

Fundamentals Concepts of Data Bases

اصول طراحی پایگاه داده ها

جلسه ۱ : مفاهیم اولیه

مدرس : اسماعیل نورانی

<http://www.nurani.ir/> - Info@nurani.Ir

لیست مراجع :

- ۱- بانک اطلاعات علمی کاربردی (جلداول) دکتر مصطفی حق جو ۱۳۸۶ (نشر: دانشگاه علم و صنعت ایران)
- 2- Database System Concepts by Abraham Silbershatz 2007
eBook Download Link : <http://www.flazx.com/ebook7976.php>
- 3- An Introduction to Database Systems , 8th Edition , by C. J. Date , 2004 Addison-Wesley
مترجم : حسین ابراهیم زاده قلزم (نشر: دانشگاه صنعتی امیر کبیر)
- ۴- مفاهیم بنیادی پایگاه داده ها سید محمد تقی روحانی رانکوهی ، ویراست دوم ۱۳۸۳

بارمبندی :

پایان ترم ۱۴ نمره

میانترم ۳ نمره

Quiz ۱ نمره

پروژه ۲ نمره

عناصر اصلی محیط پایگاه داده ها

• سخت افزار (Hardware)

• نرم افزار (Software)

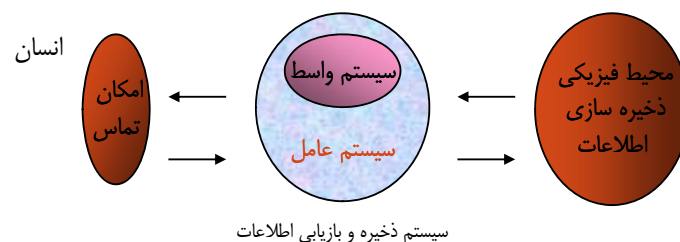
• کاربر (User)

• داده (Data)

سیستم واسط ذخیره و بازیابی - سیستم ذخیره و بازیابی اتوماتیک (کامپیوتری)



- محیط فیزیکی ذخیره سازی اطلاعات : شامل داده های ذخیره شده روی رسانه ها و خود رسانه ها می باشد .



مروری بر نسلهای تکنولوژی سیستم واسط

- نسل اول (۱۹۵۷ - ۱۹۴۸) : نسل بدون سیستم واسط
- نسل دوم (۱۹۶۶ - ۱۹۵۸) : نسل شیوه های دستیابی (Access Methods)
- نسل سوم (۱۹۷۳ - ۱۹۶۷) : نسل سیستم مدیریت داده ها (Data Management System)
- نسل چهارم (۱۹۸۰ - ۱۹۷۴) : نسل سیستم مدیریت پایگاه داده ها (Database Management System)
- نسل پنجم (تاکنون - ۱۹۸۱) : نسل بانکهای معرفت (Knowledge Bases)

نسل اول (۱۹۴۸-۱۹۵۷) : نسل بدون نرم افزار واسط

- رسانه ذخیره سازی معمولاً نوار بود .
- هرگونه تغییر در رسانه های ذخیره سازی یا ساختار فایل منطقی (فیزیکی) سبب تغییر در برنامه ایجاد کننده و پردازشگر فایل می شود .
- اشتراک داده ها مطرح نیست
- تکرار در ذخیره سازی (افزونگی - Redundancy) داده ها در بالاترین حد است .

نسل دوم : نسل شیوه های دستیابی (۱۹۵۸ - ۱۹۶۶) (Access Methods)

- نرم افزارهایی به نام شیوه های دستیابی ایجاد شدند .
- **تعریف شیوه دستیابی (AM) :** نرم افزاری است که به جنبه های فیزیکی محیط ذخیره سازی و عملیات در این محیط می پردازد ، به نحوی که دیگر برنامه کاربر نیازی به پرداختن به این جنبه ها را ندارد .
- برنامه های کاربردی در قبال تغییرات در سخت افزار ذخیره سازی مصونیت یافتند .

نسل سوم : نسل سیستم مدیریت داده ها (۱۹۶۷-۱۹۷۳) (DMS) (Data Management System) :

- نرم افزاری نسبتاً پیچیده بنام سیستم مدیریت داده ها ، واسط است بین برنامه های کاربران و محیط فیزیکی ذخیره سازی .
- میزان تکرار ذخیره سازی کاهش می یابد .
- داده های مشترک در کاربردهای متنوع بکار می روند .

