » Лесен достъп и манипулиране на данните;
- » Консистентност (съгласуваност, вярност) на данните;



Цели на СУБД

- » Физическа независимост на данните – структурата на съхранение е прозрачна за приложенията, използващи данните;
- » Логическа независимост на данните – логическата структура може да бъде променяна без това да засегне приложенията, използващи данните;
- » Контрол на излишеството (повторението) – даден факт да се съхранява на едно място;
- » Контрол на интегритета – валидиране на данните;
- » И др.