



معماری پایگاه‌های اطلاعاتی: با تأکید بر علم اطلاعات و دانش‌شناسی

دکتر فرامرز سهیلی مهندس علی رستم‌پور

«امروزه کتاب‌خوانی و علم‌آموزی، نه تنها یک وظیفه‌ی ملی، که یک واجب دینی است.»^۱

در عصر حاضر یکی از شاخصه‌های ارزیابی رشد، توسعه و پیشرفت فرهنگی هر کشوری میزان تولید کتاب، مطالعه و کتاب‌خوانی مردم آن مرز و بوم است. ایران اسلامی نیز از دیرباز تاکنون با داشتن تمدنی چندهزارساله و مراکز متعدد علمی، فرهنگی، کتابخانه‌های معتبر، علما و دانشمندان بزرگ با آثار ارزشمند تاریخی، سرآمد دولت‌ها و ملت‌های دیگر بوده و در عرصه‌ی فرهنگ و تمدن جهانی به‌سان خورشیدی تابناک همچنان می‌درخشد و با فرزندان نیک‌نهاد خویش هنرنمایی می‌کند. چه کسی است که در دنیا با دانشمندان فرزانه و نام‌آور ایرانی همچون ابوعلی سینا، ابوریحان بیرونی، فارابی، خوارزمی و ... همچنین شاعران برجسته‌ای نظیر فردوسی، سعدی، مولوی، حافظ و ... آشنا نباشد و در مقابل عظمت آنها سر تعظیم فرود نیاورد. تمامی این افتخارات ارزشمند، برگرفته از میزان عشق و علاقه فراوان ملت ما به فراگیری علم و دانش از طریق خواندن و مطالعه منابع و کتاب‌های گوناگون است. به شکرانه‌ی الهی، تاریخ و گذشته ما، همیشه درخشان و پربار است. ولی اکنون در این زمینه در چه جایگاهی قرار داریم؟ آمار و ارقام ارائه‌شده از سوی مجامع و سازمان‌های فرهنگی در مورد سرانه‌ی مطالعه‌ی هر ایرانی، برایمان چندان امیدوارکننده نمی‌باشد.

کتاب، دروازه‌ای به سوی گستره‌ی دانش و معرفت است و کتاب خوب، یکی از بهترین ابزارهای کمال بشری است. همه‌ی دستاوردهای بشر در سراسر عمر جهان، تا آنجا که قابل کتابت بوده است، در میان دست‌نوشته‌هایی است که انسان‌ها پدید آورده و می‌آورند. در این مجموعه‌ی بی‌نظیر، تعالیم الهی، درس‌های پیامبران به بشر، و همچنین علوم مختلفی است که سعادت بشر بدون آگاهی از آنها امکان‌پذیر نیست. کسی که با دنیای زیبا و زندگی‌بخش کتاب ارتباط ندارد بی‌شک از مهم‌ترین دستاورد انسانی و نیز از بیشترین معارف الهی و بشری محروم است. با این دیدگاه، به‌روشنی می‌توان ارزش و مفهوم رمزی عمیق در این حقیقت تاریخی را دریافت که اولین خطاب خداوند متعال به پیامبر گرامی اسلام (ص) این است که «بخوان!» و در اولین

۱. پیام مقام معظم رهبری به مناسبت آغاز هفته کتاب ۷۲/۱۰/۴

سوره‌ای که بر آن فرستاده‌ی عظیم‌الشان خداوند، فرود آمده، نام «قلم» به تجلیل یاد شده‌است: «اقْرَأْ وَ رَبُّكَ الْأَكْرَمُ. الَّذِي عَلَّمَ بِالْقَلَمِ» در اهمیت عنصر کتاب برای تکامل جامعه‌ی انسانی، همین بس که تمامی ادیان آسمانی و رجال بزرگ تاریخ بشری، از طریق کتاب جاودانه مانده‌اند.

دانشگاه پیام‌نور با گستره‌ی جغرافیایی ایران‌شمول خود با هدف آموزش برای همه، همه‌جا و همه‌وقت، به‌عنوان دانشگاهی کتاب‌محور در نظام آموزش عالی کشورمان، افتخار دارد جایگاه اندیشه‌سازی و خردورزی بخش عظیمی از جوانان جویای علم این مرز و بوم باشد. تلاش فراوانی در ایام طولانی فعالیت این دانشگاه انجام پذیرفته تا با بهره‌گیری از تجربه‌های گرانقدر استادان و صاحب‌نظران برجسته کشورمان، کتاب‌ها و منابع آموزشی درسی شاخص و خودآموز تولید شود. در آینده هم، این مهم با هدف ارتقای سطح علمی، روزآمدی و توجه بیشتر به نیازهای مخاطبان دانشگاه پیام‌نور با جدیت ادامه خواهد داشت. به‌طور قطع استفاده از نظرات استادان، صاحب‌نظران و دانشجویان محترم، ما را در انجام این وظیفه‌ی مهم و خطیر یاری‌رسان خواهد بود. پیشاپیش از تمامی عزیزانی که با نقد، تصحیح و پیشنهادهای خود ما را در انجام این وظیفه‌ی خطیر یاری می‌رسانند، سپاسگزاری می‌نماییم. لازم است از تمامی اندیشمندانی که تاکنون دانشگاه پیام‌نور را منزلگاه اندیشه‌سازی خود دانسته و ما را در تولید کتاب و محتوای آموزشی درسی یاری نموده‌اند، صمیمانه قدردانی گردد. موفقیت و بهروزی تمامی دانشجویان و دانش‌پژوهان عزیز آرزوی همیشگی ما است.