نسل چهارم : نسل سیستم مدیریت پایگاه داده ها

(Database ManagementSystem) (DBMS) (۱۹۷۴-۱۹۸۰)

- نسل جهش در تکنولوژی ذخیره و بازیابی اطلاعات بود .
- مستقل شدن برنامه های کاربردی از جنبه ها و خصوصیات محیط فیزیکی ذخیره سازی که آن را اصطلاحاً **استقلال داده فیزیکی** (Physical Data Independence) می نامند .
- نرم افزار پیچیده و جامع موسوم به **سیستم مدیریت پایگاه داده ها** واسط بین برنامه های کاربران و محیط داخلی فیزیکی ذخیره سازی است .

نسل پنجم نسل بانکهای معرفت و شی گرا (تاکنون-۱۹۸۱)

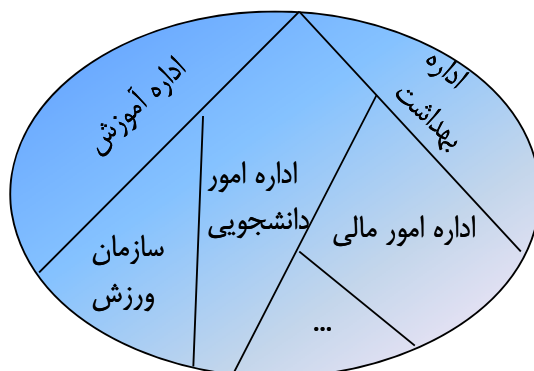
- Knowledge bases
- Object oriented databases
- Object relational database (Oracle)

پایگاه داده ها (Database)

تعریف: پایگاه داده ها مجموعه ای است از داده های ذخیره شده ، بطور یکپارچه (integrated) ، حتی الامکان بدون افزونگی (Redundancy) (با افزونگی کنترل شده) ، مبتنی بر یک ساختار داده ای (DS) مشخص ، تحت مدیریت متمرکز (Centralized Management) مورد استفاده یک یا چند کاربر ، بطور اشتراکی و همزمان.

مثال مقدماتی (صرفا درک بهتر تعریف):

محیط عملیاتی دانشگاه: زیر محیطها شامل:
هر زیر محیط مجموعه نیازهای اطلاعاتی خاص دارد.



مجموعه نیازهای اطلاعاتی ← مجموعه Application Programs : سیستم حقوق و دستمزد،

سیستم پرسنلی، سیستم اطلاعات دانشجویی، ...

