Vol 6 Issue 10 April 2017 ISSN No :2231-5063

International Multidisciplinary Research Journal

Golden Research Thoughts

Chief Editor
Dr.Tukaram Narayan Shinde

Publisher
Mrs.Laxmi Ashok Yakkaldevi

Associate Editor
Dr.Rajani Dalvi

Honorary
Mr.Ashok Yakkaldevi

Welcome to GRT

RNI MAHMUL/2011/38595 **ISSN No.2231-5063**

Golden Research Thoughts Journal is a multidisciplinary research journal, published monthly in English, Hindi & Marathi Language. All research papers submitted to the journal will be double - blind peer reviewed referred by members of the editorial board.Readers will include investigator in universities, research institutes government and industry with research interest in the general subjects.

Regional Editor

Manichander Thammishetty
Ph.d Research Scholar, Faculty of Education IASE, Osmania University, Hyderabad

International Advisory Board

Kamani Perera
Regional Center For Strategic Studies, Sri Lanka

Janaki Sinnasamy
Librarian, University of Malaya

Romona Mihaila
Spiru Haret University, Romania

Delia Serbescu
Spiru Haret University, Bucharest, Romania

Anurag Misra
DBS College, Kanpur

Titus PopPhD, Partium Christian University, Oradea,Romania

Mohammad Hailat
Dept. of Mathematical Sciences, University of South Carolina Aiken

Abdullah Sabbagh
Engineering Studies, Sydney

Ecaterina Patrascu
Spiru Haret University, Bucharest

Loredana Bosca
Spiru Haret University, Romania

Fabricio Moraes de Almeida
Federal University of Rondonia, Brazil

George - Calin SERITAN
Faculty of Philosophy and Socio-Political Sciences Al. I. Cuza University, Iasi

Hasan Baktir
English Language and Literature Department, Kayseri

Ghayoor Abbas Chotana
Dept of Chemistry, Lahore University of Management Sciences[PK]

Anna Maria Constantinovici
AL. I. Cuza University, Romania

Ilie Pintea,
Spiru Haret University, Romania

Xiaohua Yang
PhD, USA

.....More

Editorial Board

Pratap Vyamktrao Naikwade
ASP College Devrukh,Ratnagiri,MS India

R. R. Patil
Head Geology Department Solapur University,Solapur

Rama Bhosale
Prin. and Jt. Director Higher Education, Panvel

Salve R. N.
Department of Sociology, Shivaji University,Kolhapur

Govind P. Shinde
Bharati Vidyapeeth School of Distance Education Center, Navi Mumbai

Chakane Sanjay Dnyaneshwar
Arts, Science & Commerce College, Indapur, Pune

Awadhesh Kumar Shirotriya
Secretary,Play India Play,Meerut(U.P.)

Iresh Swami
Ex - VC. Solapur University, Solapur

N.S. Dhaygude
Ex. Prin. Dayanand College, Solapur

Narendra Kadu
Jt. Director Higher Education, Pune

K. M. Bhandarkar
Praful Patel College of Education, Gondia

Sonal Singh
Vikram University, Ujjain

G. P. Patankar
S. D. M. Degree College, Honavar, Karnataka

Maj. S. Bakhtiar Choudhary
Director,Hyderabad AP India.

S.Parvathi Devi
Ph.D.-University of Allahabad

Sonal Singh,
Vikram University, Ujjain

Rajendra Shendge
Director, B.C.U.D. Solapur University, Solapur

R. R. Yalikar
Director Managment Institute, Solapur

Umesh Rajderkar
Head Humanities & Social Science YCMOU,Nashik

S. R. Pandya
Head Education Dept. Mumbai University, Mumbai

Alka Darshan Shrivastava
Shaskiya Snatkottar Mahavidyalaya, Dhar

Rahul Shriram Sudke
Devi Ahilya Vishwavidyalaya, Indore

S.KANNAN
Annamalai University,TN

Satish Kumar Kalhotra
Maulana Azad National Urdu University

Address:-Ashok Yakkaldevi 258/34, Raviwar Peth, Solapur - 413 005 Maharashtra, India
Cell : 9595 359 435, Ph No: 02172372010 Email: ayisrj@yahoo.in Website: www.aygrt.isrj.org

