

فصل پنجم

مدیریت منابع داده‌ها

اهداف یادگیری

پس از مطالعه این فصل شما باید بتوانید:

- ارزش تجاری بکارگیری فرایندها و فن‌آوری‌های مدیریت منابع داده‌ها را برای سازمان توضیح دهید.
- مزایای رویکرد مدیریت پایگاه داده‌ها را برای مدیریت منابع داده‌های کسب و کار در مقایسه با رویکرد پردازش فایل مشخص کنید.
- توضیح دهید که چگونه نرم‌افزارهای مدیریت پایگاه داده‌ها به اهالی کسب و کار کمک می‌کنند و از عملیات و مدیریت یک کسب و کار پشتیبانی می‌نمایند.
- مثال‌هایی ارائه کنید تا هر یک از مفاهیم زیر را روشن کند:
 - انواع عمده پایگاه داده‌ها
 - ابزارداری داده‌ها و داده‌کاوی
 - عناصر منطقی داده‌ها
 - ساختارهای اساسی پایگاه داده‌ها
 - توسعه پایگاه داده‌ها

های اطلاعات مدیریت

سیستم و برنامه‌های
شبکه و داده‌های
سیستم، می‌توان به
اشاره کرد. برنامه‌های
می‌کنند. سیستم‌های
برنامه‌های کمکی
را به عنوان بخشی از
زبان و ویرایشگرهای
از فرایندهای کسب

CP نظارت می‌کند.
راواتی ارائه می‌کند
م و ارتباطات شبکه
ها برای اداره برنامه
های کمکی و سایر

این زبان‌ها برای
بکار می‌گیرند و
فای برنامه‌نویسی
هستند. زبان‌های
و خدمات تحت

شرکت هنری شاین: ارزش تجاری یک انبار داده‌ها

اعلیٰ شرکت‌ها داده‌های زیادی درباره مشتریان‌شان ذخیره می‌کنند. چالش فن‌آوری اطلاعات در این حوزه نحوه‌ی ادغام و ارسال این اطلاعات به گونه‌ای است که واحدهای کسب و کار بتوانند بلافاصله به تغییرات فروش و ترجیحات مشتریان واکنش نشان بدهند.

شرکت هنری شاین (www.henryschein.com) راه حل این مسأله را پیدا کرده است. این شرکت ۲٫۸ میلیارد دلاری که در زمینه بخش محصولات بهداشتی و درمانی فعالیت می‌کند با استفاده از یک تیم شش نفره از متخصصان سیستم‌های اطلاعاتی داخلی‌اش یک انبار داده طراحی نموده است. معاونت اطلاعاتی سازمان، جیم هاردینگ می‌گوید که می‌دانسته که برخورداری از مهارت‌های درست برای پروژه انبار داده‌ها ضرورت دارد، ولی در آن زمان شرکت هیچ تجربه‌ای برای طراحی چنین سیستمی نداشت. بنابراین او و گریس موناهان، معاون سیستم‌های کسب و کار افرادی را تحت عنوان «تیم شاین» استخدام نمودند.

با توجه به این که هاردینگ دو ابزار اصلی برای انبار داده‌ها انتخاب کرده بود (نرم‌افزار بازسازی داده‌های شرکت انفورماتیکا و نرم‌افزارهای query و گزارش‌دهی به کاربر از شرکت مایکرواستراتژی) آنها به دنبال کسانی بودند که قبلاً با این نرم‌افزارها کار کرده باشند. بنابراین موناهان سه نفر را وارد شرکت کرد: مدیر پروژه داریل کلی، الگوساز داده‌ها کریستین بیتز، و کارشناس رابط نرم‌افزاری رینا لیوای که مسئول طراحی رابط کاربر و تحلیل داده‌ها و نیز پشتیبانی و آموزش کاربران بود. داوون سان مسئول بازاریابی، تبدیل و بارگذاری داده‌ها بود و مدیر پایگاه داده جمیل اودین دو کار اصلی دیگر را اجرا می‌کرد. عضو دیگر تیم از گروه توسعه کاربردهای شاین به این تیم انتقال یافته بود.

بعد از جذب مهارت‌ها ضروری، اولویت بعدی کسب اطمینان از کیفیت داده‌ها بود. هاردینگ می‌گوید «این نکته تقریباً بدیهی به نظر می‌آمد ولی گاهی در چنین پروژه‌هایی کیفیت فراموش می‌شود و آن وقت انبار داده‌ها بی‌فایده می‌شود زیرا کسی به داده‌های آن اعتماد نمی‌کند.» بنابراین در ابتدای پروژه، تیم با ۱۷۵ نفر از کاربران بالقوه سیستم مصاحبه کرد تا نوع اطلاعاتی که آنها نیاز دارند و گزارش‌هایی که می‌خواهند ببینند را مشخص کند. موناهان می‌گوید این گام‌ها اهمیت طولانی داده‌ها در سیستمی که باید از حالت تراکنشی به حالت انبار داده‌ها برود روشن نمود. این کار باعث شد زمانی برای‌شان تولید شود.

کیمبال می‌گوید «فقط افراد داخلی سازمان می‌توانند بدانند که سیستم‌شان واقعاً چطور کار می‌کند، کدام داده‌ها خوب هستند و کدام داده‌ها خوب نیستند، و این که کاربر نهایی چطور می‌خواهد از داده‌ها استفاده کند.» هاردینگ می‌گوید «تضمین کیفیت داده‌ها دشوارترین بخش پروژه است، زیرا جزئیات زیادی دارد و زمان زیادی می‌گیرد و کسی که مثل داریل قبلاً چند تا از این پروژه‌ها انجام نداده باشد اهمیت آن را درک نمی‌کند.»

و یک مانع خسته‌کننده دیگر هم وجود داشت. انبار داده‌ها طوری طراحی شده بود که اطلاعات بسیار دقیقی درباره مشتریان فراهم کند. هاردینگ می‌گوید «با این کار می‌توانستیم طبق خواستمان اطلاعات را تجزیه یا ترکیب کنیم.» ولی این امر باعث کند شدن عملکرد سیستم می‌شد. بنابراین تیم جدول‌های خلاصه شده‌ای طراحی کرد تا queryها سریع‌تر کار کنند. و این جدول‌ها باید آزمایش می‌شدند. هاردینگ می‌گوید این فرایند خیلی طول کشید ولی در نهایت خیلی خوب کار کرد. این برنامه بیش از دو سال به طول انجامید.

البته ساختن یک انبار داده کاری است که هیچ وقت تمام نمی‌شود. شرکت‌های جدید خریداری می‌شوند. محصولات جدید اضافه می‌شوند. مشتریان می‌آیند و می‌روند و ویژگی‌ها و بهبودهای جدیدی به سیستم اضافه می‌شود. ولی از دیدگاه IT انبار داده‌ها تکمیل شده و ۸۵ درصد داده‌هایش را از سیستم تراکنش اصلی می‌گیرد. هدف اصلی بعدی ایجاد یک سیستم انبار داده برای عملیات شعبه اروپا و متصل کردن آن به سیستم ایالات متحده است.

هاردینگ می‌گوید پروژه وی قطعاً ارزش هزینه کردن را دارد. ولی اعداد و ارقام لازم در این خصوص را ندارد. وی می‌گوید «ما بازگشت سرمایه را به صورت رسمی برای پیگیری‌های بعدی محاسبه نکردیم. ما به این خاطر پروژه را انجام می‌دهیم که برای کسب و کارمان ارزش دارد.»

لو فرارو، معاون و مدیر کل گروه دارویی شرکت شاین می‌گوید مزایای تجاری آن فوق‌العاده هستند. وی اکنون می‌تواند سودآورترین مشتریان و مشتریان هدف برای تخفیف و افزایش فروش را شناسایی کند و می‌تواند کسب و کارش را با توجه به انواع محصول یا از لحاظ حوزه فروش بررسی کند. فرارو می‌گوید انبار داده‌ها علاوه بر این به وی کمک می‌کند مشتریان مناسب برای بازاریابی الکترونیکی را شناسایی کند که سالانه بیش از ۲۵ میلیون قطعه از آنها خرید می‌کنند.

یکی از باارزش‌ترین ویژگی‌های انبار داده این است که کاربران قادر می‌سازد در حین استفاده از سیستم بحث‌های جدیدی را به گزارش‌شان اضافه کنند. فرارو می‌گوید «بعد از این که یک گزارش اصلی را ایجاد کردید، نتیجه‌گیری کردید و بر اساس مفروضات آن داده‌ها را تحلیل نمودید، سیستم به شما اجازه می‌دهد به جای این که مجبور شوید گزارش جدیدی ایجاد کنید، همان گزارش را بهبود بدهید.» واحد IT قبلاً برای تحلیل روندهای فروش و بازار باید گزارش‌های متعددی را به صورت روزانه، هفتگی، ماهانه و فصلی ایجاد می‌کرد، ویرایش و مرور می‌کرد، اجرا می‌کرد، دانلود می‌کرد، بازنویسی می‌کرد و پرینت می‌گرفت. ولی امروز کاربران خودشان اطلاعات را در زمان کوتاهی جستجو می‌کنند، مرتب می‌کنند و تحلیل می‌نمایند. هاردینگ می‌گوید «انبار داده‌ها بخشی از فرهنگ ما شده است. این سیستم در سازمان ما مقدس تلقی می‌شود.»

پوشش‌ها

۱. برخی الزامات ایجاد یک انبار داده خوب را نام ببرید. از شرکت هنری شاین برای نمونه استفاده کنید.
۲. نرم‌افزارهای اصلی برای ساخت و بکارگیری یک انبار داده کدامند؟
۳. ارزش تجاری یک انبار داده برای شرکت هنری شاین یا هر شرکت دیگر چیست؟

قسمت ۱

مبانی فنی مدیریت پایگاه داده‌ها

مدیریت منابع داده‌ها^۱

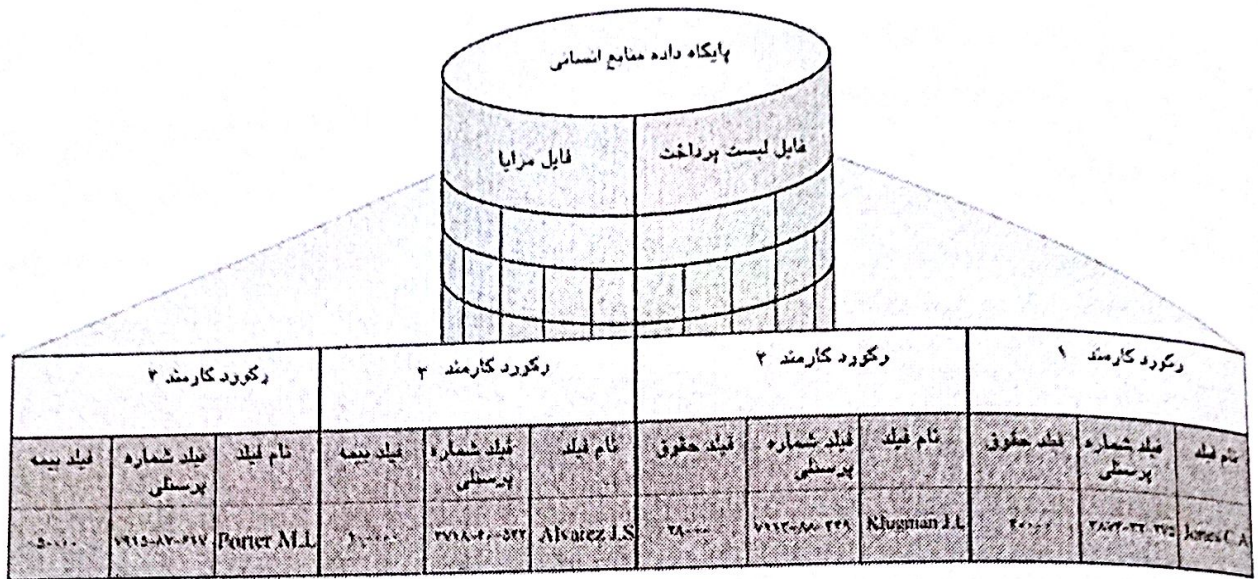
تصور کنید اگر روشی سازمان‌یافته برای ذخیره‌سازی یا بازیابی داده‌ها و اطلاعات سازمانی وجود نداشت، گرفتن اطلاعات از سیستم اطلاعاتی چقدر دشوار می‌شد. بنابراین، در تمام سیستم‌های اطلاعاتی، داده‌ها باید به شکلی منطقی سازماندهی و ساختاردهی شوند تا بتوان به سادگی به آنها دسترسی پیدا کرد، به صورت کارآ آنها را پردازش نمود، به سرعت بازیابی کرد و به صورت اثربخشی مدیریت نمود. ساختارهای داده‌ها و روش‌های دسترسی ساده، پیچیده زیادی برای سازماندهی و دسترسی کارآ به داده‌های ذخیره شده در سیستم اطلاعاتی وجود دارند. در این فصل ما این مفاهیم را در کنار کاربردهای مدیریتی و ارزش مدیریت منابع داده‌ها مورد بحث قرار می‌دهیم.

مفاهیم اساسی داده‌ها

قبل از این که جلوتر برویم، اجازه دهید برخی مفاهیم اساسی درباره چگونگی سازماندهی داده‌ها در سیستم‌های اطلاعاتی را مورد بحث قرار دهیم. در اینجا یک چارچوب مفهومی از سطوح مختلف داده‌ها بکار گرفته شده است که طبقه‌بندی‌های متفاوت و یا عناصر مختلف داده‌ها را متمایز می‌کند. بنابراین داده‌ها به صورت منطقی و به ترتیب در کاراکترها، فیلدها، رکوردها، فایل‌ها و پایگاه داده‌ها سازماندهی می‌گردند، درست همان‌طور که نوشته‌ها می‌توانند در حروف، واژه‌ها، جملات، بندها و اسناد سازماندهی شوند. مثال‌های این عناصر منطقی داده‌ها در شکل ۱-۵ نشان داده شده است.

کاراکتر

پایه‌ای‌ترین عنصر منطقی داده‌ها، کاراکتر است که از نمادهای الفبایی، عددی یا سایر نمادها تشکیل می‌شود. ممکن است تصور کنید که بیت یا بایت عناصر ابتدایی‌تری از داده‌ها هستند، ولی به یاد داشته باشید که این موارد به عناصر حافظه فیزیکی برمی‌گردد که در فصل ۳ درباره سخت‌افزار رایانه، مورد بحث قرار گرفت. از نقطه‌نظر یک کاربر (که نگرشی منطقی در مقابل نگرش فیزیکی یا سخت‌افزاری به داده‌ها دارد) یک کاراکتر پایه‌ای‌ترین عنصر داده‌ها است که می‌تواند مشاهده و یا دست‌کاری شود.



شکل ۱-۵ مثال‌هایی از عناصر منطقی داده‌ها در سیستم‌های اطلاعاتی. به مثال‌های چگونگی ارتباط فیلدها، رکوردها، فایل‌ها و پایگاه داده‌ها توجه کنید.

فیلد

سطح بعدی داده‌ها **فیلد** یا «عنصر داده» است. یک فیلد از گروهی از کاراکترهای مرتبط تشکیل می‌شود. مثلاً گروهی از کاراکترهای الفبایی در اسم شخص ممکن است یک فیلد نام ایجاد کند (یا معمولاً فیلد نام، نام خانوادگی)، و گروهی از اعداد در مقدار فروش فیلد مقدار فروش را تشکیل می‌دهد. یک فیلد داده‌ها نشان دهنده یک ویژگی^۱ (یک خاصیت یا کیفیت) از یک موجودیت^۲ (شیء، شخص، مکان، یا رویداد) است. برای مثال دستمزد یک کارمند یک ویژگی است که در فیلد داده‌های عمومی بکار می‌رود تا موجودیت یک کارمند سازمان را توصیف نماید.

