Swami Sahajanand College of Computer Science

B.C.A. SEM-V[NEP]

Subject: DATABASE TECHNOLOGY IN INDIA
Major12 - 26516

UNIT 1

DATABASE TECHNOLOGY IN INDIA

- Introduction to Database Systems in India, Understanding Data, Information, and Database, Concepts, Early Methods of Data Storage and Record Management in Indian Institutions, Evolution from Manual Records to Digital Databases in India.
- Early Adoption of Database Technology in India (1970s–1980s) Introduction of Computers in Indian Government Departments and Public Sector Units, Use of Mainframes and Batch Processing Systems in Indian Railways, Banks, and Census Operations, Initial Database Practices: Indexed Files, Hierarchical Systems in Indian Enterprises.
- Growth of RDBMS and Structured Data Storage (1990s) Entry of International RDBMS vendors (Oracle, Sybase, IBM DB2) into the Indian Market, Widespread Adoption of Relational Databases in: Public Sector Banks (State Bank of India, Punjab National Bank, etc.), Government Data Centers, Manufacturing and Telecom Industries, Shift from Legacy Systems to Structured Query Language (SQL) based systems.
- Role of Indian IT Industry in Database Development Contribution of TCS, Infosys, Wipro, HCL, and others in Database Solutions and Custom Software. Growth of Enterprise Resource Planning (ERP) systems for Indian Businesses, Outsourcing Boom and the Need for Scalable Database Backends (late 1990s - 2000s).
- Indian Government Initiatives and National Database Projects National Informatics Centre (NIC) – Building Database Solutions for E-Governance, Aadhaar and UIDAI: Development of the World's Largest Biometric Database, Election Commission, National Population Register, and Passport Seva Databases, Digital India Initiative and Database-Driven Citizen Services

❖Introduction to Database Systems in India

- A Database Management System (DBMS) is a software solution designed to efficiently manage, organize, and retrieve data in a structured manner.
- It serves as a critical component in modern computing, enabling organizations to store, manipulate, and secure their data effectively.
- From small applications to enterprise systems, DBMS plays a big role in supporting data-driven decision-making and operational efficiency.

❖Understanding Data, Information, and Database Concepts,

- **Data:** "Data can be anything for E.g. Person data, Company data, Student data(name, id and age)."

  OR

  "Data is unprocessed information"
- Data can be written in computerized or non-computerized format.
- Data must be processed in proper way to generate useful and meaningful information.

- **Information:**

  "Processed data is known as information." Information is organized or classified data.

- **Database:**

  "Database is an organized collection of table"
- A database is a tool for collecting and organizing information.
- Data that is stored in a computer so that a program can use it to answer queries.
- Database can store information about people, products, banking or anything else.
- Example: Telephone dictionary.

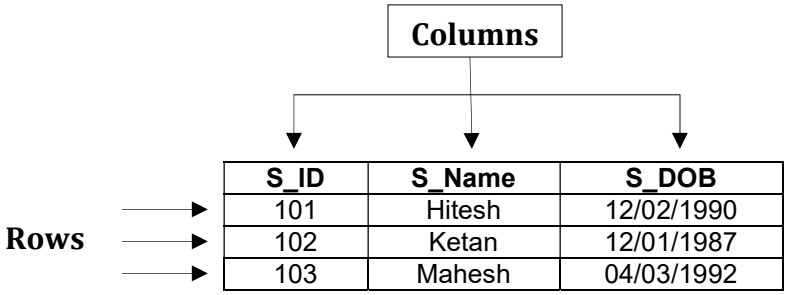
Student Database		
Std_detail_table	Std_mark_table	Std_project_table

- **Table:**

  "Table is a collection of rows and columns"
  - Table is used to store information into rows and column. A table represents an entity.
  - E.g.: a table **Student_detail** stores the information about student detail like name, id, address, DOB etc.
- **Column [Field/Attribute]:**

  "A column is a vertical entity in a table that contains all information associated with a specific field in a table."
  - In table each column has a column name and contain value.
  - E.g.: in a **Student_detail** table S_ID, S_Name, S_Address, S_DOB etc. is a column name.
- **Row [Record/ Tuple]:**

  "A row is a horizontal entity in a table."
  - Row is a combination of column values represented in sequence.



S_ID	S_Name	S_DOB
101	Hitesh	12/02/1990
102	Ketan	12/01/1987
103	Mahesh	04/03/1992

❖Early Methods of Data Storage and Record Management in Indian Institutions

- Early data storage and records management relied on physical objects and manual processes before the advent of computers. These methods included tally marks, clay tablets, punch cards, and magnetic tape, which evolved from simple inventory tracking to more complex record keeping. Records management systems, such as filing cabinets and document management systems, also emerged to handle the increasing volume of information.