ISSN: 2231-5063 IMPACT FACTOR : 4.6052(UIF) VOLUME - 6 | ISSUE - 10 | APRIL - 2017

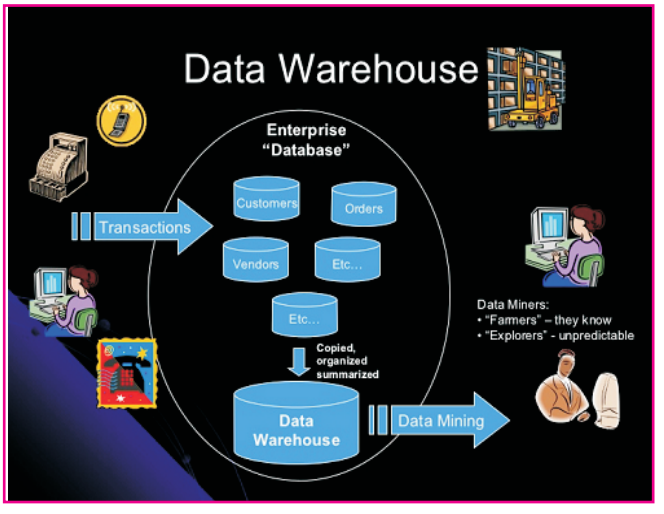
GRT GOLDEN RESEARCH THOUGHTS

DATA WAREHOUSING AND DATA MINING

Prof. Chetana C. Taware
Department Of BBA(CA) , M.S.K.College, Someshwarnagar.

ABSTRACT

Organisations are today suffering from a malaise of data overflow. The developments in the transaction processing technology has given rise to a situation where the amount and rate of data capture is very high, but the processing of this data into information that can be utilised for decision making, is not developing at the same pace. Data warehousing and data mining (both data & text) provide a technology that enables the decision-maker in the corporate sector/govt. to process this huge amount of data in a reasonable amount of time, to extract intelligence/knowledge in a near real time.



The data warehouse allows the storage of data in a format that facilitates its access, but if the tools for deriving information and/or knowledge and presenting them in a format that is useful for decision making are not provided the whole rationale for the existence of the warehouse disappears. Various technologies for extracting new insight from the data warehouse have come up which we classify loosely as "Data Mining Techniques".

Our paper focuses on the need for information repositories and discovery of knowledge and thence the overview of, the so hyped, Data Warehousing and Data Mining.