رکورد

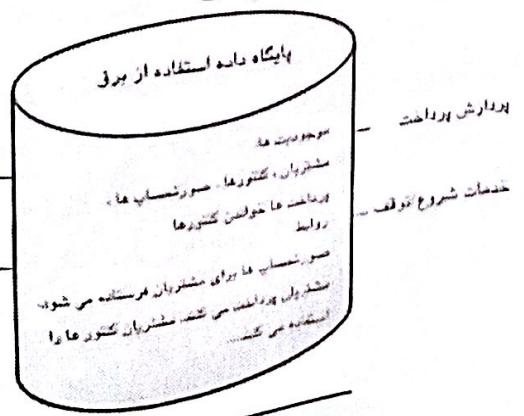
فیلدهای مرتبط داده‌ها، به شکل یک رکورد دسته‌بندی می‌شوند. بنابراین یک رکورد بیانگر ویژگی‌هایی است که یک موجودیت را توصیف می‌کنند. مثال آن رکورد لیست پرداخت شخص است، که از فیلدهای داده‌ها تشکیل شده است و ویژگی‌هایی از قبیل نام شخص، شماره تأمین اجتماعی، و نرخ پرداخت وی را بیان می‌کند. رکوردهای **با طول ثابت**^۳ شامل تعداد مشخصی از فیلدهای داده با طول ثابت هستند. رکوردهای **با طول متغیر**^۴ شامل تعدادی متغیر از فیلدها با طول فیلدهای متغیر هستند.

فایل

گروهی از رکوردهای مرتبط با هم فایل را تشکیل می‌دهد (که گاهی به آن جدول یا فایل تخت هم می‌گویند). وقتی یک جدول منفرد را مستقل از سایر فایل‌های مرتبط با آن در نظر بگیریم، به آن فایل تخت می‌گوییم. تعریف دقیق واژه فایل تخت می‌تواند هم محدود و هم گسترده باشد. یعنی یک پایگاه داده فایل تخت، باید فقط از داده‌ها و مسیر آن تشکیل شده باشد. در معنای گسترده‌تر، این عبارت به پایگاه داده‌هایی اطلاق می‌شود که فقط از یک فایل استون و ردیف‌هایش تشکیل شده باشد و به جز ساختار خود جدول، هیچ رابطه‌ای بین رکوردها وجود نداشته باشد. هر گروهی از رکوردها چه به شکل جدول باشد یا ردیف و ستون، فایل نامیده می‌شود. یعنی فایل کارکنان ممکن است شامل رکوردهایی درباره کارکنان یک شرکت باشد. فایل‌ها معمولاً بر مبنای کاربرد اولیه‌شان طبقه‌بندی می‌شوند از قبیل **فایل حقوق و دستمزد** یا **فایل موجودی کالا**، یا بر مبنای نوع داده‌های خود از قبیل **فایل سند** یا **فایل گرافیکی**. همچنین فایل‌ها بر مبنای ماندگاری‌شان مثلاً یک **فایل اصلی حقوق و دستمزد** در مقابل **فایل مبادلات** هفتگی حقوق و دستمزد طبقه‌بندی می‌شوند: یک فایل مبادلات، ممکن است شامل رکوردهای تمام مبادلاتی باشد که در طول یک دوره رخ داده‌اند و ممکن است به صورت دوره‌ای برای به روز کردن رکوردهای ماندگار یک فایل اصلی استفاده شوند. یک **فایل تاریخچه** نوعی فایل اصلی یا فایل مبادلات منسوخ شده است که برای اهداف پشتیبانی یا برای ذخیره‌سازی طولانی مدت بکار می‌رود که **ذخیره‌سازی بایگانی** نامیده می‌شود.

پایگاه داده

یک پایگاه داده^۱ مجموعه‌ای یکپارچه از عناصر داده‌های منطقی مرتبط است. یک پایگاه داده رکوردهای ذخیره شده در فایل‌های جداگانه را در یک ظرف مشترک داده‌ها بهم پیوند می‌دهد تا داده‌ها را برای استفاده‌های مختلف فراهم کند. داده‌های ذخیره شده در یک پایگاه داده، از برنامه کاربردی که آنها را بکار می‌گیرد و نوع ابزار حافظه که روی آن ذخیره شده‌اند مستقل هستند. بنابراین، پایگاه داده شامل عناصر داده‌هایی است که موجودیت‌ها و روابط میان موجودیت‌ها را توصیف می‌کنند. برای مثال، شکل ۲-۵ برخی روابط و موجودیت‌های درون یک پایگاه داده را برای یک کاربرد الکتریکی مشخص می‌کند. همچنین برخی کاربردهای کسب و کار (تهیه صورتحساب‌ها، پردازش پرداخت‌ها) که به دسترسی به عناصر داده‌های این پایگاه داده نیاز دارند را مشخص می‌کند.



شکل ۲-۵ برخی موجودیت‌ها و روابط درون یک پایگاه داده ساده شده استفاده از برق. به کاربردهای دسترسی به داده‌ها در پایگاه داده‌ها در کسب و کار توجه کنید.

ساختارهای پایگاه داده‌ها

رابط میان عناصر منفرد داده‌های ذخیره شده در پایگاه داده‌ها، بر یکی از چندین ساختار یا الگوی منطقی داده‌ها متنی می‌باشد. بسته‌های سیستم مدیریت پایگاه داده (DBMS) برای بکارگیری یک ساختار داده‌ای خاص و فراهم کردن دسترسی سریع و آسان به اطلاعات ذخیره شده در پایگاه داده‌ها طراحی شده‌اند. پنج ساختار اساسی پایگاه داده الگوهای سلسله مراتبی، شبکه‌ای، رابطه‌ای، شیء‌گرا و چندبعدی هستند. سه ساختار اولیه پایگاه داده‌ها در شکل ۳-۵ نشان داده شده است.

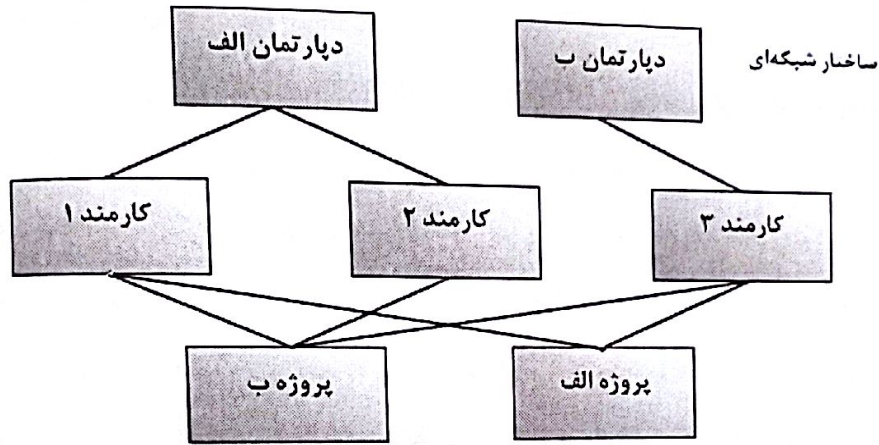
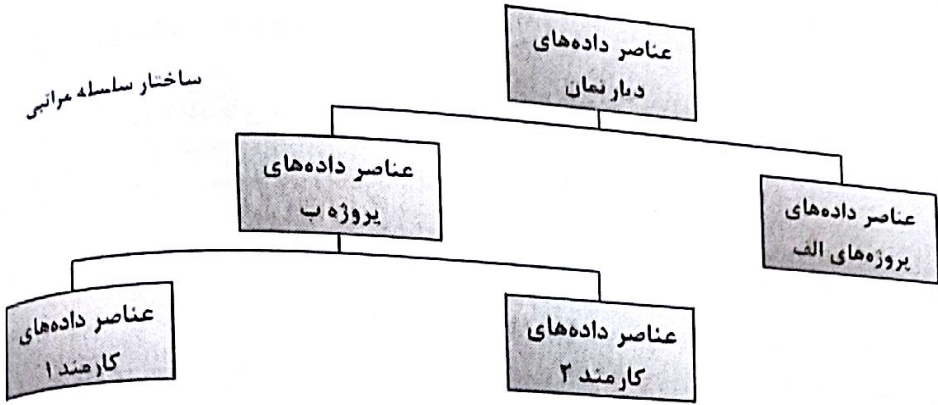
ساختار سلسله مراتبی

بسته‌های اولیه سیستم مدیریت پایگاه داده (DBMS) رایانه‌های بزرگ از ساختار سلسله مراتبی^۱ استفاده می‌کردند، که در آن روابط میان رکوردها، یک ساختار سلسله مراتبی یا درختی شکل را ایجاد می‌کرد. در الگوی قدیمی سلسله مراتبی، تمام رکوردها بهم وابسته هستند و در ساختارهایی چند سطحی مرتب می‌شوند، طوری که یکی از آنها رکورد ریشه‌ای است و بقیه در سطوح زیر مجموعه آن قرار می‌گیرند. بنابراین تمام روابط میان رکوردها یک به چند^۲ است، زیرا هر عنصر داده تنها به یک عنصر بالاتر از خود مربوط است. عنصر یا رکورد داده‌ها در بالاترین سطح سلسله مراتب (عنصر داده‌های دپارتمان در این شکل) عنصر ریشه‌ای نامیده می‌شود. برای دسترسی به هر عنصر داده باید از یک ریشه و در طول شاخه‌های درخت حرکت نمود رکورد مورد نظر (مثلاً، عنصر داده‌های کارمند) مکان‌یابی شود.

ساختار شبکه‌ای

ساختار شبکه‌ای^۳ می‌تواند روابط منطقی پیچیده‌تری را میان داده‌ها بیان کند، و هنوز هم توسط بسیاری از بسته‌های سیستم مدیریت پایگاه داده (DBMS) در رایانه‌های بزرگ مورد استفاده قرار می‌گیرد. این ساختار روابط چند به چند^۴ را میان رکوردها ممکن می‌کند. یعنی الگوی شبکه‌ای می‌تواند از طریق دنبال کردن چندین مسیر مختلف به یک عنصر داده دسترسی پیدا کند، زیرا هر عنصر یا رکورد داده می‌تواند با هر تعداد عناصر دیگر مرتبط باشد. مثلاً، در شکل ۳-۵ رکوردهای دپارتمان‌ها می‌توانند به بیش از یک رکورد کارمند، و رکوردهای کارمند می‌توانند به بیش از یک رکورد پروژه مرتبط شوند. از این رو، شخص می‌تواند تمام رکوردهای کارمندان برای یک دپارتمان خاص، یا تمام رکوردهای پروژه مرتبط با یک کارمند خاص را مکان‌یابی کند.

¹ hierarchical structure
² one to many
³ network structure



ساختر رابطه‌ای

جدول کارمند

شماره بخش	عنوان کارمند	نام کارمند	شماره کارمند
بخش A			کارمند ۱
بخش A			کارمند ۲
بخش B			کارمند ۳
بخش B			کارمند ۴
بخش C			کارمند ۵
بخش B			کارمند ۶

جدول بخش

شماره بخش	نام بخش	محل بخش	مدیر بخش
بخش A			
بخش B			
بخش C			

شکل ۳-۵. نمونه سه ساختار اصلی پایگاه داده‌ها. این ساختارها سه روش اصلی توسعه و بیان روابط میان عناصر داده‌ها در یک پایگاه داده هستند.

ساختار رابطه‌ای

الگوی رابطه‌ای^۱ پرکاربردترین ساختار در میان این سه ساختار پایگاه داده‌ها است. این الگو در بسیاری از بسته‌های DBMS در ریزرایانه‌ها و همچنین در اغلب سیستم‌های متوسط و بزرگ به کار می‌رود. در الگوی رابطه‌ای، تمام عناصر داده‌های درون پایگاه داده به عنوان عناصر ذخیره‌شده به شکل **جدول** ساده دویبعدی در نظر گرفته می‌شوند.

relational structure

که گاهی به آنها روابط هم گفته می‌شود. جداول پایگاه داده رابطه‌ای، جداولی تخت با ستون و ردیف هستند. هر ردیف بیانگر یک رکورد خاص در فایل و هر ستون بیانگر فیلد است. تفاوت اصلی میان فایل تخت و پایگاه داده آن است که داده‌های فایل تخت فقط می‌توانند به همان یک فایل نسبت داده شوند. ولی پایگاه داده می‌تواند به طور همزمان نسبت داده‌ها را با چند فایل مختلف مشخص کند و عناصر داده‌ای مختلف یک فایل را به یک یا چند فایل دیگر ربط دهد.

شکل ۳-۵ الگوی پایگاه داده رابطه‌ای را نشان می‌دهد که دپارتمان‌ها و کارمندان را توسط دو جدول روابط میان رکوردهای بیان می‌کند. سایر جداول یا روابط، برای این پایگاه داده سازمان، ممکن است روابط میان عناصر داده‌های پروژه‌ها، بخش‌ها، خطوط تولید و از این قبیل را بیان کنند. بسته‌های DBMS مبتنی بر الگوی رابطه‌ای برای فراهم کردن اطلاعات موردنیاز کاربران، می‌توانند عناصر داده‌ها از جداول مختلف را بهم پیوند بدهند. برای مثال یک بسته DBMS می‌تواند نام و حقوق یک کارمند را از جدول کارمند در شکل ۳-۵ و نام دپارتمان کارمند را از جدول دپارتمان با استفاده از فیلد مشترک شماره بخش Dept.no آنها بازیابی کند و این فیلد مشترک دو جدول را بهم پیوند می‌دهد. شکل ۴-۵ را ببینید.

عملیات رابطه‌ای

در یک پایگاه داده رابطه‌ای، سه عملیات عمده برای ایجاد مجموعه داده‌های مفید انجام می‌شود. عملیات گزینش برای ایجاد زیرمجموعه‌ای از رکوردها طبق معیار خواسته شده انجام می‌شود. مثلاً در یک پایگاه داده کارکنان ممکن است یک عملیات گزینش انجام شود تا زیرمجموعه‌ای از رکوردهای کارکنانی تهیه شود که درآمدی بالای ۳۰,۰۰۰ دلار در سال دارند و سابقه کارشان در سازمان بیشتر از سه سال است. همچنین می‌توان عملیات گزینش را به صورت تنظیم جدولی موقتی در نظر گرفت که ردیف‌های آن رکوردهایی منطبق با معیار خواسته شده دارند. عملیات پیوند برای ترکیب موقت دو یا چند جدول به کار می‌رود تا کاربر بتواند همه رکوردها را در یک جدول بزرگ‌تر ببیند. با این عملیات کاربر می‌تواند بدون این که وارد تک‌تک این جداول شود، داده‌های مورد نظرش را از چند فایل یا پایگاه داده متفاوت دریافت کند.

در نهایت، عملیات پروژه برای ایجاد زیرمجموعه‌ای از ستون‌های جداول موقتی عملیات گزینش و پیوند بکار می‌رود. همان طور که عملیات گزینش زیرمجموعه‌ای از رکوردهای منطبق با معیار خواسته شده را تنظیم می‌کند، عملیات پروژه نیز زیرمجموعه‌ای از ستون‌ها و فیلدهای مورد نظر کاربر ایجاد می‌نماید. کاربران با استفاده از عملیات پروژه می‌توانند ستون‌هایی را که برای پاسخ به یک پرسش یا تنظیم گزارش خاص لازم دارند انتخاب کنند و مجبور نیستند همه ستون‌های جدول را با هم بررسی کنند.