دانشگاه پیام‌نور

فهرست مطالب

پیشگفتار	یازده
فصل اول: مفاهیم پایه پایگاه داده‌ها	۱
هدف کلی	۱
هدف‌های یادگیری	۱
مقدمه	۲
۱-۱ پایگاه داده چیست؟	۲
۱-۱-۱ معرفی مفاهیم	۵
۲-۱ سیستم مدیریت پایگاه داده	۱۲
۳-۱ پیاده‌سازی پایگاه داده	۱۳
۱-۳-۱ رویکرد مبتنی بر فایل	۱۳
۲-۳-۱ رویکرد مبتنی بر پایگاه داده	۱۹
۴-۱ اهداف (مزایای) پایگاه داده	۲۲
۱-۴-۱ استقلال داده‌ها	۲۲
۲-۴-۱ اجتناب از ناسازگاری	۲۳
۳-۴-۱ اشتراک داده‌ها و کنترل همروندی	۲۳
۵-۱ عناصر اصلی و معماری پایگاه داده	۲۴
۱-۵-۱ عناصر تشکیل‌دهنده محیط پایگاه داده	۲۴
۶-۱ امنیت، و جامعیت تراکنش در پایگاه داده	۳۳
۱-۶-۱ قوانین تضمین جامعیت پایگاه داده	۳۴
۷-۱ تاریخچه ذخیره و بازیابی اطلاعات	۳۵
۱-۷-۱ دوره اول	۳۶
۲-۷-۱ دوره دوم	۳۶
۳-۷-۱ دوره سوم	۳۷

۳۷	خلاصه فصل اول
۳۸	خودآزمایی چهارگزینه‌ای فصل اول
۳۹	خودآزمایی تشریحی فصل اول
۴۱	فصل دوم: مدل‌سازی و نمودار ER
۴۱	هدف کلی
۴۱	هدف‌های یادگیری
۴۲	مقدمه
۴۲	۱-۲ مفاهیم طراحی مدل‌سازی
۴۳	۱-۱-۲ موجودیت
۴۴	۲-۱-۲ صفت
۴۸	۲-۲ ارتباط یا بستگی
۴۸	۳-۲ نمودار ER
۵۱	۱-۳-۲ درجه ارتباط
۵۲	۲-۳-۲ اتصال یا ماهیت نوع ارتباط
۵۳	۳-۳-۲ حد یا کاردینالیتی
۵۳	۴-۳-۲ صفت در ارتباط
۵۵	۵-۳-۲ وابستگی وجودی
۵۵	۶-۳-۲ اشتراک در صفت (ارث‌بری)
۵۹	خلاصه فصل دوم
۶۰	خودآزمایی چهارگزینه‌ای فصل دوم
۶۱	خودآزمایی تشریحی فصل دوم
۶۳	فصل سوم: ساختار داده - مدل داده رابطه‌ای
۶۳	هدف کلی
۶۳	هدف‌های یادگیری
۶۴	مقدمه
۶۴	۱-۳ ساختار رابطه‌ای
۶۵	۲-۳ مدل داده‌ای
۶۵	۱-۲-۳ مدل رابطه‌ای
۶۸	۲-۲-۳ انواع کلید در مدل رابطه‌ای
۷۱	۳-۳ قواعد جامعیت در مدل رابطه‌ای
۷۳	۴-۳ طراحی مقدماتی پایگاه جدولی (رابطه‌ای)
۸۲	خلاصه فصل سوم
۸۳	خودآزمایی چهارگزینه‌ای فصل سوم
۸۴	خودآزمایی تشریحی فصل سوم

۸۵	فصل چهارم: جبر رابطه‌ای
۸۵	هدف کلی
۸۵	هدف‌های یادگیری
۸۵	مقدمه
۸۹	۴-۱ بررسی رفتار عملگرها در مدل رابطه‌ای
۹۰	۴-۲ عملگرهای جبر رابطه‌ای
۹۱	۴-۲-۱ عملگرهای خاص
۹۷	۴-۲-۲ عملگرهای عام
۹۸	۴-۲-۳ عملگر اشتراک
۹۸	۴-۲-۴ عملگر تفاضل
۹۹	۴-۳ به‌روزرآوری داده‌ها
۱۰۲	خلاصه فصل چهارم
۱۰۲	خودآزمایی چهارگزینه‌ای فصل چهارم
۱۰۳	خودآزمایی تشریحی فصل چهارم
۱۰۵	فصل پنجم: SQL
۱۰۵	هدف کلی
۱۰۵	هدف‌های یادگیری
۱۰۶	مقدمه
۱۰۶	۵-۱ انواع داده در SQL
۱۰۶	۵-۱-۱ داده‌ی رشته‌ای
۱۰۷	۵-۱-۲ داده‌ی یونیکد
۱۰۸	۵-۱-۳ داده‌ی باینری
۱۰۹	۵-۱-۴ داده‌ی عددی
۱۱۰	۵-۱-۵ داده‌های تاریخ
۱۱۱	۵-۲ عملگرها در SQL
۱۱۲	۵-۲-۱ عملگرهای منطقی
۱۱۳	۵-۳ دستورات SQL
۱۱۳	۵-۳-۱ دستورات ایجاد پایگاه داده
۱۱۳	۵-۳-۲ تعریف دامنه جدید
۱۱۴	۵-۳-۳ دستور تغییر دامنه
۱۱۴	۵-۳-۴ دستور حذف دامنه
۱۱۴	۵-۳-۵ ایجاد بانک در SQL
۱۱۵	۵-۳-۶ دستور ایجاد جدول در بانک
۱۱۷	۵-۳-۷ دستور ایجاد تغییرات در جدول
۱۱۸	۵-۳-۸ دستور حذف جدول
۱۱۸	۵-۳-۹ دستور select