KEY WORDS: *Data Warehousing , Data Mining , decision making.*

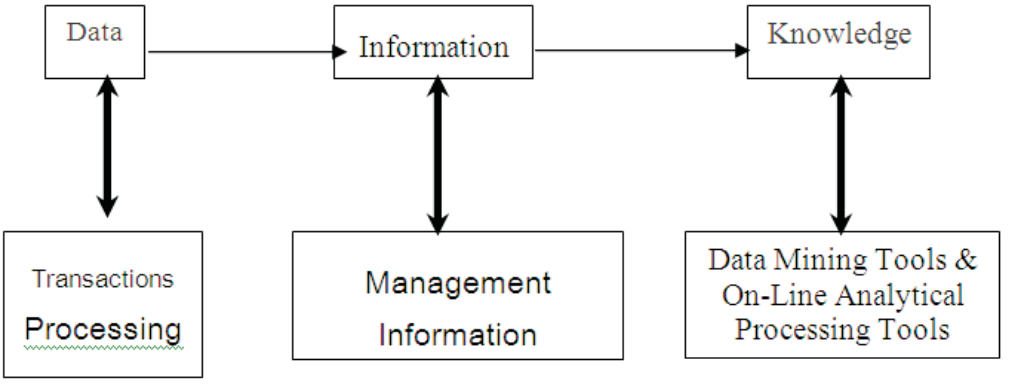
INTRODUCTION

Organizations have lately realized that just processing transactions and/or information's faster and more efficiently, no longer provides them with a competitive advantage vis-à-vis their competitors for achieving business excellence. Information technology (IT) tools that are oriented towards knowledge processing can provide the edge that organizations need to survive and thrive in the current era of fierce competition. The increasing competitive pressures and the desire to leverage information technology techniques have led many organizations to explore the benefits of new emerging technology – viz. "Data Warehousing and Data Mining". What is needed today is not just the latest and updated to the nano-second information, but the cross-functional information that can help decisions making activity as "on-line" process.

EVOLUTION OF INFORMATION TECHNOLOGY TOOLS:

The evolution of the information systems characterize the evolution of systems from data maintenance

systems, to systems that transform the data into "information" for use in the decision making process. These systems supported the information acquisition from the database of transactional data. The managerial knowledge acquisition function is/was not directly supported by these systems. The evolution of new patterns in the changing scenario could not be provided by these systems directly, the planner was supposed to do this from experience.



The Transformation of Data into Knowledge and associated tools.

Warehouse with a database:

And, these days, change is occurring at an ever-increasing rate. A key challenge is implementing an information infrastructure that allows your company to rapidly respond to change. One solution to this challenge is the data warehouse.

Data warehousing is an information infrastructure based on detail data that supports the decision-making process and provides businesses the ability to access and analyze data to increase an organization's competitive advantage.

Data warehousing is a process, not an off-the-shelf solution you buy, but hardware--database and tools integrated into an evolving information infrastructure--that changes with the dynamics of the business.

What it actually is?

A data warehouse stores current and historical data. This data is taken from various, perhaps incompatible, sources and stored in a uniform format. Several tools transform this data into meaningful business information for the purpose of comparisons, trends and forecasting. Data in a warehouse is not updates or changed in any way, but is only loaded and accessed later on. Data is organized according to subject instead of application.

Data – Warehouse Functions:

The main function behind a data warehouse is to get the enterprise-wide data in a format that is most useful to end-users, regardless of their locations. Data warehousing is used for:

1) Increasing the speed and flexibility of analysis.
2) Providing a foundation for enterprise-wide integration and access.
3) Improving or re-inventing business processes.
4) Gaining a clear understanding of customer behavior.

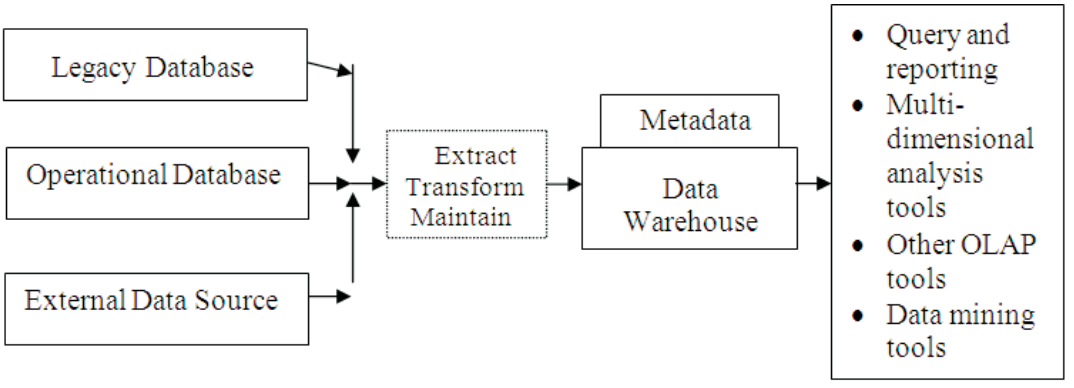
Data Warehouse Architecture:

Each implementation of a data warehouse is different in its detailed design (a schematic high-level of the architecture and its components is given in the figure below), but all are characterized by a handful of the following key components:

A data model to define the warehouse contents.