بخاطر استفاده گسترده از الگوهای رابطه‌ای، نرم‌افزارهای تجاری زیادی برای ایجاد و مدیریت آنها ارائه شده است. نرم‌افزارها برتر پایگاه داده رابطه‌ای برای رایانه‌های بزرگ Oracle 10g از شرکت Oracle، و DB2 از شرکت IBM هستند. یک نرم‌افزار رایج برای پایگاه داده رایانه‌های متوسط سرور SQL شرکت مایروسافت است. رایج‌ترین نرم‌افزار پایگاه داده برای رایانه‌های شخصی، برنامه Access شرکت مایروسافت است.

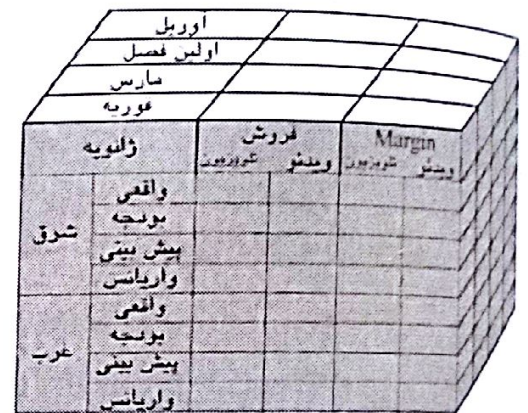
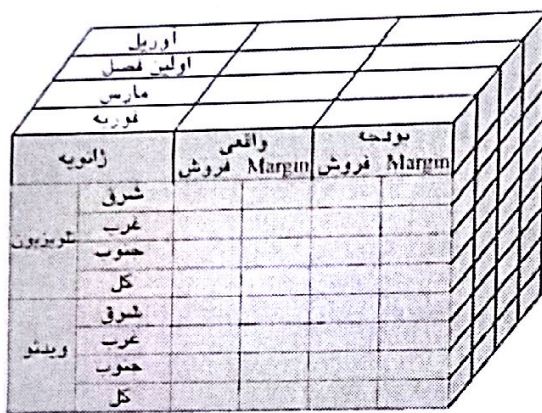
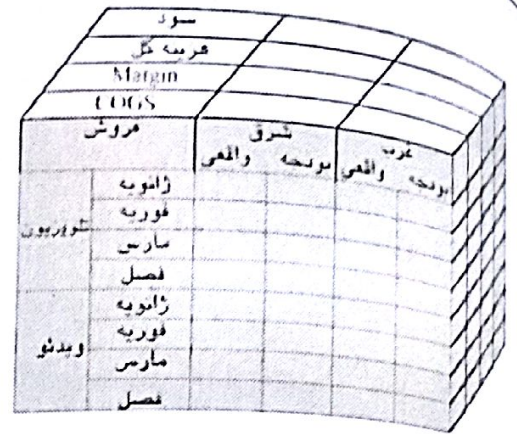
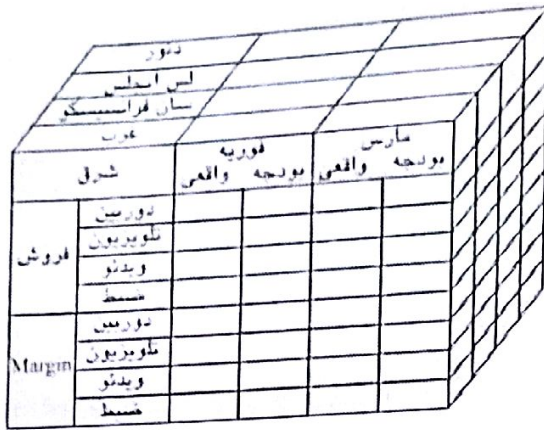
ساختار چند بعدی

ساختار چند بعدی پایگاه داده‌ها^۱ یکی از انواع الگوهای رابطه‌ای است که ساختارهای چندبعدی را برای سازماندهی داده‌ها و بیان روابط میان داده‌ها بکار می‌گیرد. می‌توانید ساختارهای چندبعدی را به عنوان مکعب‌هایی از داده‌ها و مکعب‌هایی درون مکعب‌های داده‌ها در نظر بگیرید. هر وجه مکعب به عنوان بعدی از داده‌ها در نظر گرفته می‌شود. شکل ۵-۵ مثالی است که نشان می‌دهد هر بعد داده‌ها می‌تواند بیانگر دسته متفاوتی از داده‌ها مانند نوع محصول، منطقه، کانال فروش، و زمان باشد.

هر سلول درون ساختار چند بعدی شامل داده‌های مترادفی است که با عناصر سایر ابعاد خود مرتبط هستند. برای مثال، یک سلول منفرد ممکن است شامل فروش کل برای یک محصول در یک منطقه و از طریق یک کانال فروش خاص در طول ماه باشد. یک مزیت عمده پایگاه داده‌های چندبعدی این است که آنها از روشی خلاصه و آسان برای درک و بکارگیری عناصر داده‌ها با روابط درونی متعدد برخوردارند. بنابراین پایگاه داده‌های چندبعدی به رایج‌ترین ساختار پایگاه داده‌ها برای پایگاه داده‌های تحلیلی تبدیل شده‌اند که از کاربردهای پردازش تحلیلی آن‌لاین^۲ (OLAP) پشتیبانی می‌کنند. در پردازش تحلیلی آن‌لاین پرس و جوهای پیچیده کسب و کار باید سریع پاسخ‌های داده شوند. ما کاربردهای پردازش تحلیلی آن‌لاین را در فصل ۹ مورد بحث قرار می‌دهیم.

جدول کارمند					جدول بخش			
شماره بخش	عنوان کارمند	نام کارمند	شماره کارمند	نام بخش	محل بخش	مدیر بخش	شماره بخش	
بخش A			کارمند ۱				بخش A	
بخش A			کارمند ۲				بخش B	
بخش B			کارمند ۳				بخش C	
بخش B			کارمند ۴					
بخش C			کارمند ۵					
بخش B			کارمند ۶					

شکل ۴-۵ پیوستن جدول کارمند و بخش در یک پایگاه داده‌های رابطه‌ای شما را قادر می‌سازد تا به داده‌های موجود در هر دو جدول بصورت همزمان دسترسی گزینشی داشته باشید.



شکل ۵-۵ مثالی از ابعاد مختلف یک پایگاه داده‌های چندبعدی

ساختار شیء‌گرا

الگوی پایگاه داده شیء‌گرا به عنوان یکی از فن‌آوری‌های کلیدی نسل جدید کاربردهای چندرسانه‌ای تحت وب در نظر گرفته می‌شود. همان‌طور که شکل ۵-۶ نشان می‌دهد یک شیء از ارزش‌های داده‌هایی تشکیل می‌شود که ویژگی‌های یک موجودیت را به همراه عملیاتی که می‌تواند روی داده‌ها انجام گیرد توصیف می‌کند. این قابلیت جمع‌بندی به الگوی شیء‌گرا اجازه می‌دهد تا بهتر از سایر ساختارهای پایگاه داده‌ها بتواند داده‌های پیچیده (گرافیکی، تصاویر، صدا، متن) از را اداره نماید.

علاوه بر این، الگوی شیء‌گرا از قابلیت **وراثت**^۱ برخوردار است؛ یعنی یک شیء جدید می‌تواند به صورت خودکار از طریق شبیه‌سازی برخی یا تمام خصوصیات یک یا چند شیء **والد** خلق شود. از این رو در شکل ۵-۶، شیء‌های حساب جاری و سپرده هر دو می‌توانند ویژگی‌ها و عملیات مشترکی را از شیء والد حساب بانکی به ارث ببرند. چنین قابلیت‌هایی باعث شده سیستم‌های مدیریت پایگاه داده شیء‌گرا^۲ (OODBMS) برای طراحی با کمک رایانه (CAD) و کاربردهای متعدد و فزاینده دیگر مناسب باشند. برای مثال، فن‌آوری شیء، طراحان را قادر می‌سازد

¹Inheritance

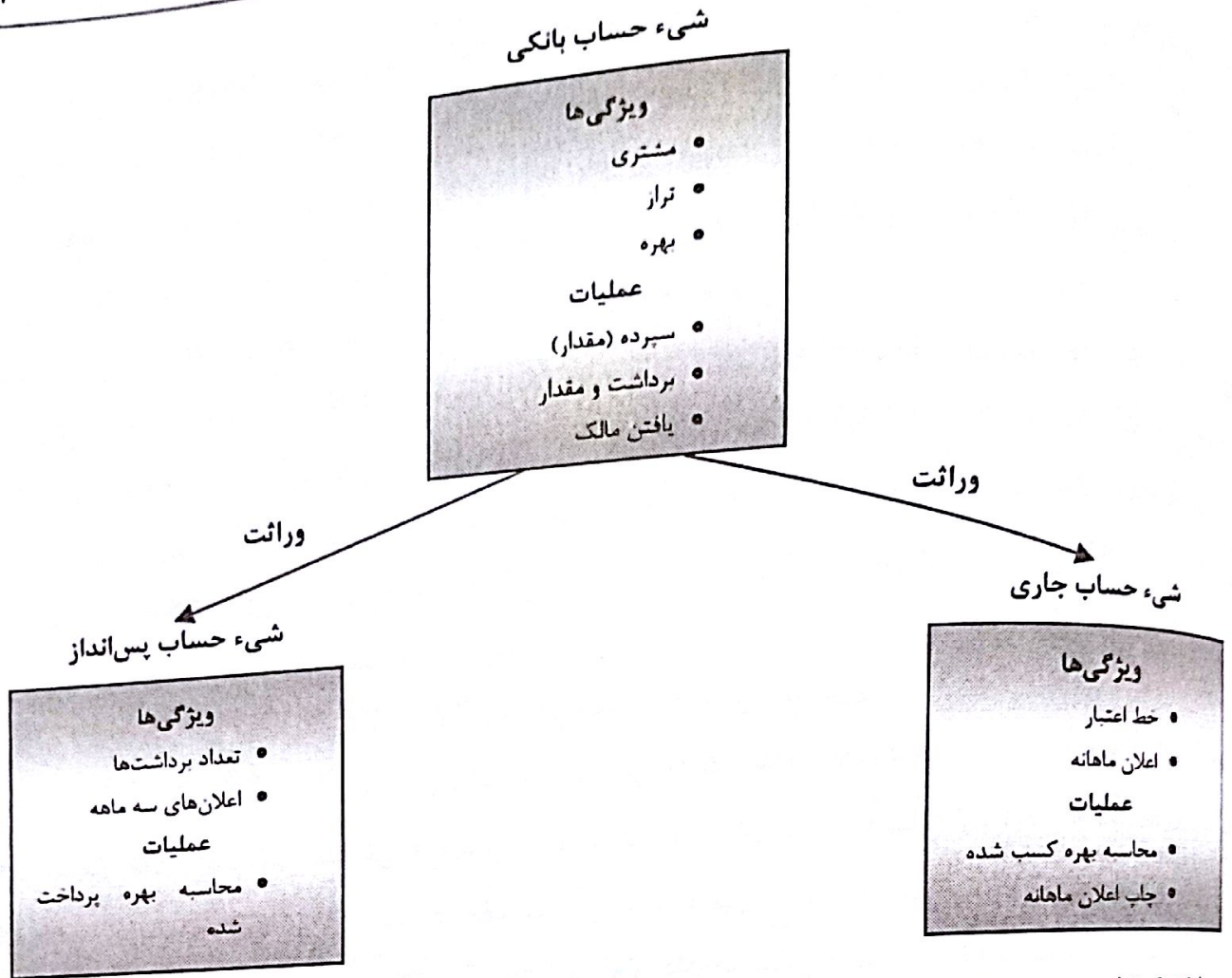
²object oriented database management system

طرح‌های محصولات را توسعه دهند، آنها را به عنوان اشیایی در پایگاه داده شیء‌گرا ذخیره نمایند، و آنها را تکرار، تعدیل کنند تا طرح‌های جدیدی برای محصولات ایجاد کنند. به علاوه، کاربردهای چندرسانه‌ای تحت وب، برای اینترنت و اینترنت‌ها و اکسترانت‌های شرکت به یک حوزه کاربردی عمده برای فن‌آوری شیء تبدیل شده است. طرفداران فن‌آوری شیء می‌گویند یک سیستم مدیریت پایگاه داده شیء‌گرا می‌تواند بسیار کارآتر از سیستم‌های مدیریت پایگاه داده رابطه‌ای با انواع داده‌های پیچیده^۱ از قبیل تصاویر گرافیکی و اسناد، کلیپ‌های ویدئویی، قطعات صوتی، و سایر زیرمجموعه‌های صفحات وب کار کند. البته، ارائه‌کنندگان اصلی سیستم مدیریت پایگاه داده رابطه‌ای نیز از طریق اضافه کردن بخش‌های شیء‌گرا به نرم‌افزار رابطه‌ای خود با آنها مقابله کرده‌اند. مثال‌های آن شامل اضافه کردن بخش شیء به محصول DB2 شرکت IBM، و کارتریج‌های مبتنی بر شیء شرکت Oracle برای محصول Oracle 10g می‌گردد.

ارزیابی ساختارهای پایگاه داده‌ها

ساختار سلسله‌مراتبی داده‌ها یک الگوی طبیعی برای پایگاه داده‌ها است که در اوایل دوران پردازش مبادلات و کاربرد رایانه برای پردازش مبادلات معمولی و ساختارمند استفاده می‌شد. داده‌های این عملیات می‌توانند به سادگی از طریق گروه‌هایی از رکوردها در رابطه‌ی سلسله‌مراتبی بیان شوند. هرچند، با پیشرفت زمان موارد فراوانی پیش می‌آمد که اطلاعات رکوردهایی مورد نیاز بود که روابط سلسله‌مراتبی نداشتند. برای مثال، در برخی سازمان‌ها، کارکنانی چند دپارتمان مختلف می‌توانند روی چند پروژه متفاوت (به شکل ۳-۵ مراجعه کنید) کار کنند. یک ساختار شبکه‌ای داده‌ها به سادگی این روابط چند به چند را اداره می‌نماید در حالی که الگوی سلسله‌مراتبی چنین قابلیت‌هایی ندارد. بنابراین ساختار شبکه‌ای بخاطر منعطف بودن‌شان رواج بیشتری در سازمان‌ها پیدا کردند. هرچند، از آن جایی که روابط ساختار شبکه‌ای نیز مانند ساختار سلسله‌مراتبی باید از قبل تعریف شوند، الگوی شبکه‌ای نمی‌تواند به سادگی تقاضاهای فوری و موردی را برای اطلاعات پاسخ گوید؛ این امر نیاز ما به الگوی رابطه‌ای را روشن می‌کند.

پایگاه داده‌های رابطه‌ای، از سوی دیگر به کاربران نهایی اجازه می‌دهند به سادگی پاسخ به تقاضاهای فوری‌شان را دریافت کنند. علت آن است که در هنگام ایجاد پایگاه داده‌های رابطه‌ای نیازی نیست که تمام روابط میان عناصر داده‌ها رابطه‌ای مشخص گردند. نرم‌افزارهای مدیریت پایگاه داده (از قبیل Oracle، BD2، Access و Approach) با استفاده از بخش‌هایی از جداول مختلف، جداولی جدید از روابط داده‌ها خلق می‌کنند. از این رو، کار با پایگاه داده‌های رابطه‌ای برای برنامه‌نویسان آسان‌تر است و نگهداری آنها نسبت به الگوهای سلسله‌مراتبی و شبکه‌ای ساده‌تر می‌باشد.



شکل ۵-۶ شیء حساب جاری و شیء پس‌انداز می‌توانند ویژگی‌ها و عملیات مشترکی را از شیء حساب بانکی به ارث ببرند.