۱۲۱	۱۰-۳-۵ استفاده از عبارات اسکالر در دستور SELECT
۱۲۳	۱۱-۳-۵ عملگر between و in در select
۱۲۴	۱۲-۳-۵ عملگر like در select
۱۲۵	۱۳-۳-۵ جستجوی مقدار NULL
۱۲۵	۱۴-۳-۵ پرس وجوهای مبتنی بر پیوند جدول‌ها
۱۲۷	۱۵-۳-۵ دستورات Select متداخل
۱۲۸	۱۶-۳-۵ دستور وجودی Exists در Select
۱۲۹	۱۷-۳-۵ توابع ستونی (جمع‌ی) در select
۱۳۱	۱۸-۳-۵ دستورات Group By و Having در Select
۱۳۳	۱۹-۳-۵ دستور بازیابی با قید «همه»
۱۳۳	۲۰-۳-۵ عملگرهای مجموعه‌ای در Select
۱۳۴	۲۱-۳-۵ عملگرها ANY , ALL
۱۳۵	۲۲-۳-۵ دستور Insert
۱۳۷	۲۳-۳-۵ دستور Update
۱۳۸	۲۴-۳-۵ دستور Delete
۱۳۹	خلاصه فصل پنجم
۱۳۹	خودآزمایی چهارگزینه‌ای فصل پنجم
۱۴۱	خودآزمایی تشریحی فصل پنجم
۱۴۳	فصل ششم: مطالعه موردی در زبان SQL
۱۴۳	هدف کلی
۱۴۳	هدف‌های یادگیری
۱۴۳	مقدمه
۱۴۴	۱-۶ پایگاه داده کتابخانه
۱۴۵	۲-۶ شرح عملیات
۱۴۷	۱-۲-۶ ایجاد پایگاه و جدول و ارتباط
۱۶۸	۲-۲-۶ دستورات T-SQL
۱۷۷	خلاصه فصل ششم
۱۷۸	خودآزمایی چهارگزینه‌ای فصل ششم
۱۷۸	خودآزمایی تشریحی فصل ششم
۱۷۹	فصل هفتم: نرمال‌سازی
۱۷۹	هدف کلی
۱۷۹	هدف‌های یادگیری
۱۷۹	مقدمه
۱۸۰	۱-۷ وابستگی تابعی
۱۸۲	۱-۱-۷ وابستگی تابعی کامل

۱۸۲	۲-۱-۷ مجموعه پوشش وابستگی
۱۸۴	۳-۱-۷ مجموعه وابستگی بهینه و مجموعه وابستگی کاهش ناپذیر
۱۸۶	۴-۱-۷ نمودار وابستگی تابعی (FDD)
۱۸۶	۲-۷ نرمال سازی
۱۹۰	۱-۲-۷ سطوح نرمال
۱۹۷	خلاصه فصل هفتم
۱۹۸	خودآزمایی چهارگزینه‌ای فصل هفتم
۱۹۹	خودآزمایی تشریحی فصل هفتم
۲۰۱	پیوست
۲۱۵	پاسخنامه
۲۱۶	منابع

پیشگفتار

با گسترش علم و فناوری و تغییرات اساسی به وجود آمده در این زمینه، لزوم تغییر در سرفصل‌های دروس حوزه‌های مختلف علمی نیز به وجود آمد، این مهم در حوزه علم اطلاعات و دانش‌شناسی نیز رخ داد و سرفصل‌هایی با ماهیت بین‌رشته‌ای و متکی بر فناوری‌های اطلاعاتی و ارتباطی در دروس این حوزه قرار گرفت. یکی از آن‌ها معماری پایگاه‌های اطلاعاتی است. معماری پایگاه اطلاعاتی توصیفی از موقعیت و جایگاه همه اجزایی است که یک پایگاه داده را تشکیل می‌دهند و آن را عملیاتی می‌کنند. معمولاً در پایگاه داده این عملیات در سه سطح فیزیکی، ادراکی یا منطقی و خارجی انجام می‌گیرد. حضور و توسعه پایگاه‌های اطلاعاتی به عنوان محمل و منابع اطلاعاتی ارزشمند سبب شد تا کتابداران و دانشجویان این رشته نیز علاوه بر نحوه جستجو و بازیابی اطلاعات در این پایگاه‌ها، نیاز به نحوه طراحی، مدیریت و راه‌اندازی پایگاه‌های اطلاعاتی را احساس کنند و متن مناسبی که بتواند این نیاز را برآورده سازد و به طور خاصی به حوزه علم اطلاعات و دانش‌شناسی بپردازد وجود نداشت، و این انگیزه‌ای برای تدوین این اثر شد. در این کتاب تلاش شده است تا با مرور متون مرتبط با این حوزه و تجربه تدریس این درس در ترم‌های متوالی متنی مناسب برای این درس تدوین شود. برخی از مثال‌ها و متن‌ها از کتاب روحانی رانکوهی (۱۳۹۲) برگرفته شده است. محتوای این کتاب در قالب هفت فصل ارائه شده است. در فصل اول مفاهیم کلی مرتبط به داده، اطلاعات و پایگاه داده مورد بررسی قرار گرفته است. در فصل دوم تلاش شده است تا بحث مدل‌سازی و نمودار ER تشریح شود. فصل سوم کتاب مدل داده رابطه‌ای و در فصل چهارم جبر رابطه‌ای