Early Methods of Data Storage:

- **Tally Marks:**
- Ancient civilizations used tally marks or tick marks on stones, wood, or other materials to keep track of inventories like food or other resources.
- **Clay Tablets:**
- Writing on clay tablets was used for record-keeping, accounting, and storing libraries of information.
- **Punch Cards:**
- Punch cards, used to program mechanical devices like looms and player pianos, were later adopted for data storage and processing.
- **Magnetic Tape:**
- Magnetic tape, initially used for audio recording, became a widely used data storage medium for computers, especially for backup and archiving.

Early Records Management:

- **Filing Systems:**
- The invention of the printing press led to the need for organized storage of documents. Filing cabinets and other systems emerged to handle the increasing volume of paperwork.
- **Records Managers:**
- As organizations generated more documents, records managers were hired to manage the physical and electronic records.
- **Document Management Systems:**
- Document management systems (DMS) began to develop during the 1980s, focusing on digitizing and storing paper documents to manage large volumes of information more efficiently.

❖Evolution from Manual Records to Digital Databases in India.

- The shift from manual records to digital databases has been a significant evolution in data management, moving from paper-based systems to electronic storage and retrieval. This transition has led to more efficient, accessible, and scalable data management, ultimately improving productivity and decision-making.

Early Stages: Manual Systems

- Before the advent of computers, data was stored manually, often in filing cabinets with paper records organized in folders. This system, while foundational, had limitations: accessibility was time-consuming, scalability was constrained by physical space, and error-prone processes were common.

The Rise of Digital Systems

- The introduction of computers brought about the dawn of digital databases, allowing for more significant data storage and complex processing. These digital systems offered several advantages over manual methods, including:
- **Increased Efficiency:**
- Data retrieval and processing became faster and more automated.
- **Improved Accessibility:**
- Digital databases allowed for remote access and easier sharing of information.
- **Enhanced Scalability:**
- Digital storage could accommodate larger datasets, and new databases could be created and deployed with minimal effort.
- **Reduced Errors:**
- Automated processes reduced the risk of human error during data entry and retrieval.

❖ **Key Aspects of the Digitization Process:**

- **Scope:**
- The NAI is digitizing a vast collection of records, spanning over 250 years and encompassing millions of pages.
- **Types of Records:**
- The digitization project includes records from various eras, including pre-1947 British era documents, Home Department papers, Foreign Department records, post-1947 government records, private collections, and cartographic materials.
- **Benefits:**
- Digitization helps protect fragile records from physical deterioration, making them more accessible to researchers and the public without requiring the handling of delicate originals.
- **Challenges:**
- The digitization process involves challenges such as ensuring uniformity in metadata formats, storage methods, and adherence to a standardized operating procedure.
- **Standards and Procedures:**
- The National Committee of Archivists (NCA) has been working to establish a standardized operating procedure (SOP) for digitization, ensuring consistency and quality across different archival institutions.
- **Software and Tools:**
- C-DAC has developed the Digitalaya software, which supports controlled harvesting of records and enables the establishment of trustworthy digital repositories.
- **Metadata and Searchability:**
- Digitized records are often enriched with metadata to facilitate searching and discovery, including Optical Character Recognition (OCR) and handwriting recognition for text extraction.
- **Security and Preservation:**
- Digital records are subject to data security risks, requiring the implementation of strong security measures, such as encryption, backups, and secure storage solutions.
- **Public Accessibility:**

- Many digitized records are made available through online portals and databases, making them accessible to researchers and the public globally.

Examples of Initiatives:

- **National Archives of India (NAI):**
- The NAI is at the forefront of the digitization effort, focusing on preserving and making available its vast archival holdings.
- **State Archives:**
- Many state archives are also actively involved in digitizing their collections, working towards a more comprehensive digital archive of India's historical records.
- **Digital India Initiative:**
- The digitization effort is part of the broader Digital India initiative, which aims to transform India into a digitally-enabled society.

Ongoing Efforts and Future Directions:

- **Standardization:**
- Efforts are underway to standardize digitization processes, metadata formats, and storage methods across different archives.
- **Public Portals:**
- More digitized records are being made available through public portals, enhancing accessibility and promoting research.
- **Collaborative Initiatives:**
- Collaborations between different archival institutions and research organizations are crucial for sharing resources and expertise.
- **Digital Preservation:**
- Ensuring the long-term preservation of digital records is a key challenge, requiring ongoing investment in infrastructure and expertise.
- **User Engagement:**
- Engaging researchers, historians, and the general public in the digitization process can help ensure the relevance and usefulness of the digital archives.