محدودیت اصلی الگوی رابطه‌ای این است که سیستم‌های مدیریت پایگاه داده رابطه‌ای نمی‌توانند با سرعت و گزایی الگوهای شبکه‌ای یا سلسله‌مراتبی حجم زیاد پردازش مبادلات کسب و کار را انجام دهند، و یا کاربردهای پیچیده و با حجم بالا را مانند الگوهای شیء‌گرا پردازش نمایند. این شکاف عملکردی با توسعه نرم‌افزارهای پیشرفته DBMS رابطه‌ای از طریق اضافه کردن بخش‌های شیء‌گرا کمتر شده است. استفاده از نرم‌افزارهای الگوهای شیء‌گرا چندبعدی برای مدیریت پایگاه داده‌ها پیوسته در حال رشد است، و این فن‌آوری‌ها نقش بزرگ‌تری را در پردازش تحلیلی آن‌لاین (OLAP) و کاربردهای اینترنتی ایفاء می‌نمایند.

توسعه پایگاه داده‌ها

نرم‌افزارهای مدیریت پایگاه داده مانند Microsoft Access و Approach شرکت Lotus به کاربران نهایی اجازه می‌دهند به سادگی پایگاه داده‌های موردنیاز خود را ایجاد کنند. البته، سازمان‌های بزرگ معمولاً کنترل و توسعه پایگاه داده‌های سازمان را به مدیران پایگاه داده^۱ (DBA) و سایر متخصصین پایگاه داده‌ها می‌سپارند. این امر، اسجام و امنیت پایگاه داده‌های سازمانی را بهبود می‌دهد. توسعه‌دهندگان پایگاه داده‌ها از *واژه‌نامه تعریف داده‌ها*

(DDL) موجود در DBMSها نظیر Oracle 10g یا DB2 شرکت IBM استفاده می‌کنند تا محتواها، روابط، و ساختارهای داده‌ها را در هر پایگاه داده ایجاد کنند، مشخص نمایند، و هنگام ضرورت این مشخصات پایگاه داده را اصلاح کنند. این اطلاعات در یک پایگاه داده شامل داده‌هایی درباره این تعریف‌ها و مشخصات که **واژه‌نامه داده**، یا مخزن فوق‌داده نام دارد، ذخیره می‌شوند. این واژه‌نامه داده توسط نرم‌افزار مدیریت پایگاه داده، مدیریت و توسط مدیران پایگاه داده نگهداری می‌شود.

یک **واژه‌نامه داده** نوعی کاتالوگ مدیریت پایگاه داده یا راهنمای **فوق داده‌ها**^۱ یعنی داده‌هایی درباره داده‌ها می‌باشد. یک واژه‌نامه داده مدیریت پایگاه داده با استفاده از نرم‌افزارهای DBMS، فوق داده‌ها یعنی داده‌هایی درباره ساختار، عناصر داده‌ها، و سایر خصوصیات پایگاه داده‌های سازمان ارائه می‌کند. برای مثال یک واژه‌نامه داده شامل نام و توصیف همه انواع رکوردهای و روابط بین‌شان، اطلاعات لازم برای دسترسی کاربران و بکارگیری برنامه‌های کاربردی و نیز نگهداری و امنیت پایگاه داده می‌شود.

مدیر پایگاه داده می‌تواند برای ارزیابی وضعیت هر جنبه‌ای از فوق داده‌های یک شرکت، دیکشنری‌های داده‌ها را مورد پرس و جو قرار بدهد. سپس مدیر می‌تواند تغییراتی در تعاریف عناصر داده‌های مورد نظرش ایجاد کند. برخی دیکشنری‌های **فعال** (در مقابل **منفعل**) داده‌ها، هر گاه که کاربران نهایی و برنامه‌های کاربردی از سیستم مدیریت پایگاه داده برای دسترسی به پایگاه داده‌های یک سازمان استفاده می‌کنند، به صورت خودکار تعاریف استاندارد عناصر داده‌ها را اعمال می‌کنند. برای مثال، یک واژه‌نامه فعال داده‌ها ممکن است به یک برنامه ورود داده‌ها اجازه ندهد که تعریف غیراستانداردی برای رکورد یک مشتری وارد نماید، یا به یک کارمند اجازه ندهد که نام یک مشتری را که طولانی‌تر از اندازه عنصر داده‌های تعریف شده است وارد نماید.

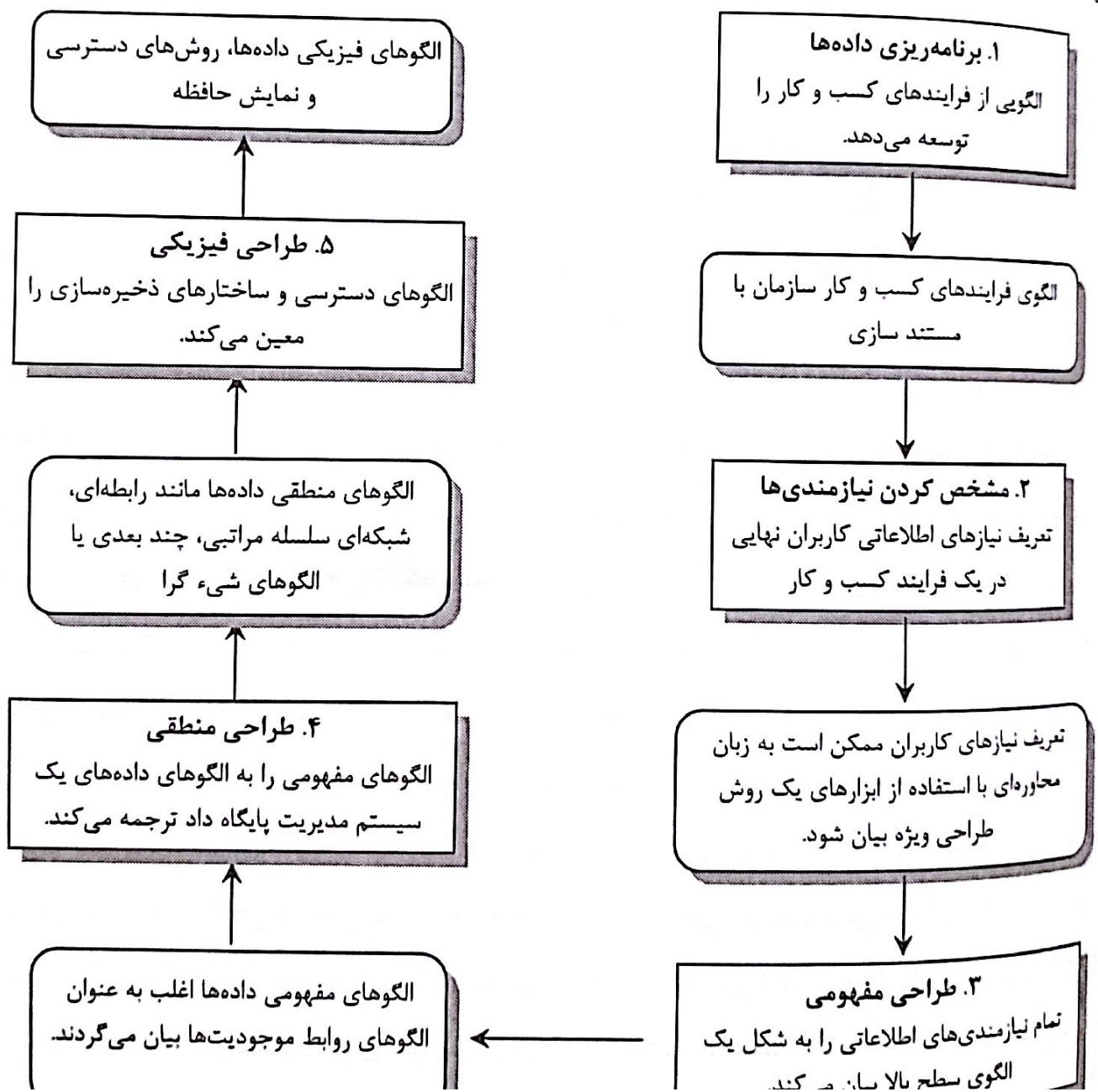
توسعه یک پایگاه داده شامل انواع داده‌های پیچیده می‌تواند کار دشواری باشد. مدیران پایگاه داده‌ها و تحلیل‌گران طرح پایگاه داده‌ها برای الگوسازی فرایندهای کسب و کار و داده‌های مورد نیاز، با کاربران نهایی و تحلیل‌گران سیستم‌ها همکاری می‌کنند. سپس آنها تعیین می‌کنند (۱) چه تعاریفی از داده‌ها باید در یک پایگاه داده باشد، و (۲) چه ساختار یا روابطی باید میان عناصر داده‌ها وجود داشته باشد.

برنامه‌ریزی داده‌ها و طراحی پایگاه داده‌ها

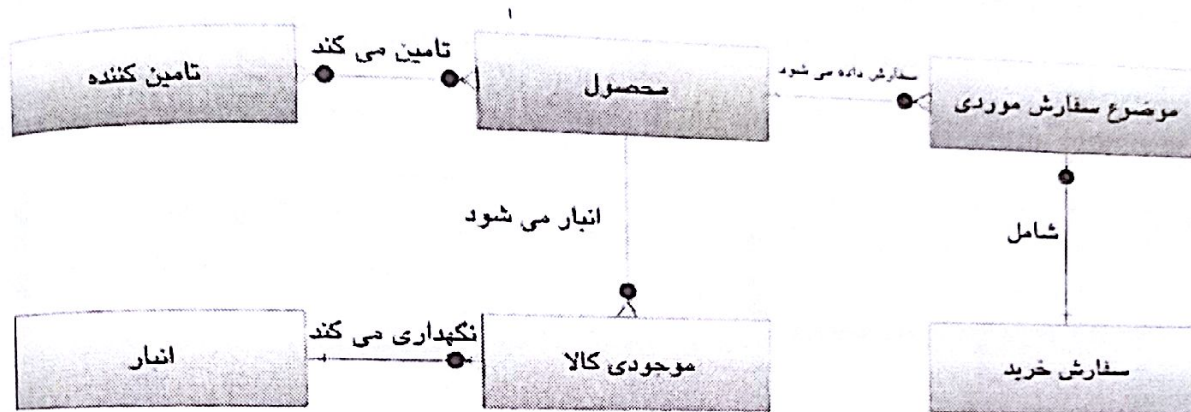
همان‌طور که شکل ۷-۵ نشان می‌دهد، توسعه پایگاه داده‌ها ممکن است با یک **فرایند برنامه‌ریزی داده‌ها**^۲ به صورت از بالا به پایین شروع شود. طراحان و مدیران پایگاه داده با مدیران شرکت و کاربران نهایی همکاری می‌کنند تا **الگوی سازمان** را تنظیم کنند و فرایند اصلی کسب و کار سازمان را تعریف نمایند. سپس آنها نیازهای اطلاعاتی کاربران نهایی در یک فرایند کسب و کار را مانند یک فرایند خرید/دریافت که تمام کسب و کارها دارند، تعریف می‌کنند.

۱ metadata
۲ data planning process

در ادامه کاربران نهایی باید عناصر کلیدی داده‌ها را که برای انجام فعالیت‌های خاص کسب و کار مورد نیاز هستند تعیین کنند. این کار معمولاً شامل توسعه نمودارهای رابطه موجودیت^۱ (ERD) است که روابط میان تمام موجودیت‌های فرایند کسب و کار را الگوسازی می‌کند. مثلاً شکل ۸-۵ برخی از این روابط را در فرایند خرید/دریافت مشخص می‌کند. کاربران نهایی و طراحان پایگاه داده می‌توانند از نرم‌افزارهای مدیریت پایگاه داده و یا نرم‌افزارهای الگوسازی کسب و کار استفاده کنند تا الگوهای ERD را برای فرایند خرید/دریافت توسعه دهند. این امر برای شناسایی عرضه‌کنندگان و محصولات مورد نیاز برای خودکارسازی خرید/دریافت و سایر فرایندهای کسب و کار مفید است و از طریق استفاده از برنامه‌های مدیریت منابع سازمان (ERP) یا نرم‌افزار مدیریت زنجیره تأمین (SCM) صورت می‌گیرد. وقتی درس تحلیل و طراحی سیستم را بگذرانید، جزئیات بیشتری درباره ERDها و سایر ابزارهای الگوسازی داده‌ها یاد می‌گیرید.



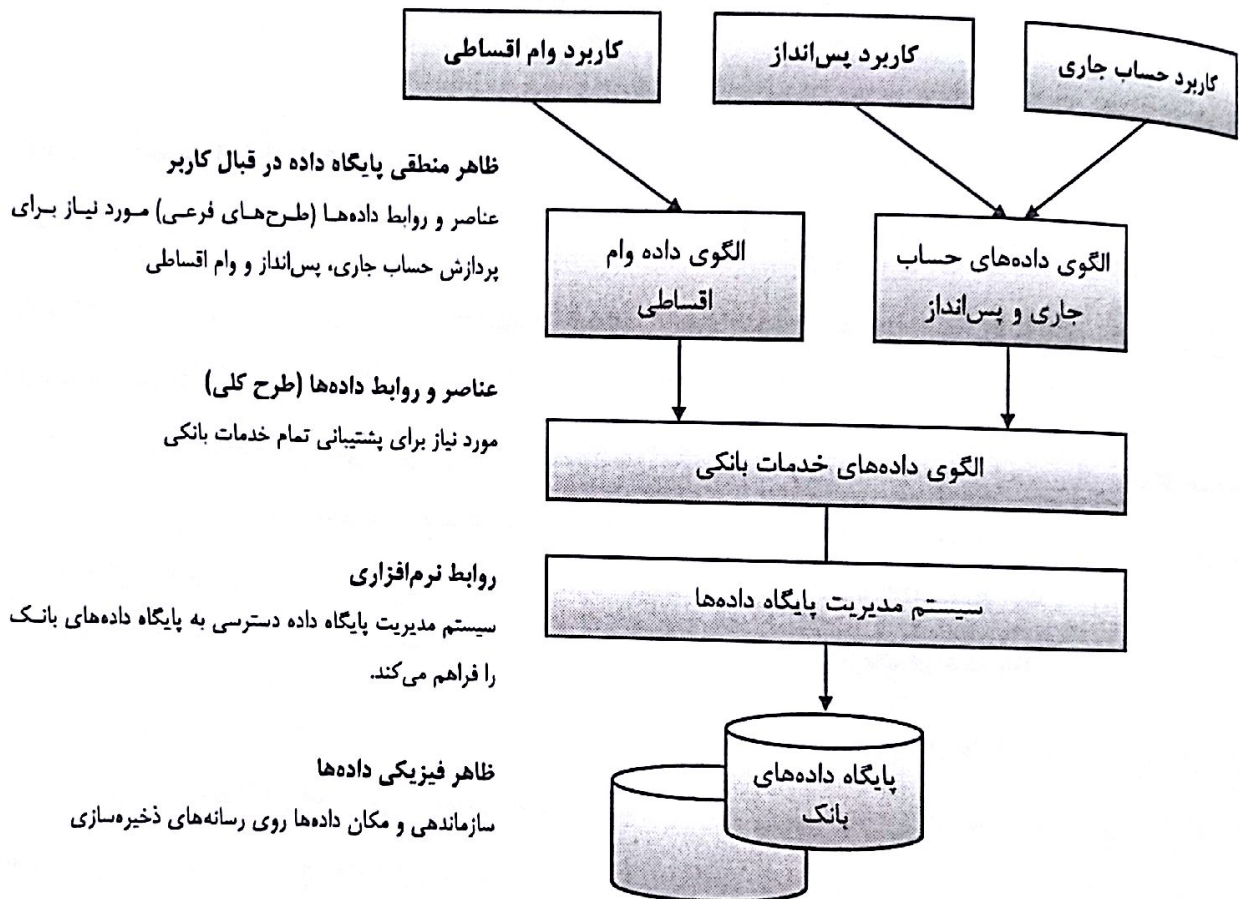
درک نگرش‌های کاربران بخش عمده‌ای از فرایند الگوسازی داده‌ها است که در آن روابط میان عناصر داده‌ها مشخص می‌شوند. هر الگوی داده‌ها روابط منطقی میان عناصر داده‌های مورد نیاز برای پشتیبانی از یک فرایند اصلی کسب و کار را تعریف می‌کند. برای مثال، آیا یک عرضه‌کننده می‌تواند بیش از یک نوع محصول برای ما فراهم کند؟ آیا یک مشتری می‌تواند بیش از یک نوع حساب با ما داشته باشد؟ آیا یک کارمند می‌تواند بر اساس کار در چندین گروه کاری پروژه‌های نرخ‌های پرداخت متفاوتی داشته باشد؟



شکل ۸-۵ این نمودار رابطه موجودیت، برخی روابط میان موجودیت‌ها (محصول، تأمین‌کننده، انبار، غیره) در یک فرایند خرید/دریافت را نشان می‌دهد.