مورد بحث و بررسی قرار گرفته است. در فصل پنجم به معرفی و نحوه کار با دستورات زبان SQL پرداخته شده و در فصل بعدی هم به صورت عملی نحوه کار با این برنامه در قالب مثال هایی از پایگاه داده کتابخانه تشریح شده است. در فصل آخر نیز نحوه نرمال سازی در پایگاه داده ها مورد بحث قرار گرفته است. نویسندگان از زحمات خانم مهندس نسیم حسنی که در تهیه برخی از تصاویر و مثال های کتاب نقش داشتند، تشکر و قدردانی می کنند. نویسندگان در تهیه این اثر سعی و تلاش کرده اند تا کار بی نقص باشد. لیکن ذات کارهای پژوهشی به گونه ای است که امکان خطاهای ناخواسته در آن ها وجود دارد. به همین جهت از کلیه صاحب نظران و خوانندگان این متن تقاضا داریم تا کاستی ها را به دیده اغماض نگریسته و نظرات و پیشنهادهای ارزشمند خود را به ما اطلاع دهند تا در ویرایش های آینده آن ها را منظور کنیم.

دکتر فرامرز سهیلی و مهندس علی رستم پور

پاییز ۱۳۹۹

fsohieli@gmail.com

فصل اول

مفاهیم پایه پایگاه داده‌ها^۱

هدف کلی

هدف کلی این فصل آشنایی با مطالب پایه‌ای مرتبط با مفاهیم پایگاه داده و سیستم مدیریت پایگاه داده^۲ است.

هدف‌های یادگیری

انتظار می‌رود پس از مطالعه این فصل یادگیرنده بتواند:

۱. پایگاه داده را معرفی کند.
۲. مفاهیم و اصطلاحات رایج در پایگاه داده را بشناسد.
۳. سیستم مدیریت پایگاه داده را بفهمد.
۴. با انواع پایگاه داده آشنا شود.
۵. اجزای تشکیل‌دهنده پایگاه داده را بشناسد.
۶. معماری پایگاه داده را بفهمد.
۷. کاربران و مدیر پایگاه داده را بشناسد.

1. Database

2. Database management system(DBMS)

مقدمه

پیش از یادگیری هر موضوعی، در ابتدا لازم است که به یادگیری اصطلاحات و مفاهیم اساسی آن موضوع پرداخت. در این فصل ابتدا به تفصیل مطالب پایه‌ای مرتبط با مفاهیم پایگاه داده، سیستم مدیریت پایگاه داده بیان می‌شود و مفاهیم و اصطلاحات ضروری برای آشنایی و درک مناسب کلیات موضوع ارائه می‌شود. با معرفی دو مشی فایلینگ و پایگاه داده، خواننده با امتیازات و توانایی‌های روش مفید و برتر ایجاد پایگاه داده آشنا می‌شود. عناصر تشکیل دهنده و معماری بانک درک عمیق‌تری از ساختار پایگاه داده را ارائه می‌دهد. با معرفی جایگاه کاربران و مدیران بانک و نیز سطح امنیت و اصول جامعیت، ویژگی‌های برتر بانک آشکار می‌شود. همچنین به‌طور خلاصه سیر تحول و پیشرفت مدل‌های پایگاه داده و برنامه‌های نرم‌افزاری مرتبط با آن شرح داده خواهد شد.

۱-۱ پایگاه داده چیست؟

بانک اطلاعاتی یا پایگاه داده یک اصطلاح بسیار رایج در دنیای کامپیوتر است که بیش از آنچه که تصور می‌شود نه تنها در دنیای کامپیوتر، بلکه در زندگی روزمره نیز جاری و قابل مشاهده است. همه ما با مفهوم واضح بانک برای نگهداری وجوه نقد و یا فلزات گران بها آشنا هستیم. کتابخانه نیز مثال شهودی دیگری از یک بانک شامل مجموعه عظیمی از داده‌های فیزیکی به نام کتاب است. در واقع می‌توان بانک را محلی برای نگهداری داده‌هایی از هر نوع دانست. مدرسه، دانشگاه یا هر اداره و سازمانی برای شناسایی افراد تحت مجموعه خود نیاز به بانکی حاوی اطلاعات این افراد خواهد داشت. جهت سرعت بخشیدن به امور و صرفه‌جویی در رفت و آمد و نیز مقایسه در مواردی چون انتخاب منزل مسکونی، دفتر اداری، رستوران و یا هر موضوع دیگری دسترسی داشتن به مجموعه‌ای از اطلاعات مورد نیاز بسیار راه‌گشا و کاربردی خواهد بود. به شیوه بسیار ساده می‌توان بانک اطلاعات را مجموعه‌ای از داده‌های به هم مرتب با ساختار منظم که برای پاسخ‌گویی به نیازهای اطلاعاتی یک سازمان گردآوری شده‌اند تعریف کرد.

از تعریف ساده فوق می‌توان چند نکته را برداشت کرد: اول آنکه پایگاه داده شامل مجموعه‌ای از داده‌های با مفهوم در مورد یک موضوع مشخص است. بنابراین این داده‌های با معنی و مرتبط به یک موضوع می‌بایست تحت یک ساختار مشخص با نظم از

پیش تعیین شده‌ای سازماندهی شوند. جهت نظم‌دهی به دارایی‌های یک پایگاه که همان اطلاعات است نیاز به یک قالب خاص وجود دارد. سازماندهی داده‌ها در قالب موردنظر نیز نیازمند به یک سیستم مدیریت خواهد بود که در آینده شرح داده خواهد شد.