❖Early Adoption of Database Technology in India (1970s–1980s)

- The 1970s and 1980s saw the nascent stages of database technology adoption in India, primarily influenced by the growth of the IT industry and government policies. Initially, the focus was on building domestic expertise in the computer industry, with institutions like TIFR and IITs playing a key role in education and research. While network and hierarchical databases were popular during this period, the relational model gained academic attention but faced initial performance challenges. The development of more efficient relational database systems, like System R, would later pave the way for their wider adoption in the 1980s.

Key Developments and Influences:

- **Government Initiatives:**
- The Indian government actively supported the development of local expertise and the creation of indigenous computer technologies.
- **Educational Institutions:**

- Institutions like TIFR and IITs played a crucial role in training IT professionals and fostering research in database technologies.
- **Early Adoption of Network and Hierarchical Databases:**
- These database models were prevalent in the 1970s, particularly on mainframe and minicomputers, according to the International Journal of Engineering Trends and Technology.
- **The Relational Model's Rise:**
- While academics explored the relational model in the 1970s, its practical adoption was hindered by performance limitations.

❖**Introduction of Computers in Indian Government Departments and Public Sector Units**

- The introduction of computers in Indian government departments and public sector units (PSUs) began with initial exploration and experimentation, gradually evolving into a more integrated and strategic approach. The initial focus was on using computers for administrative tasks and accounting, later expanding to include more diverse applications like information retrieval, data processing, and policy development.

Early Stages (Pre-1970s):

- **Limited Adoption:**
- Early adoption was primarily in areas like statistics and accounting, with limited applications in other government departments.
- **Indigenous Development Efforts:**
- The Indian government, through the Department of Electronics (DoE) and organizations like the Electronics Corporation of India Ltd. (ECIL), focused on developing indigenous computer technology.
- **Focus on Correctness and Efficiency:**
- The initial focus was on ensuring the accuracy and efficiency of accounting and data processing.

Growth and Expansion (1970s - 1990s):

- **Increased Automation:**
- The use of computers expanded to include more diverse applications, including information retrieval and data processing.
- **Establishment of Computer Centers:**
- Computer centers were set up in various government departments to support the use of technology.
- **Role of the National Informatics Centre (NIC):**
- The NIC played a key role in developing and implementing computer-based systems for the government, including PARLIS (Parliament Library and Information System) and NICNET (NIC's satellite-based network).
- **Early Initiatives in the Parliament:**
- The idea of introducing automation in the Indian Parliament was conceived in 1982, and a computer-based information retrieval system was established in 1985 according to the Rajya Sabha Secretariat report.

Modern Era (Post-1990s):

- **Increased Use of Technology:**
- The adoption of computers and information technology (IT) has grown significantly in government departments and PSUs.
- **Focus on E-Governance:**

- The government has been actively promoting e-governance initiatives, using IT to improve service delivery and citizen engagement.
- **Cloud Computing and Digital Platforms:**
- Cloud computing and digital platforms like GeM (Government e-Marketplace) have been introduced to enhance efficiency and transparency in government operations.
- **Digital Infrastructure:**
- The National Informatics Centre (NIC) has been working to upgrade the national cloud infrastructure, enabling faster and more efficient delivery of e-Governance services according to the Press Information Bureau.

❖Initial Database Practices: Indexed Files, Hierarchical Systems in Indian Enterprises

- In Indian enterprises, "Indexed Files" refers to organized collections of digital files using a hierarchical structure, often employing a system of folders and subfolders to improve organization and retrieval. This contrasts with a "Hierarchical System" which describes the organizational structure within a company itself, where employees are often grouped into levels of authority. While "Indexed Files" are a tool within the enterprise, the "Hierarchical System" is the structure the enterprise uses to operate.

Indexed Files:

- **Organization:**
- Files are arranged within a directory structure, with main folders (or "directories") potentially containing subfolders, creating a tree-like hierarchy.
- **Purpose:**
- The goal is to make finding specific files easier and faster, particularly as the number of files grows.
- **Indexing:**
- A mechanism, potentially a digital index, might be used to quickly locate files by name, date, content, or other metadata.
- **Benefits:**
- Efficient storage, improved searchability, and easier collaboration when files are organized and accessible.

