jims Jagan Institute of Management Studies



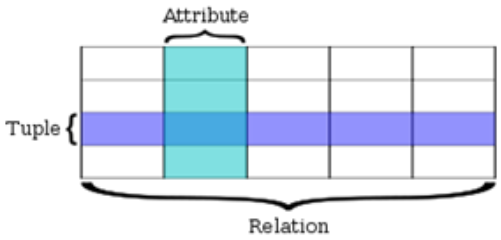

JIMS IT FLASH

Volume 8 Issue 2 Year 2014

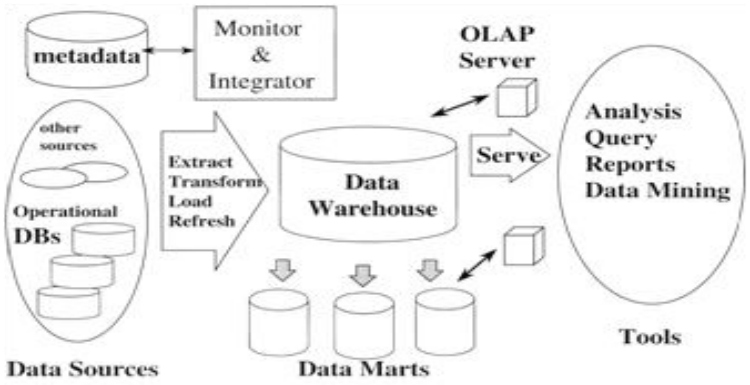
Data Mining—On What Kind of Data?

Data mining should be applicable to any kind of data repository, as well as to transient data, such as data streams. Thus the scope of our examination of data repositories will include relational databases, data warehouses, and transactional databases, advanced database systems, flat files, data streams, and the World Wide Web. Advanced database systems include object-relational databases and specific application-oriented databases, such as spatial databases, time-series databases, text databases, and multimedia databases. The challenges and techniques of mining may differ for each of the repository systems.

1. **Relational Databases:** A relational database is a collection of tables, each of which is assigned a unique name. Each table consists of a set of attributes (*columns* or *fields*) and usually stores a large set of tuples (*records* or *rows*). Each tuple in a relational table represents an object identified by a unique *key* and described by a set of attribute values. A semantic data model, such as an entity-relationship (ER) data model, is often constructed for relational databases. An ER data model represents the database as a set of entities and their relationships.



2. **Data Warehouses:-**A data warehouse is a repository of information collected from multiple sources, stored under a unified schema, and that usually resides at a single site. Data warehouses are constructed via a process of data cleaning, data integration, data transformation, data loading, and periodic data refreshing.



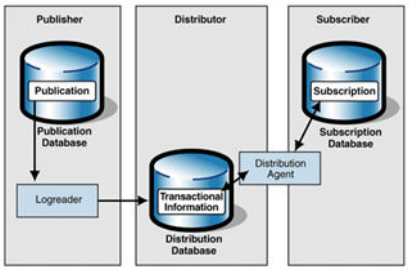
A data warehouse is usually modeled by a multidimensional database structure, where each dimension corresponds to an attribute or a set of attributes in the schema, and each cell stores the value of some aggregate measure, such as *count* or *sales amount*. The actual physical structure of a data warehouse may be a relational data store or a multidimensional data cube. A data cube provides a multidimensional view of data and allows the pre computation and fast accessing of summarized data.




JIMS IT FLASH

3. **Transactional Databases:-**Transactional database consists of a file where each record represents a transaction. A transaction typically includes a unique transaction identity number (*trans ID*) and a list of the items making up the transaction (such as items purchased in a store).

The transactional database may have additional tables associated with it, which contain other information regarding the sale, such as the date of the transaction, the customer ID number, the ID number of the salesperson and of the branch at which the sale occurred, and so on.



4. Advanced Data and Information Systems and Advanced Applications:-

Relational database systems have been widely used in business applications. With the progress of database technology, various kinds of advanced data and information systems have emerged and are undergoing development to address the requirements of new applications.

Object-Relational Databases: Object-relational databases are constructed based on an object-relational data model. This model extends the relational model by providing a rich data type for handling complex objects and object orientation. Because most sophisticated database applications need to handle complex objects and structures.

Temporal Databases, Sequence Databases, and Time-Series Databases :

A temporal database typically stores relational data that include time-related attributes. These attributes may involve several timestamps, each having different semantics. A sequence database stores sequences of ordered events, with or without a concrete notion of time.