در متون مختلف مرتبط با پایگاه داده، تعاریف متعددی برای این مفهوم بیان شده است. به عنوان مثال: پایگاه داده مجموعه‌ای از اطلاعات شامل داده‌های سازماندهی شده مربوط به یک مؤسسه یا سازمان است (سیلبرشاتس، کورتش و سودارشان^۱، ۲۰۱۱)، یا پایگاه داده مجموعه از داده‌های پایدار است که توسط سیستم‌های کاربردی موجود در مؤسسه یا سازمان مورد استفاده قرار می‌گیرد (دیت^۲، ۲۰۰۳). یا از دیدگاه علم کتابداری: پایگاه اطلاعاتی به سیستم نگهداری مجموعه‌ای از داده‌ها اطلاق می‌شود که هدف آن حفظ اطلاعات و اشاعه آن در مواقع لزوم است. اطلاعات مورداشاره، به آن چیزی گفته می‌شود که فرد با سازمان برای گردش امور جاری خود به آن نیازمند است (خسرویانی، ۱۳۷۴).

در اصطلاح‌نامه کتابداری، در تعریف پایگاه اطلاعاتی آمده است: واحدی از سوابق و بایگانی‌های قابل خواندن با ماشین که برای یک کاربرد واحد تعبیه نشده، بلکه به منزله یک مجموعه متجانس برای مقاصد مختلف به کار می‌رود (سلطانی، ۱۳۹۵). همچنین تعاریف گوناگون دیگر در سایر متون کلاسیک وجود دارد که شاید جامع‌ترین تعریف، در (روحانی رانکوهی، ۱۳۹۲) بیان شده است: «مجموعه‌ای از داده‌های ذخیره شده و پایا^۳، به صورت مجتمع (یکپارچه^۴) (نه لزوماً همیشه به صورت فیزیکی، بلکه به طور منطقی)، به هم مرتبط^۵ حتی الامکان با کمترین افزونگی^۶ (دارای یک ساختار منطقی مبتنی بر یک مدل داده^۷ توصیف شده به ویژه در چارچوب همان مدل داده و نیز داری معماری خاص)، تحت مدیریت یک سیستم کنترل متمرکز^۸، مورد استفاده یک یا چند کاربر از یک (یا بیش از یک) سیستم کاربردی»، «به طور همروند^۹ و اشتراکی^{۱۰}».

1. Silberschatz, Korth & Sudarshan

2. Date

3. Persistent

4. Integrated

5. Interconnected

6. Redundancy

7. Data Model

8. Centralized control

9. Concurrent

10. Shared

این تعریف علمی را می‌توان به صورت ساده و جامع زیر بیان کرد. «مجموعه‌ای از داده‌های ذخیره شده پایدار، که به طور مجتمع و مبتنی بر یک ساختار با حداقل میزان افزونگی، تحت مدیریت یک سیستم متمرکز که به طور اشتراکی و هم‌زمان مورد استفاده یک یا چند کاربر یا برنامه کاربردی باشد» (مقسمی، ۱۳۸۹). به طور واضح می‌توان وجه اشتراک تمامی تعاریف فوق را به این صورت برداشت کرد: مجموعه‌ای از اطلاعات به هم مرتبط در یک حوزه مشخص که توسط یک سیستم مدیریت کنترل می‌شود. اصطلاحات بانک اطلاعاتی و پایگاه داده‌ها در بسیاری از موارد به خصوص در علوم کامپیوتری، مفاهیمی هم وزن و مترادف یکدیگر دانسته شده و به صورت معادل هم به کار می‌روند.

در حوزه علم اطلاعات و دانش‌شناسی نیز تلاش‌هایی جهت بررسی وجوه اشتراک و افتراق این دو مفهوم صورت پذیرفته است. اهمیت بررسی این موضوع در مباحث کنترل واژگان در فرایند نمایه‌سازی اطلاعات و نیز گزینش مدخل و برقراری ارجاع میان اصطلاحات در دایره‌المعارف‌ها و فرهنگ‌های کتابداری و اطلاع‌رسانی است. اصطلاح‌نامه کتابداری و علم اطلاعات انجمن اطلاع‌رسانی آمریکا^۱ که از معتبرترین اصطلاح‌نامه‌ها در حوزه کتابداری و اطلاع‌رسانی است، در ویرایش ۱۹۹۸ با بررسی مدخل‌های «پایگاه اطلاعاتی» و «بانک اطلاعاتی»، هرچند این دو واژه را از لحاظ مفهومی جدا از یکدیگر نمی‌داند لیکن واژه پایگاه اطلاعاتی را بر بانک اطلاعاتی ارجحیت داده است. همچنین پایگاه اطلاعاتی چکیده‌نامه کتابداری و اطلاع‌رسانی (لیزا)^۲ اعلام می‌دارد که اصطلاح پایگاه اطلاعاتی واژه پذیرفته شده و عام است و واژه بانک داده یک اصطلاح خاص در سلسله مراتب روابط مفهومی میان واژه‌ها به شمار می‌آید.