روابط بین داده‌ها که در الگوی داده‌ها برای پشتیبانی از یک فرایند کسب و کار تعیین می‌شود، به این پرسش‌ها پاسخ می‌دهد. سپس این الگوهای داده‌ها به عنوان چارچوب‌های منطقی (به نام **طرح‌ها و زیرطرح‌های**) بکار می‌روند. این چارچوب‌ها طرح فیزیکی پایگاه داده‌ها و چگونگی توسعه برنامه‌های کاربردی برای پشتیبانی از فرایندهای کسب و کار سازمان را تعیین می‌کنند. یک طرح، یک دیدگاه کلی منطقی درباره روابط میان عناصر داده‌های درون یک پایگاه داده است، و یک زیرطرح یک دیدگاه منطقی از روابط داده‌ها است که برای پشتیبانی از برنامه‌های کاربردی خاص کاربران نهایی که به پایگاه داده‌ها دسترسی خواهند داشت، مورد نیاز است. به یاد داشته باشید که الگوهای داده‌ها دیدگاهی منطقی از داده‌ها و روابط آنها در پایگاه داده‌ها را ارائه می‌کنند.

طرح فیزیکی پایگاه داده نگاهی فیزیکی به داده‌ها دارد (که نگرش درونی نیز نامیده می‌شود) و چگونگی ذخیره‌سازی و دسترسی به داده‌ها را در ابزارهای حافظه یک سیستم رایانه‌ای توصیف می‌کند. مثلاً شکل ۹-۵ این نگرش‌های متفاوت به پایگاه داده‌ها و رابط نرم‌افزاری یک سیستم پردازش پایگاه داده بانک را نشان می‌دهد. در این مثال، حساب جاری، پس‌انداز، و وام اقساطی فرایندهای کسب و کاری هستند که الگوهای داده‌های آنها بخشی از الگوی داده‌های خدمات بانکی هستند و به عنوان یک چارچوب منطقی داده‌ها برای تمام خدمات بانکی بکار می‌روند.



شکل ۹-۵ مثالی از ظاهر فیزیکی و منطقی پایگاه داده و رابط نرم‌افزاری یک سیستم اطلاعاتی خدمات بانکداری

قسمت ۲

مدیریت منابع داده‌ها

مدیریت منابع داده‌ها

داده‌ها یکی از منابع حیاتی سازمان هستند که باید مانند سایر دارایی‌های مهم کسب و کار مدیریت شوند. سازمان‌های کسب و کار امروزی بدون داده‌های با کیفیت درباره عملیات داخلی و محیط خارجی خود قادر به بقا یا موفقیت نمی‌شوند.

با هر کلیک موس، داده‌های جدیدی خلق می‌شود و یا داده‌های ذخیره شده، از وب سایت‌های کسب و کار بازیابی می‌شوند. پلست تقاضاهای انبوه برای ابزارهای قدرتمند ذخیره‌سازی این داده‌ها می‌شود. محرک رشد آنها ضرورت شدیدی است که شرکت‌ها برای تحلیل هر بیت از اطلاعاتی که می‌توانند برای مزیت‌های رقابتی که از انبارهای عظیم داده‌های خود روشن می‌کنند دارند. این امر به کارکردهای ذخیره‌سازی و مدیریت داده‌ها نقشی استراتژیک در عصر اطلاعات داده است.

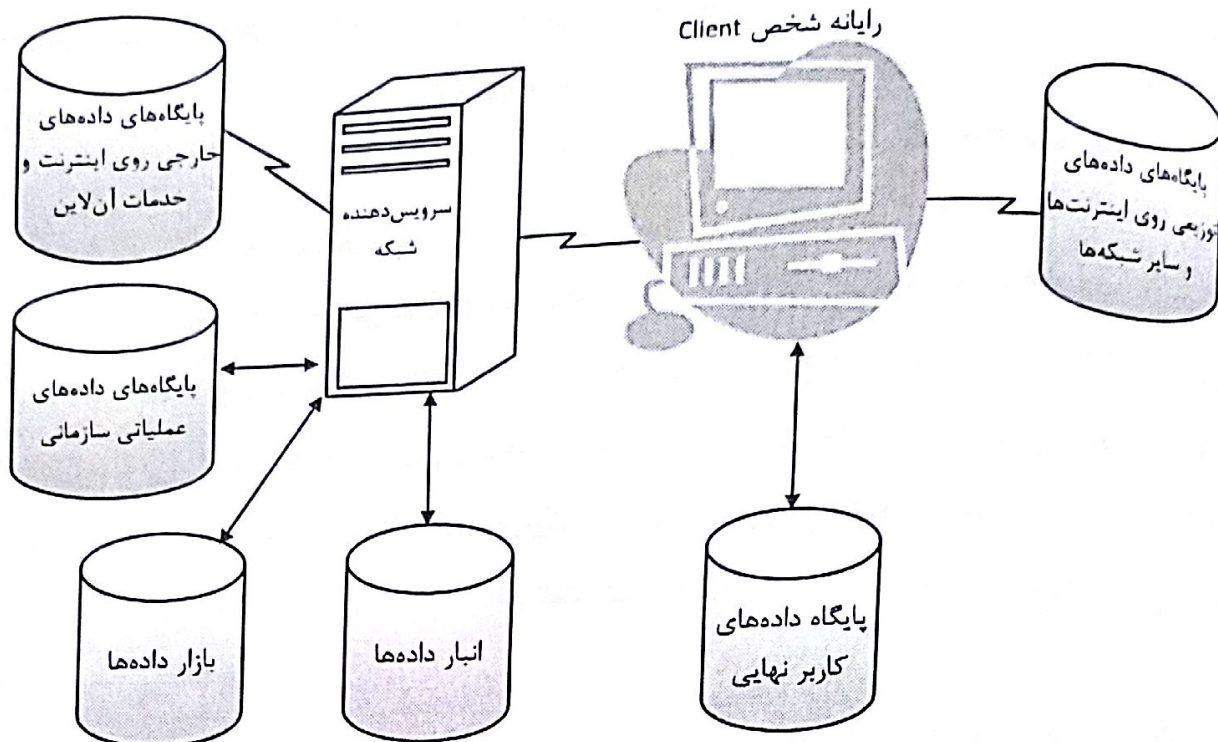
به این دلیل است که سازمان‌ها و مدیران‌شان نیازمند انجام مدیریت منابع داده‌ها هستند، فعالیتی مدیریتی که فن‌آوری‌های سیستم‌های اطلاعاتی از قبیل مدیریت پایگاه داده‌ها^۱، انبارداری داده‌ها^۲، و سایر ابزارهای مدیریت داده‌ها را برای مدیریت منابع داده‌های سازمان بکار می‌گیرد تا نیازهای اطلاعاتی ذی‌نفعان در کسب و کار را برآورده سازد. این فصل اقدامات مدیریتی مرتبط با کاربردهای فن‌آوری‌های مدیریت منابع داده‌ها و همچنین روش‌های مدیریت دارایی‌های داده‌ای یک سازمان را برای برطرف کردن نیازهای اطلاعاتی کسب و کار نشان خواهد داد.

انواع پایگاه داده‌ها

توسعه مستمر فن‌آوری اطلاعات و کاربردهای کسب و کاری آن منجر به ظهور چندین نوع عمده از پایگاه داده‌ها شده است. شکل ۱۰-۵ چندین دسته‌بندی مفهومی پایگاه داده‌ها را که ممکن است در بسیاری سازمان‌ها یافت شوند نشان داده است. اکنون اجازه دهید نگاهی مختصر به برخی از آنها داشته باشیم.

^۱ database management

^۲ data warehousing



شکل ۱۰-۵ مثال‌هایی از انواع اصلی پایگاه داده‌ها که توسط سازمان‌ها و کاربران نهایی مورد استفاده قرار می‌گیرند.

پایگاه داده‌های عملیاتی

پایگاه داده‌های عملیاتی^۱، جزئیات داده‌های مورد نیاز برای پشتیبانی از فرایندهای کسب و کار و عملیات شرکت را ذخیره می‌کنند. آنها پایگاه داده‌های موضوعی^۲ (SADB)، پایگاه داده‌های مبادله‌ای (TDB)، و پایگاه داده‌های تولید (PDB) نیز نامیده می‌شوند. مثال‌های آن عبارتند از پایگاه داده‌های مشتریان، پایگاه داده منابع انسانی، پایگاه داده موجودی کالا، و سایر پایگاه داده‌هایی که شامل داده‌های تولید شده در عملیات کسب و کار می‌باشند. برای مثال، یک پایگاه داده منابع انسانی شبیه آنچه در شکل ۱-۵ نشان داده شده است ممکن است شامل داده‌های مربوط به هر کارمند و زمان کار آنها، پرداخت‌ها، مزایا، ارزیابی‌های عملکرد، وضعیت آموزشی و توسعه‌ای، و سایر داده‌های مرتبط با منابع انسانی باشد. شکل ۱۰-۵ برخی از پایگاه داده عمومی عملیاتی که جهت استفاده شرکت‌های کوچک که با کمک نرم‌افزار مدیریت پایگاه Access ایجاد و مدیریت می‌شود، نشان می‌دهد.

پایگاه داده‌های توزیعی

بسیاری از سازمان‌ها کپی‌ها یا بخش‌هایی از پایگاه داده‌ها را روی سرورهای شبکه سایت‌های گوناگون تکرار و توزیع می‌کنند. این پایگاه داده‌های توزیعی^۳ می‌توانند روی سرورهای شبکه وب جهانی، روی اینترنت‌ها و اکسترانت‌های شرکت، یا روی سایر شبکه‌های شرکت قرار گیرند. پایگاه داده‌های توزیعی ممکن است کپی‌هایی از پایگاه داده‌های عملیاتی یا تحلیلی، پایگاه داده‌های مباحثه و یا فوق رسانه‌ای، یا هر نوع دیگری از پایگاه داده‌ها باشند. تکرار و توزیع

¹ operational database

² subject area database

³ distributed database

پایگاه داده‌ها عملکرد ایستگاه‌های کاری کاربران نهایی را بهبود می‌دهد. کسب اطمینان از این که داده‌های یک پایگاه داده توزیعی سازمان همواره و همزمان به زور می‌شوند یک چالش عمده مدیریت پایگاه داده‌های توزیعی است. شکل ۵-۵ مثال‌هایی از پایگاه داده‌های عملیاتی که می‌توانند توسط نرم‌افزارهای مدیریت پایگاه داده‌ها مانند مایکروسافت Access برای کسب و کارهای کوچک ساخته و مدیریت شوند.

یک مزیت دیگر پایگاه داده‌های توزیعی، الزامات ذخیره‌سازی آنها است. معمولاً یک پایگاه داده‌ی بزرگ، بر اساس روابط منطقی بین داده‌ها و مکان آنها در چند پایگاه داده کوچک‌تر توزیع می‌شود. مثلاً شرکتی با چند شاخه عملیاتی مختلف، ممکن است داده‌هایش را به گونه‌ای توزیع کند که شعبه هر شاخه عملیاتی با مکان پایگاه داده آن شاخه یکی باشد. با توجه به این که پایگاه داده‌های یک سیستم توزیعی می‌توانند به هم پیوند بخورند، هر شعبه کنترل داده‌های محلی خودش را در دست دارد ولی تمام شعبه‌های دیگر شرکت هم می‌توانند در صورت نیاز به تمام پایگاه داده‌ها دسترسی داشته باشند.

البته پایگاه داده‌های توزیعی بدون چالش هم نیستند. اولین چالش آنها حفظ صحت داده‌ها است. اگر شرکتی پایگاه داده‌هایش را در مکان‌های مختلف توزیع کند، هر تغییر داده‌ای در یک شعبه باید در تمام پایگاه داده‌های دیگر به‌روز شود. این به‌روزروی را با دو روش می‌توان انجام داد: **تعدیل** یا **تکرار**.

به‌روزروی پایگاه داده به روش **تعدیل** عبارت است از بکارگیری یک نرم‌افزار تخصصی که هر پایگاه داده را بررسی می‌کند و تغییرات آنها را شناسایی می‌کند. پس از این که تغییرات شناسایی شدند، فرایند تعدیل با اعمال تغییرات لازم در هر کدام کاری می‌کند که تمام این پایگاه داده‌های توزیعی شبیه هم شوند. فرایند تعدیل بسیار پیچیده است و بسته به تعداد و اندازه پایگاه داده‌های توزیعی ممکن است زمان و منابع رایانه‌ای زیادی نیاز داشته باشد.

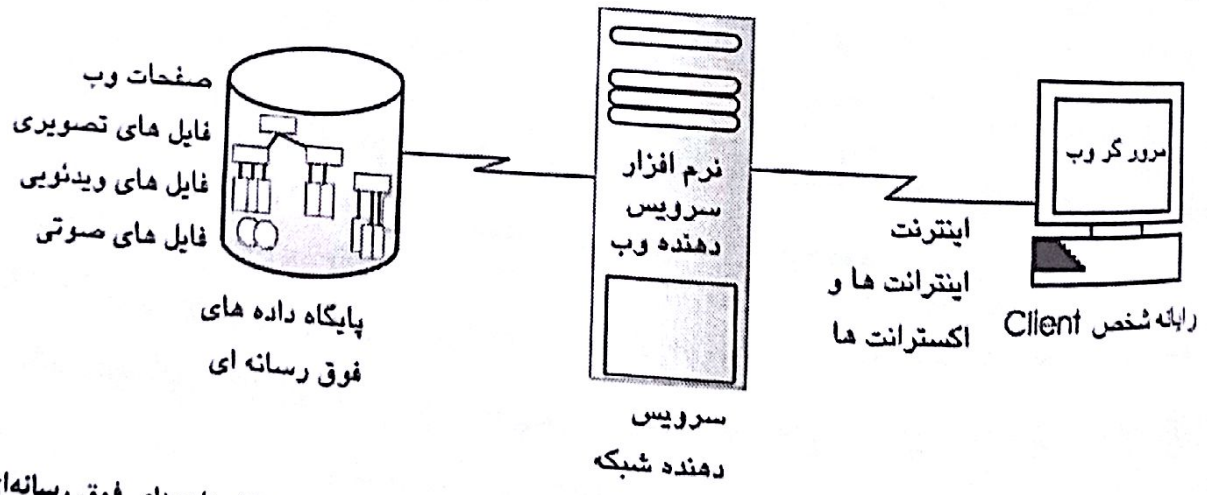
در مقابل، فرایند **تکرار** پیچیدگی کمتری دارد. در این فرایند یک پایگاه داده به عنوان مرجع شناسایی می‌شود و طی دوره‌های مشخص پایگاه داده مرجع را روی تمام سایر پایگاه داده‌ها تکرار می‌کند تا همه آنها شبیه هم شوند. تکرار تغییرات در داده‌ها باید در پایگاه‌های محلی بازنویسی شوند. به هر حال، استفاده مناسب از تعدیل و تکرار می‌تواند تمام پایگاه داده‌های توزیعی را به‌روز و دقیق نگه دارد.