A carefully designed warehouse database, whether hierarchical, relational, or multidimensional. While choosing

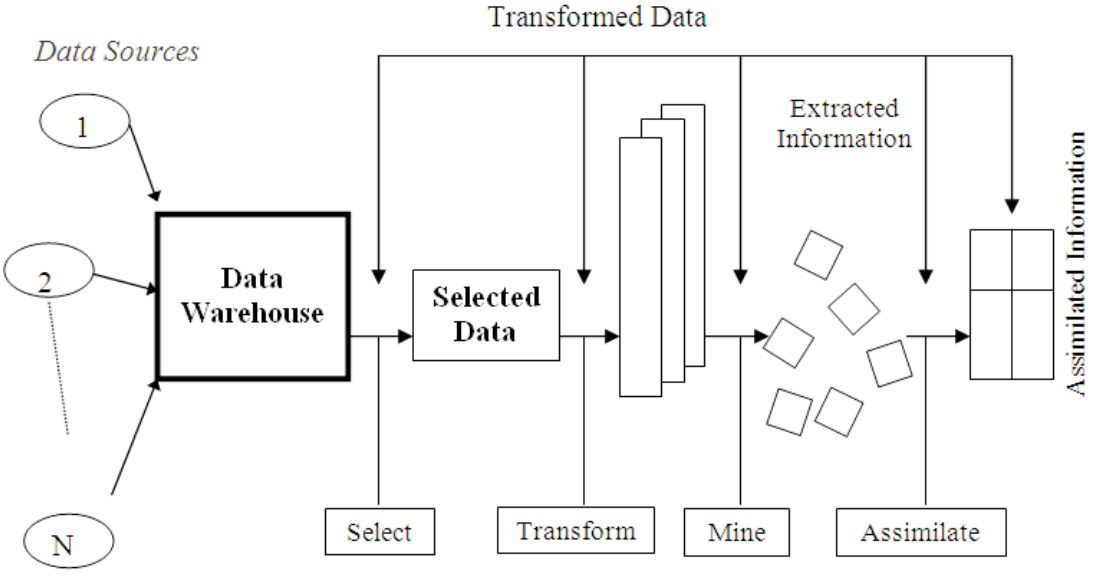
a DBMS it must be kept in view that the database management system should be powerful enough to handle huge amount of data running up to terabytes.
A front end for Decision Support System (DSS) for reporting and for structured and unstructured analysis.



Schematic view of the Data Warehouse Architecture.

Data Mining:

Data base mining or Data mining (DM) (formally termed Knowledge Discovery in Databases – KDD) is a process that aims to use existing data to invent new facts and to uncover new relationships previously unknown even to experts thoroughly familiar with the data. It is like extracting precious metal (say gold etc.) and/or gems, hence the term "mining", it is based on filtration and assaying of mountain of data "ore" in order to get "nuggets" of knowledge. The data mining process is diagrammatically exemplified in Figure below



The Data Mining Process.

Data Mining and Data Warehousing:

1) The goal of a data warehouse is to support decision making with data.
2) Data mining can be used in conjunction with a data warehouse to help with certain types of decisions.
3) Data mining can be applied to operational databases with individual transactions.

4) To make data mining more efficient, the data warehouse should have an aggregated or summarized collection of data.
5) Data mining helps in extracting meaningful new patterns that cannot be found necessarily by merely querying or processing data or metadata in the data warehouse.

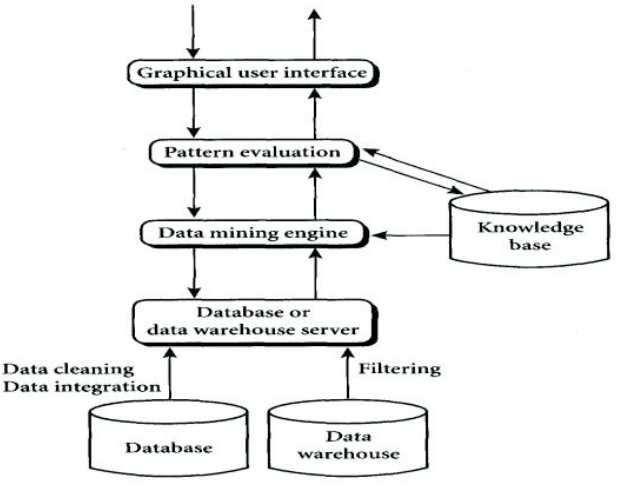
Data Mining as a Part of the Knowledge Discovery Process:

Knowledge Discovery in Databases, frequently abbreviated as KDD, typically encompasses more than data mining. The knowledge discovery process comprises six phases:

Data selection Data about specific items or categories of items, or from stores in a specific region or area of the country, may be selected.
Data cleansing process then may correct invalid zip codes or eliminate records with incorrect phone prefixes.
Enrichment typically enhances the data with additional sources of information.
Data transformation and encoding may be done to reduce the amount of data.
Goals of Data Mining and Knowledge Discovery

The goals of data mining fall into the following classes:

Prediction: Data mining can show how certain attributes within the data will behave in the future.
Identification: Data patterns can be used to identify the existence of an item, an event, or an activity.
Classification: Data mining can partition the data so that different classes or categories can be identified based on combinations of parameters.
Optimization: One eventual goal of data mining may be to optimize the use of limited resources such as time, space, money, or materials and to maximize output variables such as sales or profits under a given set of constraints.



CONCLUSION:

A data warehouse takes the organisations operational data, historical data and external data. consolidates it into a separately designed database (which can either be relational or multi-dimensional in nature) .

Manages it into a format that is optimised for end users to access and analyse.

When a data warehouse has been constructed, it provides a complete picture of the enterprise. It provides an unparalleled opportunity to the management to learn about their customers.

The data warehouse technology together with online transaction processing and data mining, allows the management to provide better customer service, create greater customer loyalty and activity, focus customer acquisition and retention of the most profitable customer, increase revenue, reduce operating cost;

provides tools that facilitate sounder decision making; improves worker/management knowledge and productivity; spares the operational database from ad-hoc queries with the resulting performance degradation and clears the legacy database system, while moving the corporate system architecture forward.

With the incorporation of new data delivery and presentation techniques, like hypertext mark up language (HTML), Open Database Connectivity (ODBC) etc. the database mining (Data & Text) operation has gained wide spread recognition as a viable tool for business intelligence gathering. Advances in the document mining technology (database mining of free form text/data, in contrast to the "classical" approach to data mining of fixed length records) are making the data mining technology more powerful.

Last but never the least, the Internet has emerged as the largest data warehouse of unstructured and free form data. The new technologies are geared towards mining this great data warehouse.

REFERENCES:

1)Computers Today – Smart Facts about Data Warehousing by Atanu Roy
2)Using Information Technology by William Sawyers Hutchinson
3)Data Base System Concepts by Silberschatz, Korth and Sudharshan
4)http://www.infogoal.com/
5)http://www.cse.buffalo.edu/
6)http://www.datawarehousingonline.com.

/

Publish Research Article
International Level Multidisciplinary Research Journal
For All Subjects

Dear Sir/Mam,

We invite unpublished Research Paper,Summary of Research Project,Theses,Books and Book Review for publication,you will be pleased to know that our journals are

Associated and Indexed,India

- International Scientific Journal Consortium
- OPEN J-GATE

Associated and Indexed,USA

- EBSCO
- Index Copernicus
- Publication Index
- Academic Journal Database
- Contemporary Research Index
- Academic Paper Databse
- Digital Journals Database
- Current Index to Scholarly Journals
- Elite Scientific Journal Archive
- Directory Of Academic Resources
- Scholar Journal Index
- Recent Science Index
- Scientific Resources Database
- Directory Of Research Journal Indexing

Golden Research Thoughts
258/34 Raviwar Peth Solapur-413005,Maharashtra
Contact-9595359435
E-Mail-ayisrj@yahoo.in/ayisrj2011@gmail.com
Website : www.aygrt.isrj.org