تعاریف ارائه شده از دایره‌المعارف بین‌المللی کتابداری و اطلاع‌رسانی^۳ برای این دو مفهوم نیز به شرح زیر است: پایگاه اطلاعاتی: «به طور کلی به هر مجموعه نظام‌مند از اطلاعات در هر قالبی که باشد، حتی به صورت چاپی، می‌توان پایگاه اطلاعاتی اطلاق کرد. داده‌ها در یک پایگاه اطلاعاتی به صورت پرونده‌های رایانه‌ای یا بر روی صفحه فشرده

1. ASIS Thesaurus of Information Science and Librarianship

2. Library and Information Science Abstract (LISA)

3. International Encyclopedia of Information and Library Science

ذخیره می‌شوند. هر پایگاه اطلاعاتی ممکن است حاوی اطلاعات کتاب‌شناختی، عددی یا آماری، و جز آن باشد. داده‌ها در هر پایگاه اطلاعاتی معمولاً به گونه‌ای ساختاردهی می‌شوند که امکان بازیابی خودکار آن‌ها وجود داشته باشد» و تعریف بانک اطلاعاتی: «این اصطلاح معمولاً مترادف پایگاه اطلاعاتی به کار می‌رود، اما در برخی موارد از اصطلاح بانک اطلاعاتی به منظور مشخص کردن مجموعه‌های غیرکتاب‌شناختی و داده‌های عددی استفاده می‌شود». پرواضح است که هرچند برای این دو مفهوم دو تعریف ارائه شده است اما تفاوت بنیادی بین آن‌ها مشاهده نمی‌شود. با تجزیه و تحلیل آماری در میان مقالات و مرور کتاب‌های نگاشته‌شده در سال‌های اخیر در می‌یابیم که باتوجه به اینکه این دو مفهوم دارای ساختار واحدی بود و نیز یک هدف مشخص را دنبال می‌کنند (مجموعه‌ای نظام‌مند از داده‌ها که امکان جستجو و بازیابی اطلاعات در آن‌ها وجود دارد باهدف دسترسی سریع و کارآمد به اطلاعات) تنها تفاوت بیان‌شده، در نوع اطلاعات قابل دستیابی در این دو نظام است: (در پایگاه اطلاعاتی کلیه انواع اطلاعات و در بانک اطلاعاتی تنها اطلاعات عددی ذخیره می‌شود) و به این دلیل اصطلاح پایگاه اطلاعاتی بیشتر مورد استفاده قرار گرفته است (کوشا، ۱۳۸۱).

در تعاریف پایگاه داده از مفاهیمی چون داده، اطلاعات و سیستم مدیریت پایگاه داده و موارد دیگری نام برده شده است که نیاز به تشریح دارند. در ادامه این فصل هریک از این موارد به تفصیل موردبررسی قرار خواهند گرفت.

۱-۱-۱ معرفی مفاهیم

در این قسمت برای رسیدن به یک ادبیات مشترک، به تشریح برخی از مفاهیم پایه در حوزه پایگاه داده پرداخته شده است

- **داده‌ها:** واژه داده ریشه لاتین دارد و مفرد آن واژه Datum است که در متون فنی معمولاً به صورت جمع به کار می‌رود. در منابع مختلف، تعاریف متعددی برای داده معرفی شده است که عمده آن‌ها به این مسئله اشاره دارند که داده واقعیاتی هستند در مورد اشیاء، افراد و رخدادها که می‌توانند ذخیره شوند و معمولاً آگاهی ایجاد نمی‌کنند و به صورت خام هستند (پردازش و سازماندهی بر روی آن‌ها انجام نشده است).

به عنوان مثال: در پایگاه داده کتابخانه نام یک فرد تنها یک داده محسوب می‌شود و هیچ اطلاعی از اینکه این نام به عضو کتابخانه یا پرسنل تعلق دارد و یا اینکه نام نویسنده، مؤلف و یا مترجم و ویراستار یک کتاب هست یا خیر وجود ندارد. یا مقدار ۱۸۰ که یک عدد است داده‌ای خام است که آگاهی‌ای برای خواننده ایجاد نمی‌کند که این عدد به چه پدیده‌ای مربوط است، مثلاً قد یک فرد هست یا تعداد نفرات یک گروه. دقت کنید که منظور از پردازش لزوماً انجام محاسبات ریاضی نیست، بلکه سازماندهی کردن و مرتبط کردن داده به یک پدیده مشخص که برای ما آگاهی ایجاد کند، داده‌ها را از حالت خام بودن خارج کرده و به اطلاعات تبدیل خواهد شد (روحانی رانکوهی، ۱۳۹۲)

مؤسسه ملی استاندارد آمریکا (ANSI)^۱، تعریف زیر را برای داده ارائه کرده است. واقعیت‌ها، پدیده‌ها، معلومات و مفاهیم که به صورت مناسب جهت برقراری ارتباط، تفسیر یا پردازش توسط انسان یا امکانات خودکار نمایش داده می‌شود. پس می‌توان گفت داده نمایشی از هر کاراکتری است که با نسبت دادن معنایی به آن بتوان با انجام عملیاتی روی آن اطلاعاتی به دست آورد (روحانی رانکوهی، ۱۳۹۲).

• **اطلاعات^۲:** اطلاعات عبارت است از داده‌های سازماندهی شده که معمولاً آگاهی ایجاد کرده و شناختی را منتقل می‌کنند و می‌توان آن‌ها را مبنایی برای تصمیم‌گیری قرار داد. مؤسسه ملی استاندارد آمریکا^۳ تعریف جامعی از اطلاع را بیان کرده که در ادامه آورده شده است (روحانی رانکوهی، ۱۳۹۲): «اطلاعات معنایی است که انسان به داده منتسب می‌کند، از طریق قراردادهای شناخته شده‌ای که در نمایش داده به کار می‌روند». به بیان ساده‌تر شاید بتوان گفت که از پردازش و سازماندهی داده‌ها یا تفسیر داده‌ها اطلاعات به دست می‌آید که در مباحث پایگاه داده، تفسیر داده معمولاً با انتساب یک مقدار به صفت مربوط به یک موجودیت به دست می‌آید. برای مثال در مورد موجودیت کتاب (صفت نام کتاب = ترسیم نقشه علم) یک فقره اطلاعات است یا در مورد موجودیت دانشجو (نام دانشجو = رضا مرادی) یک فقره اطلاعات است.