یک چالش دیگر پایگاه داده‌های توزیعی نیاز آنها به قدرت رایانه‌ای و پهنای باند بیشتر برای دسترسی به بیشتر مورد بحث قرار می‌دهیم. مسأله پهنای باند را در فصل ۶ در قسمت ارتباطات راه‌دور و شبکه‌ها

پایگاه داده‌های خارجی به اطلاعات غنی پایگاه‌های داده‌های خارجی^۱ در مقابل کارمزد، از طرف شرکت‌های خدمات تجاری دسترسی به اطلاعات غنی و همچنین با هزینه و یا بدون هزینه از طرف منابع فراوان بر روی وب جهانی فراهم می‌گردد. وب سایت‌ها برای دستیابی از صفحات پیوند داده شده از اسناد چندرسانه‌ای را از طریق پایگاه داده‌های فوق‌رسانه‌ای^۲ در دسترس شما قرار می‌دهند. داده‌ها به شکل آمار اقتصادی و فعالیت‌های جمعیت‌شناسی از بانک داده‌های آماری به شما می‌شوند و یا شما می‌توانید چکیده یا کپی کامل صدها روزنامه، مجله، خبرنامه، مقاله‌های تحقیقاتی، و سایر مواد انتشار یافته و سایر گاهنامه‌ها را از پایگاه داده‌های کتاب‌شناسی و متن کامل ببینید و یا آنها را روی رایانه خود برماری کنید.

پایگاه داده‌های فوق رسانه‌ای

بسیار سریع وب سایت‌ها روی اینترنت، اینترنت‌ها و اکسترانت‌های شرکت، بکارگیری پایگاه داده‌هایی از اسناد فوق صوتی و فوق رسانه‌ای را بطور قابل ملاحظه‌ای، افزایش داده است. یک وب سایت این اطلاعات را در یک پایگاه داده‌ی فوق رسانه‌ای شامل صفحات چندرسانه‌ای پیوند داده شده (متن، تصویر، و عکس، کلیپ‌های ویدئویی، بخش‌های صوتی، و از این قبیل) ذخیره می‌کند. یعنی از دیدگاه مدیریت پایگاه داده‌ها، مجموعه‌ای صفحات چندرسانه‌ای متصل بهم روی یک وب سایت، بیشتر یک پایگاه داده متشکل از صفحات فوق‌رسانه‌ای بهم مرتبط است تا یک پایگاه داده‌های بهم مرتبط.



شکل ۱۱-۵ اجزای یک سیستم اطلاعاتی مبتنی بر وب شامل مرورگرهای وب، سرورها، و پایگاه داده‌های فوق رسانه‌ای.

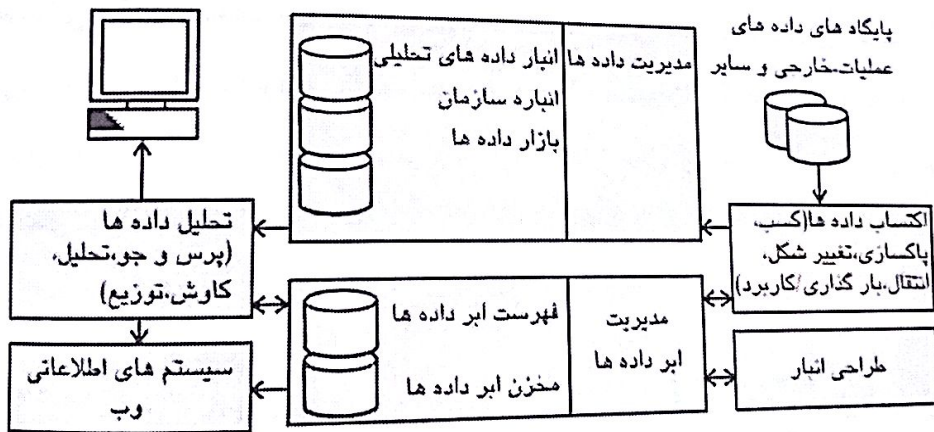
شکل ۱۱-۵ چگونگی استفاده شما را از مرورگرهای وب روی رایانه شخصی خود برای تماس با سرورهای شبکه وب نشان می‌دهد. این سرور نرم‌افزار سرور وب را اجرا می‌کند تا به صفحات وب مورد نظر شما دسترسی پیدا کند و آنها را برای شما منتقل کند. وب سایتی که در شکل ۱۱-۵ نشان داده شده است از یک پایگاه داده فوق رسانه‌ای

¹external database
Generated by CamScanner

استفاده می‌کند که از صفحات وب توصیف شده در کد زبان نشانه‌گذاری شده ابرمتن (HTML) یا برچسب‌های زبان نشانه‌گذاری شده توسعه یافته (XML)، فایل‌های تصویری، فایل‌های ویدئویی و صوتی تشکیل شده است. نرم‌افزار سرور وب، برای مدیریت و انتقال فایل‌های فوق‌رسانه‌ای برای بارگذاری توسط برنامه‌های اتصالی^۱ چند رسانه‌ای مرورگر وب شما مانند یک سیستم مدیریت پایگاه داده‌ها عمل می‌کند.

انبار داده و داده‌کاوی

یک **انبار داده**^۲ داده‌هایی را که از پایگاه داده‌های عملیاتی، خارجی و سایر پایگاه داده‌های سازمان کسب می‌شوند، ذخیره می‌نماید. انبار داده‌ها نوعی منبع مرکزی برای داده‌هایی است که پاک‌سازی شده، تغییر شکل داده و فهرست‌بندی شده‌اند تا توسط مدیران و سایر اهالی کسب و کار برای داده‌کاوی^۳، پردازش تحلیلی آن لاین، و سایر روش‌های تجزیه و تحلیل کسب و کار، تحقیقات بازار، و پشتیبانی تصمیم بکار روند. انبار داده‌ها ممکن است به **بازار داده‌ها**^۴ تقسیم شوند، یعنی زیرمجموعه‌هایی از داده‌های یک انبار را با در خصوص جنبه‌ای خاص از شرکت، مانند یک بخش یا یک فرایند کسب و کار.

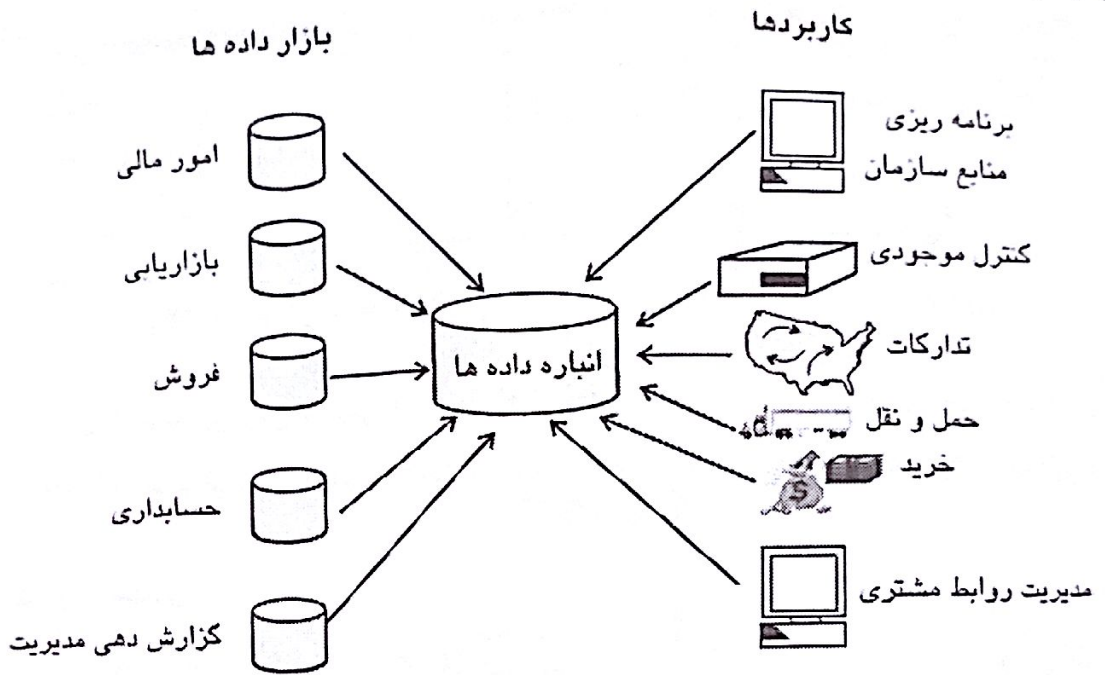


شکل ۱۲-۵ اجزای یک سیستم کامل انبار داده‌ها

شکل ۱۲-۵ اجزای یک سیستم کامل انبار داده‌ها را نشان می‌دهد. توجه کنید که چگونه داده‌ها از پایگاه داده‌های گوناگون عملیاتی و خارجی کسب، پاک‌سازی، و تبدیل به داده‌هایی می‌شوند که می‌توانند برای تجزیه و تحلیل بهتر بکار روند. این فرایند اکتساب ممکن است شامل فعالیت‌هایی از قبیل ترکیب داده‌ها از منابع مختلف فیلتر کردن داده‌های مطلوب، تصحیح داده‌های نادرست، تبدیل داده‌ها به عناصر جدید داده‌ها و جمع‌آوری داده‌ها در یک زیر مجموعه جدید داده‌ها باشد.

plug in
data warehouse
data mini-

سپس این داده‌ها در انبار داده‌های سازمان ذخیره می‌شوند، جایی که می‌توانند به بازار داده‌ها و یا به انبار داده‌های تحلیلی^۱ انتقال یابند که داده‌ها را به شکلی مفیدتر برای انواع بخصوص تجزیه و تحلیل، فوق داده‌ها (داده‌هایی که داده‌های داخل انبار داده‌ها را تعریف می‌کنند) در یک مخزن فوق داده‌ها توسط یک راهنمای فوق داده‌ها ذخیره و فهرست‌بندی می‌شوند. نهایتاً این که انواع نرم‌افزارهای تحلیلی می‌توانند برای پرس و جو، گزارش‌گیری، کاوش، و تجزیه و تحلیل داده‌ها برای عرضه از طریق اینترنت و سیستم‌های وب اینترنت به کاربران نهایی کسب و کار ارائه شوند. شکل ۱۳-۵ را ببینید.



شکل ۱۳-۵ یک انبار داده‌ها و زیر مجموعه‌های آن داده‌هایی را نگهداری می‌کند که از پایگاه داده‌های عملیاتی گوناگون برای تجزیه و تحلیل، تحقیقات بازار، پشتیبانی تصمیم، و کاوش داده‌ها استخراج شده‌اند.

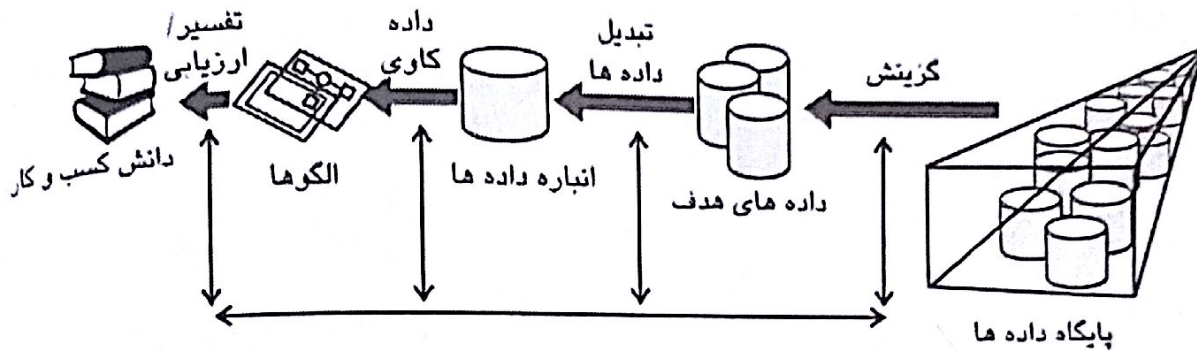
داده‌کاوی

داده‌کاوی یکی از کاربردهای اصلی پایگاه داده‌های انبار داده‌ها است. در داده‌کاوی، داده‌های درون انبار داده‌ها برای شناسایی الگوی پنهان در سوابق فعالیت‌های شرکت تحلیل می‌شوند. این تحلیل می‌تواند به مدیران کمک کند تا درباره تغییرات استراتژیک در عملیات کسب و کار برای بدست آوردن مزایای رقابتی در بازار تصمیم‌گیری کنند. شکل ۱۴-۵ را ببینید.

داده‌کاوی می‌تواند همبستگی‌ها، الگوها و روندهای جدید را در حجم عظیمی از داده‌های کسب و کار (اغلب چندین تریلیون از داده‌ها) کشف و در انبار داده‌ها ذخیره نماید. نرم‌افزار داده‌کاوی از الگوریتم‌های پیشرفته تشخیص الگو، مانند فنون ریاضی و آماری متنوع استفاده می‌کند تا کوهی از داده‌ها را بررسی و غربال کند و اطلاعات استراتژیک کسب و کار را که قبلاً ناشناخته بودند آشکار کند. مثلاً در بسیاری از شرکت‌ها داده‌کاوی برای مقاصد زیر بکار می‌رود:

¹ analytical data store

- انجام تجزیه و تحلیل سبد بازار را برای تعیین گروه محصولات جدید.
- کشف علل ریشه‌ای مسائل کیفیت یا مسائل تولید و ساخت
- جلوگیری از کاهش مشتری‌ها کنند و جذب مشتریان جدیدی
- فروش مجدد به مشتریان موجود.
- شناسایی و تشکیل پرونده مشتریان با دقت بیشتر



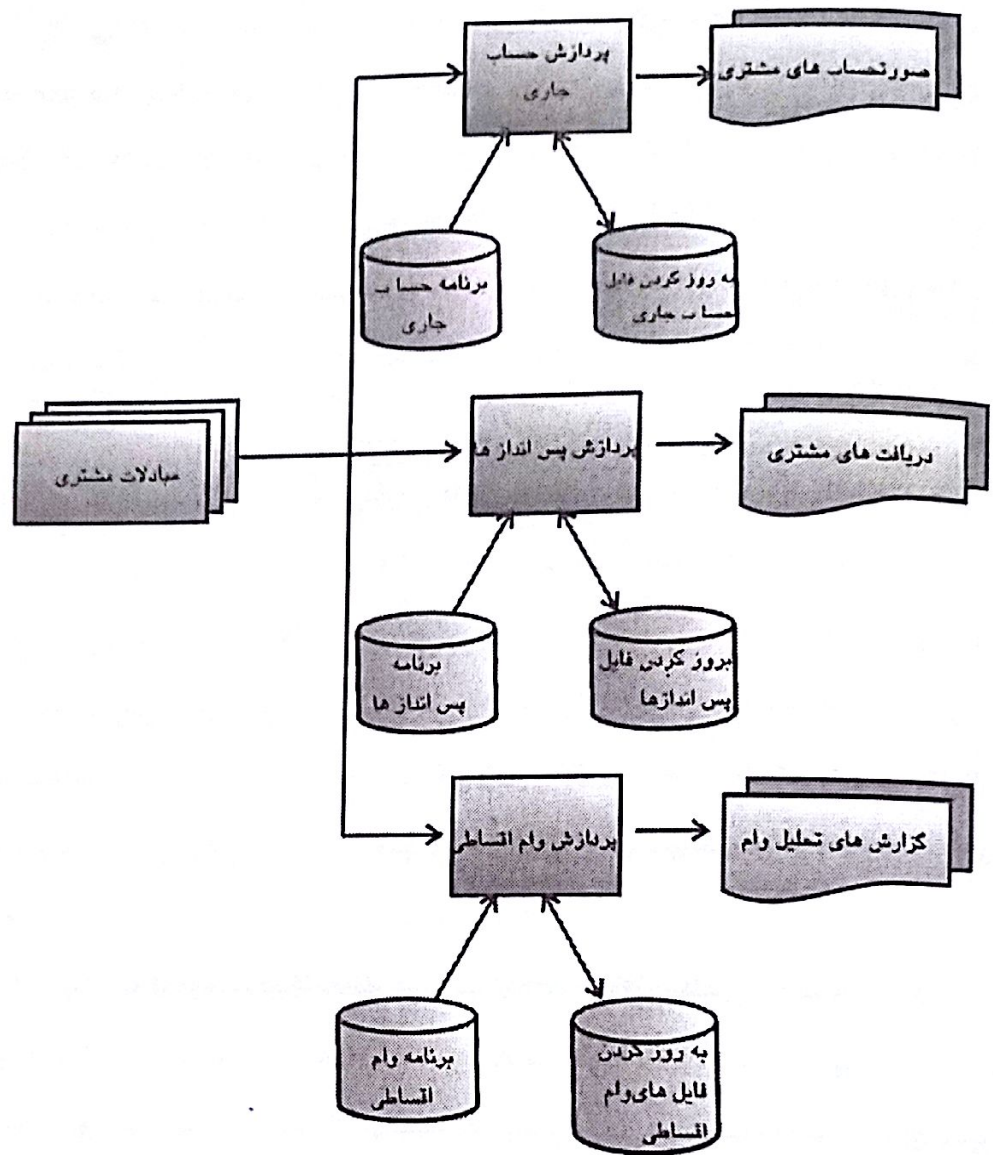
شکل ۱۴-۵ چگونگی استخراج دانش کسب و کار از یک انبار داده‌ها بوسیله داده کاوی

ما بعداً داده کاوی را، مانند پردازش تحلیلی آن لاین (OLAP)، و سایر فن‌آوری‌های تحلیل داده‌های درون پایگاه‌ها، انبارهای داده برای ارائه پشتیبانی‌های حیاتی از تصمیمات کسب و کار در فصل ۹ مورد بحث قرار خواهیم داد.