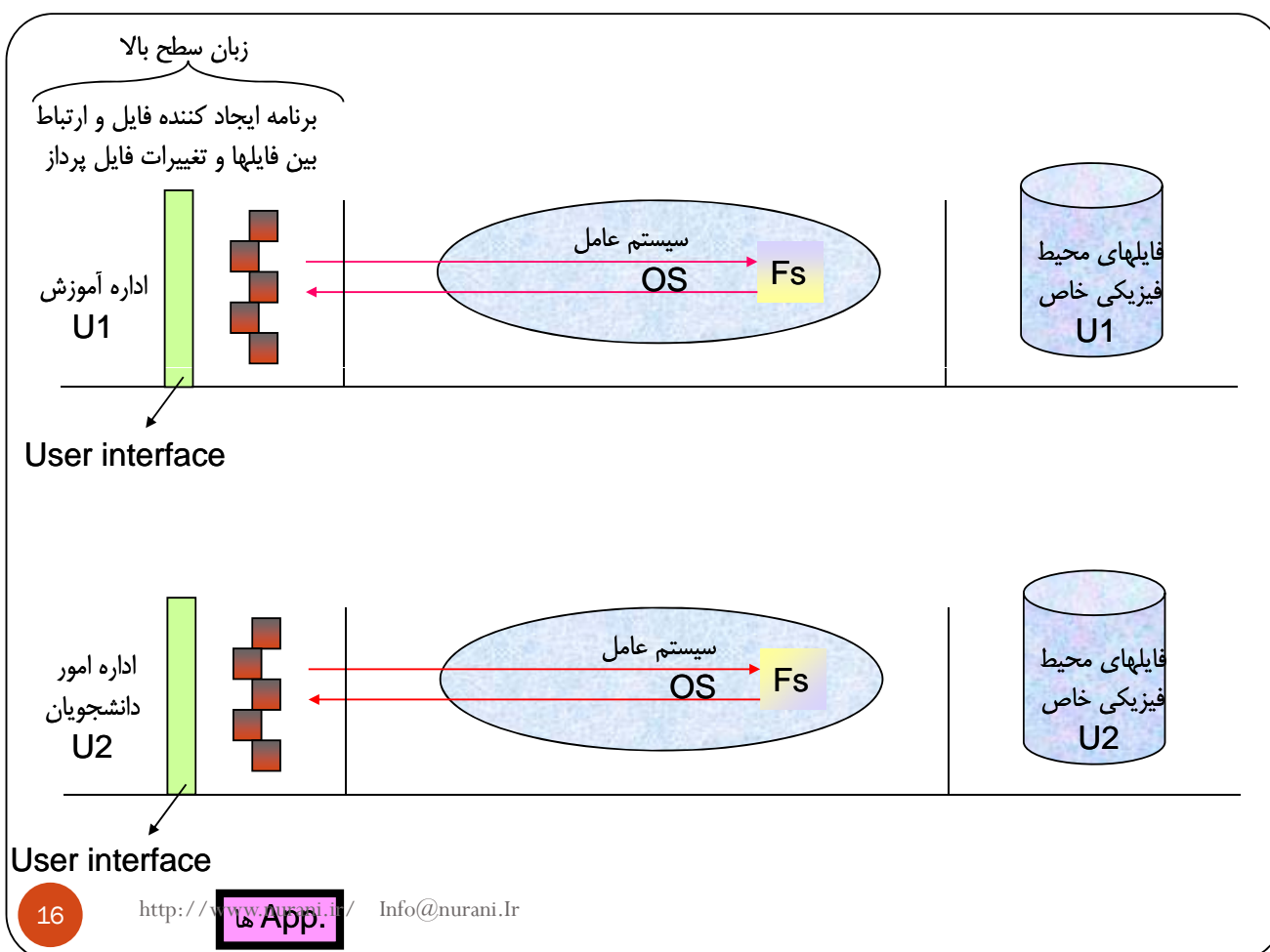
دو روش کلی برای ایجاد سیستم اتوماتیک این محیط عملیاتی وجود دارد:

– روش فایلینگ (غیر بانکی) (Non database policy)

– روش بانکی (Database policy)

روش فایلینگ

- هر یک از زیر محیطها بطور جداگانه مورد مطالعه و تجزیه و تحلیل قرار می گیرند.
- طراحی تعدادی فایل