Spatial Databases and Spatiotemporal Databases :

Spatial databases contain spatial-related information. Spatial data may be represented in raster format, consisting of *n*-dimensional bit maps or pixel maps. A spatial database that stores spatial objects that change with time is called a spatiotemporal database, from which interesting information can be mined.

Text Databases and Multimedia Databases :

Text databases are databases that contain word descriptions for objects. These word descriptions are usually not simple keywords but rather long sentences or paragraphs.

Multimedia databases store image, audio, and video data. They are used in applications such as picture content-based retrieval, voice-mail systems, video-on-demand systems, the World Wide Web, and speech-based user interfaces that recognize spoken commands. Multimedia databases must support large objects, because data objects such as video can require gigabytes of storage. Specialized storage and search techniques are also required. Because video and audio data require real-time retrieval at a steady and predetermined rate in order to avoid picture or sound gaps and system buffer overflows, such data are referred to as continuous-media data.

Jagan Institute of Management Studies




JIMS IT FLASH

JIMS IT NEWS FLASH

- **Smartphone could replace hotel keys:** Got a Smartphone? Never lose your hotel key, or even have to stop at the registration desk, again. That's the vision of a hotel chain that plans to send digital keys to guests' phones via an app instead of making them check in and get the traditional (and famously lose-able) plastic swipe cards. Arriving guests could bypass the front desk and go straight to their rooms.
- **Lower prices for unlimited e-books :**Being an e-bookworm can get pricey. Two new services will lend you an unlimited number of titles to read on your phone or tablet. online publishing service, offers 100,000-plus titles
- **Facebook goes beyond 'male' and 'female' with new gender options**
- **Lenovo sales hit $10.8 billion as company eyes smart phone biz :**Lenovo posted record sales of $10.8 billion on Thursday and emphasized plans to expand aggressively in China -- the world's biggest smart phone market . The Chinese company said net income grew by 30% to $265.3 million in the final three months of the year. Operating profit rose by 37%, clocking in at $334.2 million for the same period.
- **Nokia Set to Introduce First Smartphone With Android :** It has been reported that NOKIA will be unveiling its first ever Android smart phone at the Mobile World Conference in Spain later this month . This news comes shortly after the cellular maker completed the $7.4 billion sale of its smart phone unit to Microsoft.
- **How Indians have contributed to technology :** Satya Nadella's elevation as CEO of Microsoft marks the acme of global corporate leadership attained in recent years by first generation Indian migrant.
- **BBM gets voice calling on Android , iphone:** BlackBerry has started rolling out a new version of BBM or Blackberry manager, its popular instant messaging app for Android, iPhone and BlackBerry users. BBM users on Android and iPhone will now be able to make free voice calls to BBM contacts over a Wi-Fi or cellular data connection with BBM Voice, as well as access BBM Channels, a broadcast service.
- **Voice calling mobile app Viber sold for $900 million:** Japanese e-commerce company, Rakuten said that it will buy instant messaging app provider Viber Media for $900 million, hoping to tap the Cyprus-based company's rapidly expanding business in

jims Jagan Institute of Management Studies




JIMS IT FLASH

New Microsoft CEO - Satya Nadella

Microsoft Corp. announced on **Feb 4,2014** that its Board of Directors has appointed **Satya Nadella** as **Chief Executive Officer** and **member of the Board of Directors** effective immediately. Bill Gates would step down as chairman and advise the new CEO on technology, marking an epochal change of control at the company that drove the PC revolution.