پردازش فایل سنتی

چه احساسی پیدا می‌کردید اگر مدیر اجرایی یک شرکت بودید و به شما گفته می‌شد بدست آوردن اطلاعات مورد نظرتان درباره کارکنان‌تان، بسیار دشوار و هزینه‌بر است؟ فرض کنید معاون خدمات اطلاعاتی دلایل زیر را برای شما بیان می‌کرد:

- اطلاعاتی که شما می‌خواهید در چندین فایل مختلف است و هر فایل به روشی متفاوت سازمان‌دهی شده است.
 - هر فایل برای استفاده در نرم‌افزار کاربردی متفاوتی تنظیم شده و هیچ یک از این نرم‌افزارها اطلاعاتی را که شما نیاز دارید تولید نخواهند کرد.
 - هیچ برنامه کاربردی برای کمک به شما در استخراج اطلاعات مورد نیاز از این فایل‌ها، موجود نیست.
- وقتی سازمان از سیستم‌های پردازش فایلی استفاده می‌کند که در آن داده‌ها در فایل‌هایی مستقل رکوردهای داده سازماندهی، ذخیره و پردازش می‌گردند کاربران نهایی به همین شکل دچار یأس و ناامنی می‌شوند. در رویکرد پردازش فایل سنتی که سال‌ها برای پردازش داده‌های کسب و کار استفاده می‌شد، هر نرم‌افزار کسب و کار به گونه‌ای طراحی شده بود که فقط یک یا دو فایل بخصوص از داده‌ها را بکار ببرد که فقط شامل انواع خاصی از رکوردهای داده‌ها می‌شد. برای مثال، یک نرم‌افزار برای پردازش حساب‌های جاری در بانک طراحی شده بود تا یک فایل داده‌ها شامل رکوردهای خاصی از داده‌ها را برای حساب‌های جاری مشتریان بانک به روز کند. همچنین، نرم‌افزار پردازش وام اقساطی بانک نیازمند دسترسی و به روز کردن فایل‌های خاص داده‌ها شامل رکوردهای داده‌هایی درباره مشتریان وام اقساطی بانک بود. شکل ۱۵-۵ را ببینید.



نکل ۵-۱۵ مثال هایی از سیستم های پردازش فایل در بانکداری. به نحوه استفاده از برنامه های رایانه ای جداگانه و فایل های مستقل داده ها برای کاربرد در حساب های پس انداز، وام اقساطی، و حساب جاری در یک رویکرد پردازش فایل توجه کنید.

مسائل پردازش فایل

پایته رویکرد پردازش فایل برای پشتیبانی از اطلاعات مورد نیاز برای مدیریت کسب و کارهای نوین بسیار پرزحمت، هزینه بر و انعطاف ناپذیر شد و همان طور که به زودی می بینید، با رویکرد مدیریت پایگاه داده جایگزین شد. سیستم های پردازش فایل مسائل و مشکلات عمده زیر را داشتند.

تکرار داده ها! فایل های مستقل داده ها مقدار زیادی داده های همسان داشتند. داده های یکسان (مانند نام و آدرس مشتری) در چندین فایل ثبت و ذخیره می شد. این **تکرار داده ها** هنگام به روز کردن اطلاعات باعث ایجاد مشکل می شدند زیرا برنامه های جداگانه **نگهداری فایل** باید ایجاد و هماهنگ می شدند تا اطمینان حاصل شود هر فایل به نحو مناسبی به روز شده است.

¹ data redundancy

فقدان یکپارچگی داده‌ها. نگه داشتن داده‌ها در فایل‌های مستقل، فراهم کردن اطلاعات فوری را که مستلزم دسترسی به داده‌های ذخیره شده در چندین فایل مختلف بود، برای کاربران نهایی بسیار دشوار می‌ساخت. برنامه‌های رایانه‌ای ویژه باید برای بازیابی داده‌ها از هر فایل مستقل نوشته می‌شدند. این کار برای بعضی سازمان‌ها بسیار دشوار، زمان‌بر و پرهزینه بود به طوری که فراهم کردن چنین اطلاعاتی برای کاربران نهایی یا مدیریت سازمان غیرممکن می‌شد. اگر اطلاعات واقعاً ضروری بودند کاربر نهایی می‌بایست به صورت دستی اطلاعات مورد نیاز را از گزارش‌های گوناگون تولید شده توسط برنامه‌های کاربردی جداگانه، استخراج می‌کرد و برای گزارش‌های مورد نظر مدیریت سازمان آماده می‌نمود.

عدم استقلال داده‌ها. در سیستم‌های پردازش فایل، اجزای اصلی یک سیستم - سازمان‌دهی فایل‌ها، مکان فیزیکی ذخیره‌سازی آنها روی سخت‌افزار، و نرم‌افزار کاربردی مورد استفاده برای دسترسی به فایل‌ها - به شدت به یکدیگر وابسته بودند. برای مثال، برنامه‌های کاربردی عموماً با یک قالب (فرمت) خاص ذخیره‌سازی داده‌های در فایل‌ها کار می‌کردند. بنابراین، تغییر در قالب و ساختار داده‌ها و رکوردها در یک فایل مستلزم تغییر تمام برنامه‌هایی بود که از این فایل‌ها استفاده می‌کردند. این زحمات نگهداری برنامه، یک بار بزرگ برای سیستم‌های پردازش فایل بود این امر سبب شد، سیستم پردازش فایل به هیچ وجه خوب کار نکند و همچنین منجر به ناسازگاری‌های بسیاری در فایل‌های داده‌ها شد.

فقدان درستی یا استانداردسازی داده‌ها. در سیستم‌های پردازش فایل، احتمال زیادی وجود داشت که عناصر داده‌هایی مانند تعداد سهام و آدرس‌های مشتریان توسط کاربران نهایی و در کاربردهای مختلف به شکل‌های متفاوتی تعریف شوند. این امر منجر به مشکلات و ناسازگاری‌های جدی در توسعه برنامه‌ها برای دسترسی به چنین داده‌هایی بود. به علاوه، درستی (صحت و کامل بودن) داده‌ها مورد ظن بود زیرا هیچ کنترلی بر استفاده و نگهداری آنها توسط کاربران نهایی مجاز وجود نداشت. از این رو، فقدان استانداردها مشکلات عمده‌ای در توسعه و نگهداری برنامه‌های کاربردی، و در امنیت و انسجام فایل‌های داده‌های مورد نیاز سازمان ایجاد می‌نمود.

رویکرد مدیریت پایگاه داده

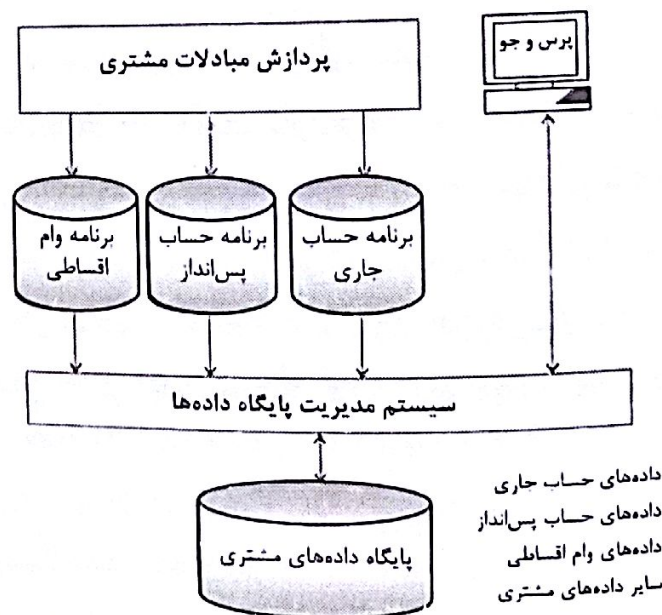
رویکرد مدیریت پایگاه داده به عنوان زیربنای روش‌های نوین مدیریت داده‌های سازمانی جایگزین رویکرد پردازش فایل شد. رویکرد مدیریت پایگاه داده، رکوردهای قبلی داده‌ها را که در فایل‌های جداگانه قرار داشتند در پایگاه داده‌ها یکپارچه می‌سازد به طوری که توسط بسیاری از برنامه‌های کاربردی قابل دسترسی باشند. علاوه بر این یک سیستم مدیریت پایگاه داده (DBMS) به عنوان یک رابط نرم‌افزاری میان کاربران و پایگاه داده‌ها عمل می‌کند و به کاربران در دسترسی ساده به داده‌های درون یک پایگاه داده کمک می‌کند. بنابراین، مدیریت پایگاه داده عبارت است از بکارگیری نرم‌افزارهای مدیریت پایگاه داده برای کنترل نحوه ایجاد، بازرسی و نگهداری پایگاه داده برای فراهم کردن اطلاعات مورد نیاز کاربران نهایی.

برای مثال، رکوردهای مشتریان و سایر انواع معمول داده‌ها برای چندین کاربرد مختلف در بانکداری مورد نیاز هستند از قبیل پردازش حساب جاری، سیستم‌های خودپرداز، کارت‌های اعتباری بانک، حساب‌های پس‌انداز، و حساب‌های وام اقساطی. این داده‌ها می‌توانند به جای این که در فایل‌های جداگانه برای هر یک از این کاربردها نگهداری شوند در یک پایگاه داده مشتری یکپارچه شوند. شکل ۱۶-۵ را ببینید.

نرم‌افزار مدیریت پایگاه داده

یک سیستم مدیریت پایگاه داده (DBMS) ابزار نرم‌افزاری اصلی در رویکرد مدیریت پایگاه داده است، زیرا ساخت، نگهداری و استفاده از پایگاه داده‌های سازمان و کاربران نهایی آن را کنترل می‌کند. همان‌طور که در شکل ۱۱-۵ دیدیم، نرم‌افزارهای DBMS برای ریزرایانه‌ها از قبیل Microsoft Access، Lotus Approach و Oracle Paradox و به شما این امکان را می‌دهند یک پایگاه داده را روی رایانه شخصی خود، روی سرور شبکه، و یا در وب جهانی مستقر نموده و مدیریت کنید. در سیستم‌های رایانه‌ای بزرگ و سرورها، DBMS یک بسته نرم‌افزاری سیستمی با اهمیت است که توسعه، استفاده، و نگهداری پایگاه داده‌ها را برای سازمان‌هایی که از رایانه استفاده می‌کنند، کنترل می‌نماید. از بینمای رایج نرم‌افزارهای DBMS برای رایانه‌های بزرگ و سرورها می‌توان Database Universal DB2 از شرکت IBM، Oracle 10g از شرکت Oracle و My SQL که یک DBMS متن باز و رایج است، را نام برد. اجزا و وظایف نسبی سیستم‌های مدیریت پایگاه داده‌ها در شکل ۱۷-۵ خلاصه شده‌اند.

به وظیفه عمده یک سیستم مدیریت پایگاه داده‌ها عبارتند از (۱) ایجاد و توسعه پایگاه داده‌های جدید و کاربردهای جدید برای پایگاه داده‌ها، (۲) حفظ کیفیت داده‌های درون پایگاه داده‌های یک سازمان، و (۳) بکارگیری پایگاه داده‌ها برای فراهم کردن اطلاعات مورد نیاز کاربران نهایی سازمان. شکل ۱۸-۵ را ببینید.



شکل ۱۶-۵ مالی از روش مدیریت پایگاه داده در سیستم اطلاعاتی بانکداری. به نحوه استفاده برنامه‌های حساب پس‌انداز، حساب جاری، و وام اقساطی از سیستم مدیریت پایگاه داده‌ها به مشارکت گذاشتن پایگاه داده مشتری توجه کنید. همچنین توجه کنید که سیستم مدیریت پایگاه داده‌ها به کاربر امکان می‌دهد بدون استفاده از برنامه‌های کاربردی جستجوی مستقیم و فوری در پایگاه داده انجام دهند.

ایجاد و توسعه پایگاه داده ها^۱ شامل تعریف و سازماندهی محتوا، روابط، و ساختار داده های مورد نیاز برای ساخت یک پایگاه داده است. **توسعه کاربردهای پایگاه داده ها^۲** شامل استفاده از DBMS برای توسعه الگوهای اولیه پرس و جوها، فرمها، گزارشها و صفحات وب است که برای یک کاربرد تجاری تنظیم شده‌اند. **نگهداری پایگاه داده ها^۳** شامل استفاده از سیستم‌های پردازش مبادلات و سایر ابزارها برای اضافه کردن، پاک کردن، به روزآوری و تصحیح داده‌های درون یک پایگاه داده است. کاربردهای اولیه یک پایگاه داده توسط کاربران نهایی شامل بکارگیری قابلیت‌های جستجو و بازرسی پایگاه داده^۴ یک DBMS برای دسترسی به داده‌های درون پایگاه داده و بازیابی و نمایش اطلاعات مورد نظر و تولید گزارشها، فرمها و سایر اسناد است.