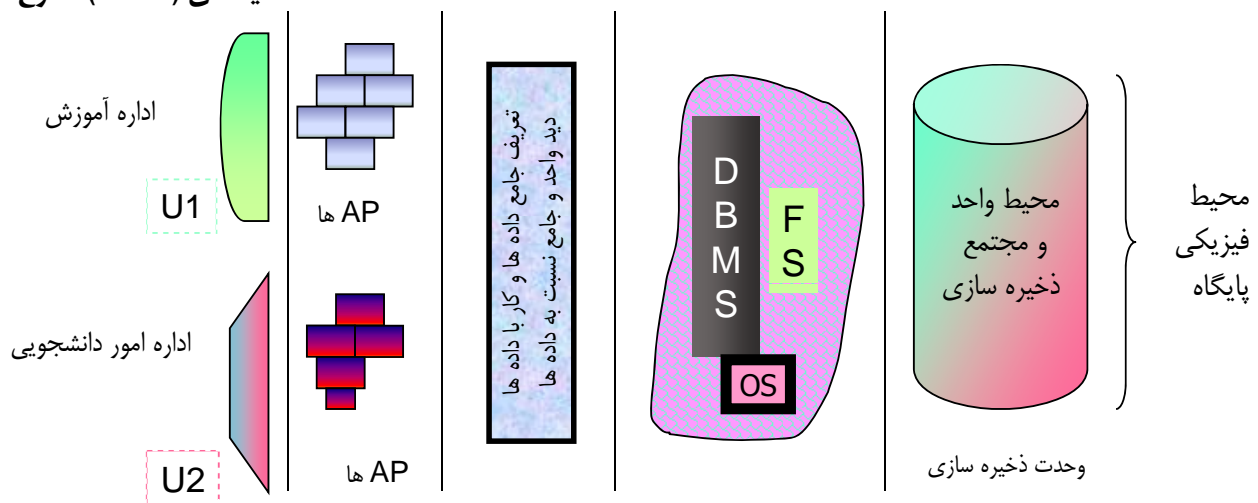
معایب روش غیر بانکی

- **نا مجتمع بودن** (یکپارچه نبودن) داده های ذخیره شده (سیستمهای پراکنده)
- درجه قابل ملاحظه ای از **افزونگی** (Redundancy) در ذخیره سازی
- خطر بروز پدیده **ناسازگاری داده ها** (Data inconsistency) (ناهمخوانی)
- **حجم بالای برنامه سازی**: تعداد بیشتر برنامه ها و بالا بودن حجم برنامه ها
- وابستگی برنامه های کاربردی به محیط فیزیکی ذخیره سازی اطلاعات.

(عدم وجود استقلال فیزیکی)

روش بانکی (Database Policy):

دیدهای (View) متنوع



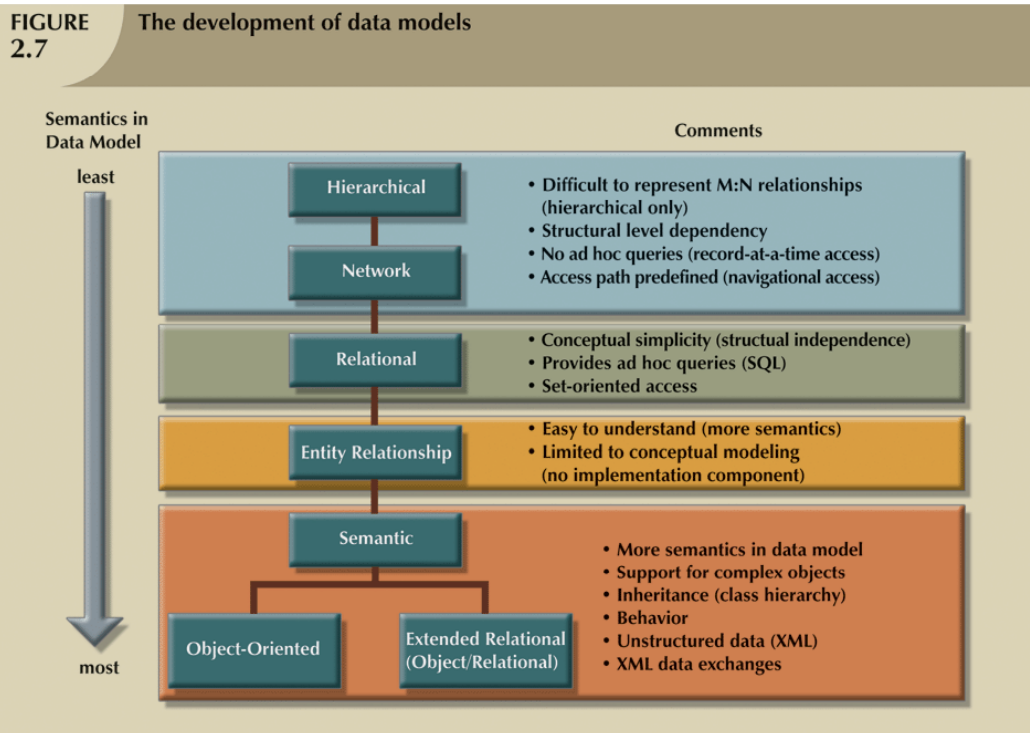
مدلهای طراحی پایگاه داده

- مدل‌های قدیمی :
- ۱- مدل سلسه مراتبی (Hierarchical Model)
 - برای سازماندهی داده ها از درخت استفاده شده است.
- ۲- مدل شبکه ای (Network Model)
 - برای سازماندهی داده ها از گراف استفاده شده است
- مدل رابطه ای (Relational)
 - به صورت جدول (Table) پیاده سازی میشود.
 - به صورت مجموعه ای از رابطه ها (Relation) طراحی می شود.
 - موفق ترین و پرکاربردترین مدل پایگاه داده می باشد.