1. American national standards institute

2. Information

3. ANSI

اطلاعات پدیده‌ای کاملاً فراگیر و قابل دستیابی در ابعاد و سطوح مختلف است، که از قبل قابل تصور نبوده و مقدار آن بسیار زیاد است. نیاز به اطلاعات که نیستی و پایداری را نمی‌توان برای آن تصور کرد، پدیده‌ای است که هیچ گریزی از آن نیست و هیچ‌گاه جریان رشد و افزایش آن متوقف نمی‌شود. با افزایش جمعیت و پیچیده‌تر شدن بافت جامعه انسانی، تولید اطلاعات و نیاز انسان به اطلاعات هم گسترش یافت (بشیری و عزیزی، ۱۳۸۴، ۱۷). در گرایش‌ها و رشته‌های مختلف این واژه دارای معانی متفاوت است، اما معانی عمومی زیر را می‌توان برای این واژه، ارائه کرد:

۱. آگاهی به دست آمده از یک مطالعه، پژوهش، تجربه یا دستورالعمل.
 ۲. آگاهی به دست آمده از وقایع یا موقعیت‌های خاص که از طریق ارتباط، جمع‌آوری یا دریافت شده است.
 ۳. مجموعه‌ای از واقعیت‌ها یا داده‌ها.
 ۴. عمل آگاه کردن یا آگاه شدن.
 ۵. داده‌های رایانه‌ای پردازش و نگهداری شده یا انتقال یافته.
- تعریفی جامع از اطلاعات در دایرةالمعارف کتابداری و اطلاع‌رسانی (۲۰۰۳) آمده است: «اطلاعات شامل داده‌هایی است که به شکلی دارای معنی و مفهوم پردازش شده است». در این شکل، اطلاعات ترکیبی از داده‌ها به شکلی قابل فهم است که قادر است ایجاد ارتباط کند و مورد استفاده قرار گیرد.
- ویسمان^۱ برای اطلاعات شش متغیر در نظر گرفته است (نقل در ویلیامز، ۱۹۶۵):
۱. کمیت اطلاعات که می‌تواند با تعداد مدارک، صفحات، واژه‌ها، حروف، ترسیم‌ها، تصاویر و غیره، اندازه‌گیری شود.
 ۲. محتوا که عبارت از معنی و مفهوم اطلاعات است.
 ۳. ساختار که مربوط به شکل ظاهری یا سازمان اطلاعات و رابطه منطقی بین عبارات یا عناصر است.
 ۴. زبان که مشتمل بر علائم، حروف الفبا، رموز و ترکیباتی است که به وسیله آن‌ها، عقاید بیان می‌شود.

۵. کیفیت که کامل بودن، صحت، مناسب بودن و به هنگام بودن اطلاعات را توصیف و مشخص می‌کند.

۶. عمر، که عبارت از مدت زمانی است که اطلاعات دارای ارزش است.

باکلمند^۱ سه معنی مختلف را برای واژه اطلاعات تعریف می‌کند:

- اطلاعات به عنوان یک فرایند: آنچه که یک فرد می‌داند تغییر می‌کند، هنگامی که او از آن آگاه می‌شود. در این معنی اطلاعات، «عمل آگاه کردن» است.
 - اطلاعات به عنوان دانش: دانش انتقال یافته درباره واقعیت، موضوع یا رویدادی خاص.
 - اطلاعات به عنوان شیء: واژه اطلاعات همچنین برای چیزهایی از قبیل داده‌ها و مدرک‌ها استفاده می‌شود که به آن‌ها به عنوان اطلاعات ارجاع داده می‌شود، چراکه اطلاع‌دهنده یا آگاهی‌دهنده تلقی می‌شوند (باکلمند، ۱۹۹۱).
- از میان انواع اطلاعاتی که باکلمند برمی‌شمارد، آنچه که مورد نظر ماست، اطلاعات به عنوان شیء است. درواقع اطلاعات به صورت شیء است که در نظام‌های ذخیره و بازیابی اطلاعات، جمع‌آوری می‌شود.

- دانش^۲: دانش ترکیبی سازمان یافته از داده‌ها است که با مجموعه‌ای از قواعد، رویه‌ها و عملیات آموخته شده از طریق تجربه و تمرین درون‌سازی شده است (داورپناه، ۱۳۸۲). دانش نه داده است و نه اطلاعات، هرچند که به هر دو مربوط است اما تفاوت آن‌ها لزوماً ماهوی نیست و فقط از نظر مراتب باهم متفاوت‌اند. دانش ترکیبی سازمان یافته از داده‌ها است که با مجموعه‌ای از قواعد، رویه‌ها و عملیات آموخته شده از طریق تجربه و تمرین درون‌سازی شده است. بنابراین دانش را درک، آگاهی یا شناختی که درطول زمان از طریق مطالعه، پژوهش، مشاهده و تجربه به دست می‌آید، تعریف کرده‌اند. دانش تلفیق تفکر با اطلاعات است. درواقع تفسیری فردی از اطلاعات برپایه تجربیات، مهارت‌ها و توانمندی‌های شخصی است (بولینگر، ۲۰۰۱: نقل در داورپناه، ۱۳۸۱).

زینس^۱ (۱۳۹۰) دسته‌بندی متفاوتی از این سه مقوله ارائه کرده‌اند که در اینجا به آن‌ها پرداخته شده است.