اجزای معمول نرم‌افزار مدیریت سیستم پایگاه داده

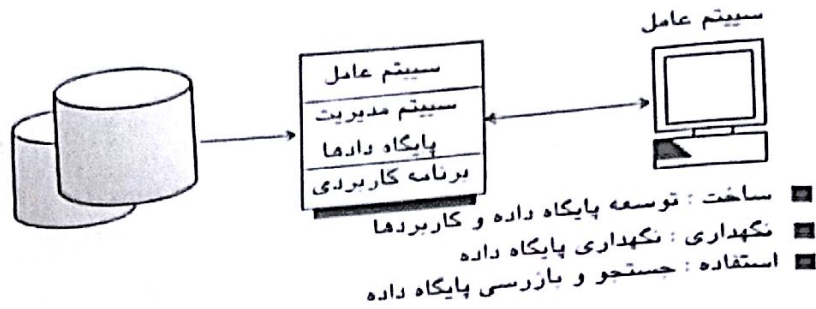
• تعریف پایگاه داده	زبان و ابزارهای گرافیکی برای تعریف موجودیتها، روابط، محدودیت‌های یکپارچگی، و حقوق مجوزها.
• دسترسی بی‌رویه	زبان و ابزارهای گرافیکی برای دسترسی به داده‌ها بدون کد دادن پیچیده.
• توسعه کاربردها	ابزارهای گرافیکی برای توسعه فهرستها، فرم‌های ورود داده‌ها و گزارشها.
• رابط زبان رویه‌ای	زبانی که دسترسی‌های بی‌رویه را با قابلیت‌های کامل یک زبان برنامه‌نویسی ترکیب می‌کند.
• پردازش مبادلات	سازوکارها را کنترل می‌کند تا از تداخل کاربران همزمان جلوگیری کند و داده‌های از دست رفته در اثر اشتباه را پوشش دهد.
• تنظیم پایگاه داده	ابزاری برای نظارت و بهبود عملکرد پایگاه داده

شکل ۱۷-۵ اجزا و کارکردهای نرم‌افزاری معمول در یک سیستم مدیریت پایگاه داده

جستجو و بازرسی پایگاه داده

قابلیت جستجو و بازرسی یک پایگاه داده مزیت عمده رویکرد مدیریت پایگاه داده است. کاربران نهایی می‌توانند از طریق درخواست اطلاعات از یک پایگاه داده با استفاده از ویژگی پرس و جو^۵ یا بکارگیری یک گزارش‌ساز^۶، از DBMS استفاده کنند. آنها می‌توانند پاسخی فوری به شکل نمایش‌های ویدئویی یا گزارش‌های چاپی دریافت نمایند. هیچ برنامه‌نویسی پیچیده‌ای هم مورد نیاز نیست. ویژگی زبان پرس و جو^۷ به شما اجازه می‌دهد به سادگی پاسخ‌های فوری به درخواست‌های موردی‌تان دریافت کنید. ویژگی گزارش‌ساز به شما امکان می‌دهد به سرعت یک قالب‌بندی گزارش برای اطلاعاتی که می‌خواهید به شکل گزارش ارائه شود، مشخص کنید.

- database development
- database application development
- database maintenance
- database interrogation
- query
- report generator
- query language



شکل ۱۸-۵ سه کاربرد اصلی نرم‌افزار مدیریت نرم‌افزار مدیریت پایگاه داده برای ساخت، نگهداری و استفاده پایگاه داده‌ها در یک سازمان

پرس و جوهای SQL^۱ یا زبان پرس و جوی ساختاریافته یک زبان پرس و جوی استاندارد بین‌المللی است که در بسیاری از نرم‌افزارهای DBMS یافت می‌شود. در اغلب موارد SQL ساختار زبانی است که برای سوال پرسیدن استفاده می‌شود و DBMS داده‌های لازم برای پاسخ به آن را بازیابی می‌کند. شکل اصلی پرس و جو با SQL به صورت زیر است:

انتخاب کن را از که

بعد از «انتخاب کن» فهرست فیلدهایی را که می‌خواهید بازیابی شوند وارد می‌کنید. بعد از «از» فهرست فایل‌ها و جدولی که می‌خواهید داده‌ها از آنها بازیابی شوند مشخص می‌کنید. بعد از «که» شرایط مورد نظران برای محدود کردن دامنه جستجو را تعیین می‌کنید تا فقط رکوردهای مورد نظر شما بازیابی شوند.

منطق بولی. یک کاربر برای استفاده از تمام قدرت SQL باید مفاهیم پشت **منطق بولی**^۲ را بشناسد. منطق بولی که توسط جرج بول در اواسط قرن هجدهم ایجاد شد ما را قادر می‌سازد تا جستجو برای اطلاعات خاص را بینه کنیم تا فقط داده‌های مورد نظرمان را بگیریم.

منطق بولی از سه عمل‌گر منطقی تشکیل شده است: (۱) و، (۲) یا، (نه). کاربر پایگاه داده با استفاده از این عملگرها در پرس و جوهای SQL می‌تواند جستجوی خود را بهینه کند تا اطمینان حاصل کند که فقط اطلاعات مورد نظرش بازیابی می‌شوند. این عمل‌گرهای منطقی را می‌توان در هنگام جستجو در اینترنت (یعنی بزرگ‌ترین پایگاه داده دنیا) نیز بکار برد. بیایید مثالی را درباره نحوه بکارگیری این عمل‌گرها بررسی کنیم.

فرض کنید می‌خواهیم در اینترنت اطلاعاتی درباره گربه‌ها بدست بیاوریم. می‌توانیم فقط واژه «گربه‌ها» را بنویسیم و احتمالاً تعدادی از سایت‌های بالقوه مفید بایابی خواهند شد. مسأله این جاست که علاوه بر وبسایت‌های مربوط به گربه‌ها، وبسایت‌های مربوط به گربه‌ها و سگ‌ها، و حیوانات خانگی هم نمایش داده خواهند شد و شاید حتی سایت مربوط به گروهی نمایشی به نام گربه‌ها هم بازیابی شود. برای این که مجبور نشوید برای یافتن مطلب مورد نظران همه این سایت‌ها را بگردید، می‌توانید از منطق بولی برای بهینه کردن جستجوی‌تان استفاده کنید و بنویسید:

گربه‌ها یا گربه‌سانان و نه سگ‌ها یا گروه نمایش

¹ structured query language

² Boolean logic

با استفاده از این روش جستجو می‌توانید وبسایت‌هایی کلمه گربه یا گربه‌سانان را دارند پیدا کنید و هر وبسایتی که کلمه سگ یا گروه نمایش را هم حذف کند با این رویکرد، می‌توانید از بازدید سایت‌های مربوط به سگ‌ها یا گروه‌های نمایشی اجتناب کنید این روش نتایج مفیدتری به شما می‌دهد و وقت‌تان را برای بازدید از سایت‌های نامرتبط تلف نمی‌کند.

پرس و جوهای گرافیکی و محاوره‌ای. بسیاری از کاربران نهایی (و متخصصین سیستم‌های اطلاعاتی) در عبارت بندی صحیح برای SQL و سایر زبان‌های پرس و جو پایگاه داده دچار مشکلاتی می‌شوند. بنابراین، اغلب نرم‌افزارهای مدیریت پایگاه داده ویژه کاربران نهایی روش‌های اشاره و کلیک از طریق رابط گرافیکی کاربر را ارائه می‌کنند، که بکارگیری‌شان ساده‌تر است و توسط نرم‌افزار به دستورات SQL ترجمه می‌شوند. بسته‌های نرم‌افزاری دیگری نیز عرضه می‌شوند که از عبارات پرس و جو به **زبان محاوره‌ای** استفاده می‌کنند که شبیه به محاوره به زبان انگیسی (با سایر زبان‌ها) است.

نگهداری پایگاه داده

فرایند **نگهداری پایگاه داده** توسط سیستم‌های پردازش مبادلات و سایر نرم‌افزارهای کاربران نهایی و با پشتیبانی DBMS انجام می‌گیرد. کاربران نهایی و متخصصین اطلاعات همچنین می‌توانند انواع برنامه‌های کمکی که توسط DBMS برای نگهداری پایگاه داده ارائه می‌شوند، بکار بگیرند. پایگاه داده‌های یک سازمان باید مرتباً به روز شوند تا مبادلات جدید کسب و کار (از قبیل فروش صورت گرفته، کالای تولید شده، یا موجودی کالای حمل شده) و سایر رویدادها را منعکس کنند. انواع تغییرات متفرقه دیگر نیز باید ایجاد شوند تا داده‌ها (مانند تغییرات آدرس و نام کارمند یا مشتری) را به روز و تصحیح کنند، و ما را از دقت و صحت داده‌های درون پایگاه داده‌ها مطمئن سازند. ما سیستم‌های پردازش مبادلات را در فصل یک معرفی کردیم و جزئیات بیشتر در مورد آنها را در فصل ۷ مورد بحث قرار می‌دهیم.

توسعه برنامه‌های کاربردی

بسته‌های نرم‌افزاری DBMS نقشی عمده‌ای نیز در توسعه برنامه‌های کاربردی ایفاء می‌کنند. کاربران نهایی، تحلیل‌گران سیستم، و سایر توسعه‌دهندگان کاربرد می‌توانند از زبان‌های داخلی برنامه‌نویسی نسل چهارم و ابزارهای توسعه نرم‌افزارها که از بسته‌های DBMS ارائه می‌شوند استفاده نمایند تا برنامه‌هایی با کاربردهای ویژه را توسعه دهند. برای مثال، شما می‌توانید به سادگی از یک نرم‌افزار DBMS برای تنظیم صفحه‌های ورود داده، فرم‌ها، گزارش‌ها، یا صفحات وب یک برنامه کسب و کار استفاده کنید طوری که خود برنامه با دسترسی به پایگاه داده‌های شرکت داده‌های مورد نیازش را بازیابی و به‌روزرسانی کند به علاوه یک DBMS، کار توسعه‌دهندگان نرم‌افزارهای کاربردی را تسهیل می‌کند، زیرا دیگر مجبور نیستند هر بار که برنامه‌ای می‌نویسند جزئیات رویه‌های کار با داده‌ها را با استفاده از زبان‌های برنامه‌نویسی قراردادی ایجاد کنند به جای آن، می‌توانند ویژگی‌هایی از قبیل عبارات زبان کار با داده‌ها^۱ (LMD) را در نرم‌افزار خود بکار ببرند که باعث می‌شود خود DBMS کارهای لازم را روی داده‌ها انجام دهد.

^۱ data manipulation language

کلامه

مدیریت منابع داده‌ها. مدیریت منابع داده‌ها یک فعالیت مدیریتی است که از فن‌آوری اطلاعات و ابزارهای نرم‌افزاری برای مدیریت منابع داده‌های یک سازمان استفاده می‌کند. در ابتدا برای مدیریت منابع داده‌ها از رویکرد پردازش فایل استفاده می‌شد که در آن داده‌ها تنها در فایل‌هایی خاص از رکوردهای داده‌ها سازماندهی می‌شدند و در دسترس قرار می‌گرفتند و برای پردازش توسط برنامه‌های خاص کاربردی در کسب و کار طراحی شده بودند. این رویکرد برای مدیریت فرایندها و سازمان‌های کسب و کار نوین بیش از حد پرزحمت، هزینه‌بر و انعطاف‌ناپذیر بود. از این رو، رویکرد مدیریت پایگاه داده‌ها توسعه داده شد تا مشکل سیستم‌های پردازش فایل را حل کند.

مدیریت پایگاه داده‌ها. رویکرد مدیریت پایگاه داده‌ها بر ذخیره‌سازی و پردازش داده‌ها تأثیر می‌گذارد. داده‌های مورد نیاز برنامه‌های مختلف، به جای آن که در چندین فایل داده‌های مستقل ذخیره شوند، در پایگاه داده‌های عمومی ترکیب و یکپارچه می‌شوند. همچنین، رویکرد مدیریت پایگاه داده‌ها بر به روزکردن و نگهداری پایگاه داده‌های معمول، به مشارکت گذاشتن داده‌های درون یک پایگاه داده توسط برنامه‌های کاربردی کاربران، فراهم کردن قابلیت گزارش‌دهی و پرس و جو/ پاسخ تأکید دارد تا کاربران نهایی بتوانند به سادگی گزارش‌ها و پاسخ‌های سریع به اطلاعات مورد نظرشان را دریافت نمایند.

نرم‌افزارهای پایگاه داده. سیستم‌های مدیریت پایگاه داده بسته‌های نرم‌افزاری هستند که ساخت، استفاده و نگهداری پایگاه‌های داده‌ها را تسهیل می‌کنند. این نرم‌افزارها ابزارهایی را ارائه می‌کنند تا کاربران نهایی، برنامه‌نویسان، و مدیران پایگاه‌های داده‌ها بتوانند پایگاه داده‌ها را خلق و اصلاح کنند، در پایگاه داده‌ها جستجو و بازرسی انجام دهند، گزارش تولید کنند، نرم‌افزارهای کاربردی ایجاد کنند و نگهداری پایگاه داده‌ها را انجام دهند.

انواع پایگاه داده‌ها. انواع گوناگونی از پایگاه داده‌ها توسط سازمان‌های کسب و کار مورد استفاده قرار می‌گیرند که عبارتند از پایگاه داده‌های عملیاتی، توزیعی، و خارجی. انبارهای داده‌ها، منبع مرکزی داده‌های سایر پایگاه داده‌ها هستند که پاک‌سازی، تبدیل و فهرست‌بندی شده‌اند تا هنگام تجزیه و تحلیل داده‌های کسب و کار و پشتیبانی از تصمیم‌گیری بکار گرفته شوند. این امر شامل داده کاوی، به معنای تلاش برای یافتن الگوها و روندها در انبار داده‌ها است. پایگاه داده‌های فوق‌رسانه‌ای در وب جهانی و اینترنت‌ها و اکسترانت‌های شرکت، صفحات چند رسانه‌ای فوق‌بهره‌ای در وب سایت‌ها را ذخیره می‌کنند. نرم‌افزارهای سرور وب می‌توانند این پایگاه داده‌ها را برای دسترسی سریع و نگهداری پایگاه داده‌های وب مدیریت کنند.

دستیابی به داده‌ها. داده‌ها باید به روشی منطقی روی ابزارهای فیزیکی حافظه سازماندهی شوند تا بتوانند به صورت کارآمد پردازش گردند. به این دلیل، داده‌ها اغلب تحت عناصری منطقی از داده‌ها از قبیل کاراکترها، فیلدها، رکوردها، فایل‌ها و پایگاه داده سازماندهی می‌شوند. ساختارهای پایگاه داده از قبیل الگوهای سلسله‌مراتبی، شبکه، رابطی، و شیء‌گرا برای سازمان دادن به روابط میان رکوردهای داده‌های ذخیره شده در پایگاه داده‌ها استفاده می‌شوند. پایگاه داده‌ها و فایل‌ها، می‌توانند هم به طور متوالی و هم مستقیم سازمان یابند و می‌توانند هم با روش‌های پردازش دستیابی مستقیم یا دستیابی ترتیبی در دسترس قرار گیرند و نگهداری شوند.

پیدا کنید و هر
های مربوط به
ا برای بازدید از

ی اطلاعاتی) در
بنسب این، اغلب
ی کاربر را ارائه
سای نرم‌افزاری
ه محاوره به زبان

و با پشتیبانی
کی که توسط
ه روز شوند تا
شده) و سایر
س و نام کارمند یا
س سیستم‌های
ر می‌دهیم

ایی، تحلیل‌گران
سعه نرم‌افزارها
ای مثال، شما
وب یک برنامه
ش را بازیابی و
را دیگر مجبور
سرارادادی ایجاد
ر ببرند که باعث

• توسعه پایگاه داده‌ها. توسعه پایگاه داده‌ها می‌تواند به سادگی از طریق بسته‌های مدیریت پایگاه داده‌های ریز رایانه‌ها و نرم‌افزارهای کاربردی کوچک کاربران نهایی انجام گیرد. البته، توسعه پایگاه داده‌های شرکت‌های بزرگ نیازمند تلاش در جهت برنامه‌ریزی از بالا به پایین داده‌ها است. این کار ممکن است شامل توسعه الگوهای روابط موجودیت‌ها و شرکت، پایگاه داده‌های موضوعی، و الگوهای داده‌ای باشد که عناصر و روابط منطقی داده‌ها را منعکس می‌کند و برای پشتیبانی عملیات و مدیریت فرایندهای اصلی کسب و کار سازمان مورد نیاز می‌باشد.