تکامل مدل‌های داده ای

TABLE 2.1 Evolution of Major Data Models				
GENERATION	TIME	MODEL	EXAMPLES	COMMENTS
First	1960s–1970s	File System	VMS/VSAM	Used mainly on IBM mainframe systems Managed records, not relationships
Second	1970s	Hierarchical and Network Data Model	IMS ADABAS IDS-II	Early database systems Navigational access
Third	Mid-1970s to present	Relational Data Model	DB2 Oracle MS SQL-Server MySQL	Conceptual simplicity Entity Relationship (ER) modeling and support for relational data modeling
Fourth	Mid-1980s to present	Object-Oriented Extended Relational	Versant FastObjects.Net Objectivity/DB DB/2 UDB Oracle 10g	Support complex data Extended relational products support objects and data warehousing Web databases become common
Next Generation	Present to future	XML	dbXML Tamino DB2 UDB Oracle 10g MS SQL Server	Organization and management of unstructured data Relational and object models add support for XML documents

تکامل مدل‌های داده ای



21

<http://www.nurani.ir/> Info@nurani.Ir

مدل رابطه ای (Relational)

- در سال ۱۹۷۰ توسط کاد (E. F. Codd) در شرکت IBM توسعه داده شد.

جدول (رابطه)

- ماتریسهایی که متشکل از چندین سطر و ستون هستند.
- هر سطر در یک رابطه یک تاپل (Tuple) نامیده می شود.

22

<http://www.nurani.ir/> Info@nurani.Ir

ارتباط دو جدول در مدل رابطه ای

FIGURE 2.3 Linking relational tables

Table name: AGENT (first six attributes)

Database name: Ch02_InsureCo

AGENT_CODE	AGENT_LNAME	AGENT_FNAME	AGENT_INITIAL	AGENT_AREACODE	AGENT_PHONE
501	Alby	Alex	B	713	228-1249
502	Hahn	Leah	F	615	882-1244
503	Okon	John	T	615	123-5589

Link through AGENT_CODE

Table name: CUSTOMER

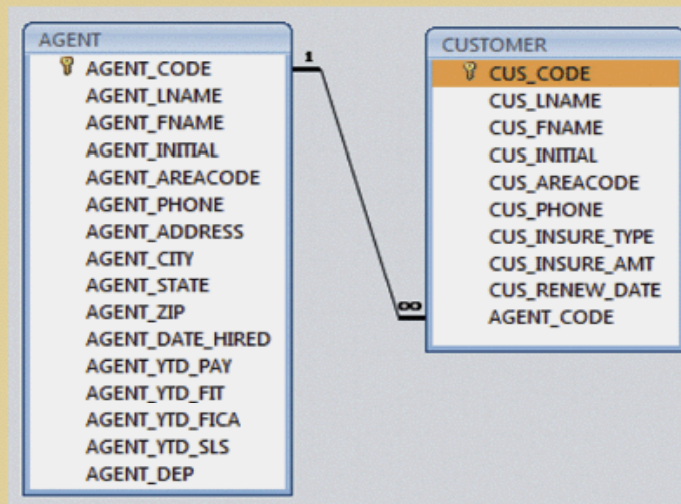
CUS_CODE	CUS_LNAME	CUS_FNAME	CUS_INITIAL	CUS_AREACODE	CUS_PHONE	CUS_INSURE_TYPE	CUS_INSURE_AMT	CUS_RENEW_DATE	AGENT_CODE
10010	Ramas	Alfred	A	615	844-2573	T1	100.00	05-Apr-2008	502
10011	Dunne	Leona	K	713	894-1238	T1	250.00	16-Jun-2008	501
10012	Smith	Kathy	W	615	894-2285	S2	150.00	29-Jan-2009	502
10013	Olowski	Paul	F	615	894-2180	S1	300.00	14-Oct-2008	502
10014	Orlando	Myron		615	222-1672	T1	100.00	28-Dec-2008	501
10015	O'Brian	Amy	B	713	442-3381	T2	850.00	22-Sep-2008	503
10016	Brown	James	G	615	297-1228	S1	120.00	25-Mar-2009	502
10017	Williams	George		615	290-2556	S1	250.00	17-Jul-2008	503
10018	Farriss	Anne	G	713	382-7185	T2	100.00	03-Dec-2008	501
10019	Smith	Olette	K	615	297-3809	S2	500.00	14-Mar-2009	503