• **قلمروهای ذهنی در مقابل قلمرو عینی:** «سه پدیده «داده»، «اطلاع» و «دانش» از دو الگوی وجودی متمایز، به نام‌های قلمروهای عینی و قلمروهای ذهنی، برخوردارند. بر همین قیاس، بین دانش ذهنی و دانش عینی تمایز قائل می‌شویم. توجه داشته باشید، «دانش ذهنی» معادل «دانش ذهن یا دانسته فردی» است و «دانش عینی» معادل «دانش به مثابه دانش یک شیء یا یک چیز» است. دانش ذهنی در جهان درونی فرد (به عنوان یک اندیشه) وجود دارد، حال اینکه دانش عینی در جهان بیرونی فرد (یعنی، آن گونه که در کتاب‌ها منتشر می‌شود، در کتابخانه دیجیتال ارائه و در ابزارهای الکترونیکی ذخیره می‌شود) وجود دارد. جهت تسهیل درک مفاهیم، اصطلاحات «دانش همگانی‌شده» به مفهوم دانش جمعی (یعنی دانش یا علمی که در قلمرو جمعی قابل درک و استفاده است) را جایگزین اصطلاح «دانش عینی» خواهیم کرد. تمایز بین دانش ذهنی و دانش همگانی‌شده، همچون تمایز تفاوت بین دانش خصوصی و دانش عمومی است. «دانش خصوصی» دانش شخصی متعلق به فرد است. این نوع اندیشه‌ها محتوایی هستند که تنها فرد آن‌ها را می‌داند، مانند احساس‌ها و پندارهای درونی. «دانش جمعی» به اندیشه‌هایی ارجاع می‌دهد که فرد آن‌ها را دانش می‌پندارد و در حیطه محتوای دانسته‌شده توسط دیگران نیز هستند. مثال‌های ساده‌ای چون: $2+2=4$ ، یا تهران پایتخت ایران است».

• **داده - اطلاعات - دانش در قلمرو ذهنی:** «در قلمرو ذهنی، «داده»ها محرک‌های حسی هستند که از راه حواس، دریافت (یا درک) می‌شوند. «اطلاعات» معنای این محرک‌های حسی است (یعنی درک تجربی). برای مثال، صداهایی که می‌شنویم، داده هستند. معنای این صداها (برای مثال، صدای موتور ماشین روشن) «اطلاع» است. شیوه دیگری را می‌توان برای تعریف این دو مفهوم در نظر گرفت - که حتی بهتر به نظر می‌رسد. «داده»ها محرک‌های حسی، یا معنای آن‌ها (یعنی، درک تجربی) هستند. بر این اساس، در مثال بالا، صداهای بلند، و به همان گونه درک حرکت موتور ماشین، «داده» هستند. اطلاعات دانش تجربی است. بر این اساس، در مثال بالا، دانش

اینکه موتور اکنون روشن است و ماشین در حال دور شدن است، «اطلاعات» است. این موردی است که مبتنی بر تجربه است. اطلاعات یک نوع دانش است و نه یک مرحله واسطه‌ای بین داده و دانش. دانش اندیشه‌ای در ذهن فرد است که بنابر باور فرد، از درستی قابل توجیه برخوردار است. دانش، بسته به مورد، می‌تواند تجربی یا غیر تجربی باشد. برای مثال در منطق یا ریاضی (به‌طور مثال، هر سه ضلعی دارای سه طرف است)؛ دانش دینی (مثلاً، خدا وجود دارد) و غیره. توجه داشته باشید که دانش، محتوای اندیشه در ذهن فرد است، که بنابر باور فرد، از درستی قابل توجیهی برخوردار است. درحالی که «دانستن» حالتی از ذهن و دارای سه مشخصه است: ۱. فرد باور دارد که درست است ۲. فرد می‌تواند آن را توجیه کند، ۳. آن حالت درست است یا به‌نظر می‌رسد درست باشد».

- داده - اطلاعات و دانش در قلمرو همگانی (عمومی): «در قلمرو همگانی، داده، اطلاعات و دانش ساخته (مصنوع) انسان هستند. این سه با نشانه‌های (علائم) تجربی، بازنمون می‌شوند (یعنی، نشانه‌هایی که فرد می‌تواند از طریق حواس‌های خود آن‌ها را حس کند). این نشانه‌ها می‌توانند شکل‌های متفاوتی از قبیل: علائم کنده‌کاری، شکل‌های نقاشی، کلمات چاپ، علائم دیجیتالی، اشعه‌های نور، امواج صدا، و ... به خود بگیرند. داده‌های همگانی، اطلاعات همگانی و دانش همگانی، منعکس‌کننده المثنی شناختی خود هستند. معنا، در قلمرو عینی داده‌ها مجموعه‌ای از علائم است که بازنمون محرک تجربی یا ادراک‌ها است. اطلاعات، مجموعه‌ای از علائم است که بازنمون دانش تجربی است، و دانش مجموعه‌ای از علائم است که بازنمون معنا (یا محتوای) اندیشه‌هایی است که فرد به‌گونه قابل توجیه، باور دارد که آن‌ها درست هستند».

- داده‌های پایا و مانا: داده‌های ذخیره‌شده در پایگاه داده را داده‌های باثبات می‌نامند. داده‌های واردشده به یک تابع ریاضی ورودی گفته می‌شود که پس از انجام یک‌سری پردازش تبدیل به خروجی می‌شوند. این نوع داده‌های ورودی - خروجی ماهیت پایدار ندارند و بسته به شرایط دچار تغییر هستند. اما داده‌های واردشده به سیستم پایگاه داده که از طریق مدیریت این سیستم پذیرفته می‌شوند، تحت شرایط جانبی دچار حذف و یا تغییر نخواهند شد و تنها زمانی که درخواستی به مدیریت سیستم داده شود، تغییر اعمال می‌شود. برای مثال سیستم دانشگاه که حاوی اطلاعات دانشجویان شامل مشخصات،