Hierarchical Systems in Indian Enterprises:

- **Structure:**
- Many Indian enterprises follow a hierarchical organizational structure, with different levels of management and authority.
- **Layers:**
- This can include top management, middle management, and a lower level of employees, with each level having a specific role and responsibility.
- **Culture:**
- Hierarchical structures are sometimes seen as deeply ingrained in Indian culture, where seniority and respect for authority are valued.
- **Challenges:**
- While hierarchical structures can provide clarity and control, they can also create communication barriers and potentially limit innovation.
- **Examples:**
- Large corporations, government organizations, and many traditional businesses in India typically utilize hierarchical structures.

❖Growth of RDBMS and Structured Data Storage (1990s)

- The 1990s witnessed the explosive growth of Relational Database Management Systems (RDBMS) and the rise of structured data storage, driven by the increasing use of the internet and the rise of e-commerce. This era saw the dominance of relational databases like Oracle, SQL Server, and IBM DB2, which provided a structured and reliable way to manage and access data.

Key Features and Trends:

- **Relational Database Dominance:**
- RDBMSs, based on the relational data model and using SQL, became the standard for managing structured data.
- **SQL as the Language of Data:**
- SQL (Structured Query Language) became the primary language for interacting with databases, enabling developers to query and manipulate data.
- **ACID Compliance:**
- RDBMSs provided ACID (Atomicity, Consistency, Isolation, Durability) properties, ensuring data integrity and reliability, crucial for financial systems and other applications.
- **Data Warehousing:**
- Early data warehousing efforts in the 1990s involved using RDBMSs to store large amounts of aggregated data for analysis and decision support.
- **Rise of the Internet and Web Applications:**
- The explosion of the internet and the growth of e-commerce led to increased demand for database systems to support dynamic and interactive web applications.
- **Emergence of MySQL:**
- MySQL, introduced in 1995, offered a more open-source and affordable alternative to the commercial RDBMS systems.
- **Evolution of Object-Relational Databases:**
- Some database vendors also began to incorporate object-relational features to better handle object-oriented data types.

❖Entry of International RDBMS vendors (Oracle, Sybase, IBM DB2) into the Indian Market

- The international RDBMS vendors Oracle, Sybase, and IBM DB2 entered the Indian market with varying degrees of success, primarily focusing on the enterprise sector. Oracle established itself as a market leader, while Sybase and DB2 also gained a foothold, particularly in specialized applications.

ORACLE

- Oracle, a leading international vendor of Relational Database Management Systems (RDBMS), has established a strong presence in the Indian market, becoming the leading RDBMS vendor in India. Its success can be attributed to factors like a focus on the mid-market segment, its strong R&D presence in India, and its adaptation to the unique needs of the Indian market. Oracle has also expanded its offerings to include cloud services, further solidifying its position.
- Oracle has been a leading RDBMS vendor in India for a considerable period, holding a dominant market share.
- In 2008, Oracle was named the leading RDBMS vendor in India and Asia Pacific, excluding Japan, with a significant market share.

- This dominance is reflected in their strong revenue figures, with India being one of their top revenue markets.

Sybase

- The entry of international vendors like Sybase (now owned by SAP) and Oracle into the Indian market is a well-documented phenomenon, driven by the increasing demand for sophisticated technology solutions by Indian enterprises. These companies compete for market share by offering diverse IT solutions, including enterprise resource planning (ERP), business intelligence, and financial services.
- Sybase (now SAP):
- Sybase, after its acquisition by SAP, continues to be a significant player in the Indian market, offering database solutions and other technology products.

BM DB2

- Both Oracle and IBM DB2 have established a strong presence in the Indian market, each offering different strengths and targeting various customer segments. Oracle has been a leading RDBMS vendor in India for a while, while IBM DB2 remains a significant player, particularly in specific sectors like banking and finance.

❖**Widespread Adoption of Relational Databases**

- Relational databases are widely adopted due to their ability to efficiently manage structured data, support complex queries, and enforce data integrity. They have been the backbone of enterprise computing for decades and are used in various industries, from e-commerce and banking to customer relationship management.

Here's why they are so popular:

Data Integrity and Consistency:

- Relational databases enforce data types and relationships, ensuring high levels of accuracy and data consistency.

Flexibility and Scalability:

- They offer flexibility in query using SQL, allowing for complex data manipulation and retrieval. Their ability to handle large volumes of data and scalability make them a suitable choice for various applications.

Mature Technology and Extensive Resources:

- Relational databases are a well-established technology with a vast community of developers and users, providing extensive support and resources.

Strong Transactional Support:

- They provide strong transactional support, ensuring that data changes are consistent and reliable, especially important for applications like financial systems.

SQL Support:

- The use of Structured Query Language (SQL) simplifies data management and manipulation, making relational databases accessible and easy to use.

Versatility and Compatibility:

- Relational databases are versatile and compatible with numerous other technologies and software systems, making them a good choice for many organizations.