23

<http://www.nurani.ir/> Info@nurani.Ir

ارتباط دو جدول در مدل رابطه ای

FIGURE 2.4 A relational diagram



24

<http://www.nurani.ir/> Info@nurani.Ir

تعریف

- **Meta Data (داده در مورد داده)**
 - داده هایی به غیر اطلاعات ورودی و خروجی که در پایگاه داده ذخیره می شوند.
 - ۱- لغتنامه داده ها (Data Dictionary):
 - در یک سیستم بانک اطلاعاتی اسامی زیادی مورد استفاده قرار می گیرند.
 - هماهنگی در نام داده ها و معنای آنها ضروری است.
 - لغتنامه داده ها همه اسمای استفاده شده در سیستم و معنای آنها را در بر می گیرد.
 - اسامی از قبیل نام جداول نام Object ها، صفات (Attribute) و
 - ۲- کاتالوگ سیستم (System Catalog)
 - علاوه بر اسامی داده ها و معنای آنها، اطلاعات دیگری از قبیل سطح دسترسی کاربران به داده ها، تاریخ ایجاد و به روز رسانی داده ها و ... در یک سیستم پایگاه داده ذخیره می شود.
 - کاتالوگ سیستم در برگیرنده لغتنامه داده ها نیز می باشد.

تراکنش (Transaction)

- هر گونه برنامه ای که توسط کاربر در محیط بانک اطلاعات اجرا میشود تراکنش نام دارد.
- چهار کنترل زیر لازم است روی تمامی تراکنش ها در بانک اطلاعات اعمال گردد تا صحت و جامعیت آن تضمین شود.
- تراکنش به دو صورت می تواند پایان پذیرد:
 - پایان ناموفق که آنرا «سقوط» (**Abort**) می نامند.
 - پایان موفق که آنرا «انجام» (**Commit**) می نامند.
- به کنترلهای فوق اصطلاح خواص **ACID** گفته می شود.
 - یکپارچگی (**Atomicity**)
 - همخوانی (**Consistency**)
 - انزوا (**Isolation**)
 - پایداری (**Durability**)

تراکنش (Transaction)

- یکپارچگی (Atomicity)
 - این ویژگی به خاصیت «همه یا هیچ» معروف است.
 - یا تمامی دستورالعملهای یک تراکنش باید اجرا شود و یا هیچ یک.
 - بعنوان مثال در تراکنشی که از حساب X پول برداشت کرده و در حساب Y واریز می کند، بدیهی است که به هر علتی اگر برداشت پوا انجام بگیرد و واریز به حساب دیگر با شکست مواجه شود، بخش اول تراکنش نیز باید لغو شود. (Rollback)
- همخوانی (Consistency)
 - هر تراکنش، پایگاه داده را از یک وضعیت صحیح به وضعیت صحیح دیگری منتقل کند.
 - بعنوان مثال مبلغ برداشت شده از حساب با مبلغ واریز شده مغایرت نداشته باشد.

تراکنش (Transaction)

- انزوا (Isolation)
 - بر طبق این ویژگی اثر تراکنش های همروند بر یکدیگر چنان است که گویی در انزوا انجام می شوند.
 - تراکنش های همروند در پایگاه داده توسط بخشی به نام واحد کنترل همروندی (Concurrency Control) صورت می گیرد.
- پایداری (Durability)
 - بر اساس این خاصیت، تراکنش هایی که مرحله انجام (Commit) می رسند، بایستی اثرشان ماندنی بوده و هرگز به طور تصادفی از بین نروند.

تعریف

- زبان پرس و جو (Query Language)
 - یک نوع زبان بیانی (Declarative) برای آنکه کاربر بتواند درخواست خود را بیان کند.
 - زبان های بیانی نسبت به زبانهای روالی (Procedural) مثل پاسکال و .. کاربرد راحتی داشته و در مقابل امکانات کمتری دارند.
- زبان میزبان (Host Language)
 - یک زبان سطح بالا مانند دلفی یا C# که امکان استفاده از زبان بیانی مانند SQL را درون خود فراهم می کنند.
- CASE (Computer Aided Software Engineering)
 - ابزارهای کمکی مورد استفاده در طراحی و پیاده سازی بانک اطلاعات

تعریف

- انواع زبانهای استفاده شده در پایگاه داده یا به عبارت دیگر بخشهای مختلف زبان فرعی داده ای (DSL)(Data Sub Language)
 - زبان تعریف داده ها (DDL)(Data Declaration Language)
 - کارهایی از قبیل تعریف جداول در مدل رابطه ای
 - زبان کار با داده ها (DML)(Data Manipulation Language)
 - عملیاتی از قبیل استخراج اطلاعات از جداول
 - زبان کنترل داده ها (DCL)(Data Control Language)
 - اعمال نظارتی همچون حقوق دسترسی کاربران