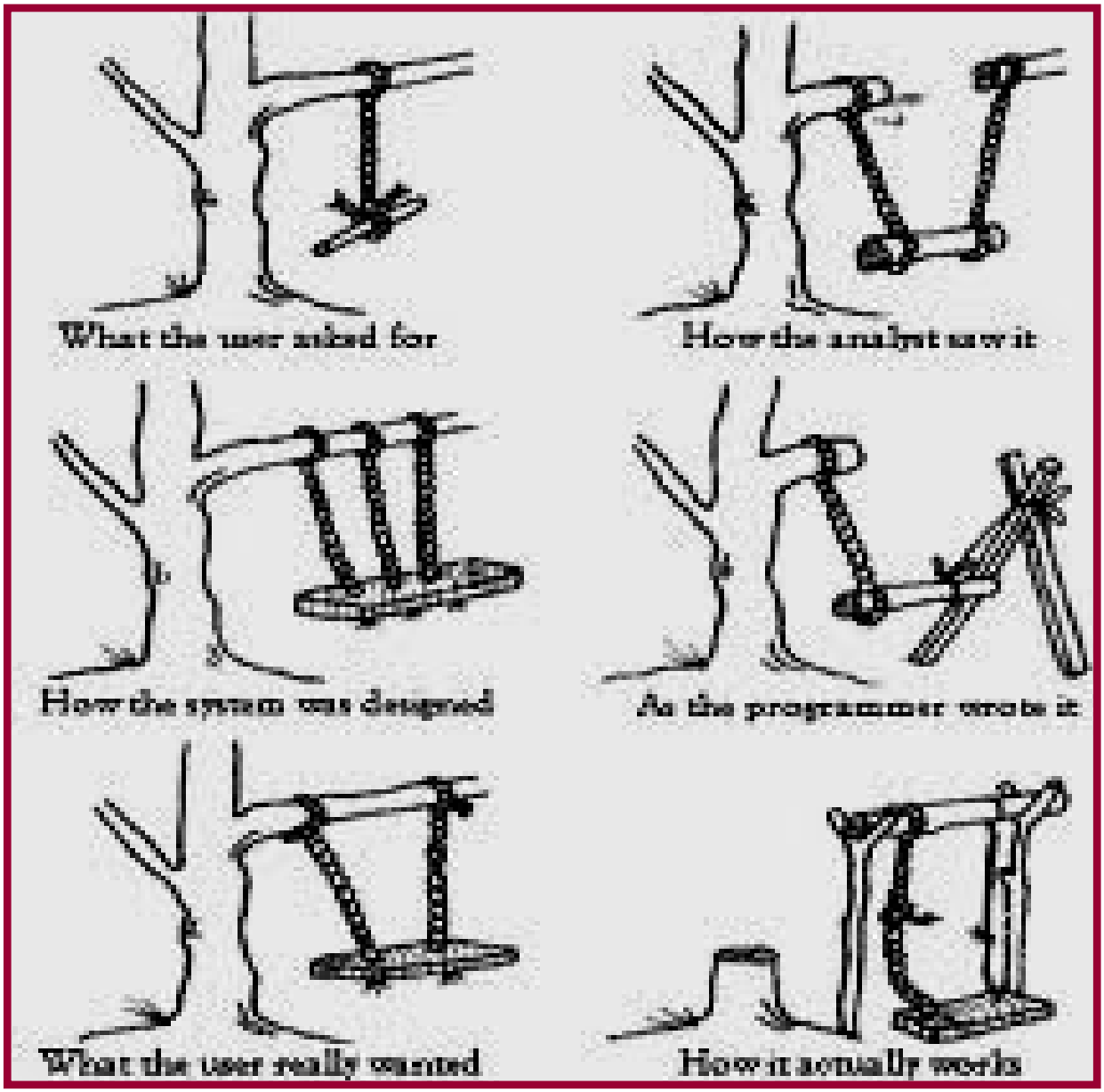
Nadella, a 46-year old born in India who led the creation of Microsoft's Internet-based, or "cloud" computing services, is only Microsoft's third CEO in 39 years, taking over from Steve Ballmer, who inherited the job from Gates in 2000.Since joining the company in 1992, Nadella has spearheaded major strategy and technical shifts across the company's portfolio of products and services, most notably the company's move to the cloud and the development of one of the largest cloud infrastructures in the world supporting **Bing, Xbox, Office and other services**.

Intel wants to be the 'operating system' for big data

Intel is continuing to build out its array of software tools for the Hadoop open-source big data processing framework, with an emphasis on the security and reliability features demanded by large enterprises.

A Data Platform tools suite will become available in the next quarter as a free-of-charge but self-supported Enterprise Edition, as well as a subscription Premium Edition that provides features such as proactive security fixes, regular enhancements and live support.

Intel is competing with the likes of Horton works and Cloud era and others in the commercial Hadoop market. The rise of such vendors underscores the fact that Hadoop is "really at a crossroads," said Jason Fedder, general manager of channels, marketing and business operations for Intel's data center software division. Linux only took off once companies began investing in hardening its features and also "coalesced around keeping it open," he added. "The same thing's happening in the big data domain."



It's not that we use technology, we live technology

-Godfrey Reggio

To be a professional means getting things done right in the first instance. Thus a computer professional is one whose program executes correctly in the first instance without errors.

-V.B Aggarwal

Editorial Team:-