Application of Data Warehouse in Real Life: State-of-the-art Survey from User Preferences' Perspective

Muhammad Bilal Shahid, Umber Sheikh, Basit Raza, Munam Ali Shah, Ahmad Kamran, Adeel Anjum
Department of Computer Science
COMSATS Institute of Information Technology
Islamabad, Pakistan

Qaisar Javaid
Department of Computer Science and Software Engineering
International Islamic University
Islamabad, Pakistan

***Abstract*—In recent years, due to increase in data complexity and manageability issues, data warehousing has attracted a great deal of interest in real life applications especially in business, finance, healthcare and industries. As the importance of retrieving the information from knowledge-base cannot be denied, data warehousing is all about making the information available for decision making. Data warehouse is accepted as the heart of the latest decision support systems. Due to the eagerness of data warehouse in real life, the need for the design and implementation of data warehouse in different applications is becoming crucial. Information from operational data sources are integrated by data warehousing into a central repository to start the process of analysis and mining of integrated information and primarily used in strategic decision making by means of online analytical processing techniques (OLAP). Despite the applications of data warehousing techniques in number of areas, there is no comprehensive literature review for it. This survey paper is an effort to present the applications of data warehouse in real life. It focuses to help the scholars knowing the analysis of data warehouse applications in number of domains. This survey provides applications, case studies and analysis of data warehouse used in various domains based on user preferences.**

Keywords*—*Data warehouse (DW); Data warehouse applications; Decision support systems; OLAP; Preference based

I. INTRODUCTION

Operational and transactional systems are the new generation systems which are different from 1970's decision support systems (DSS) [1]. In order to complete the life cycle, DSS needs the shadow of a Data Warehouse (DW). A DW pools the available data which is spread all over the organization, and makes a unify pool (like data structure) having the presence of similar and linked formats [2].

Data warehousing takes off in the 1980s as an answer to the very little or no availability of information propagated by online application systems, online applications were praised by a very limited domains of users, and integration was not there even [3]. Historical data kept by online applications are very little as they deposit their historical data for high performance in faster way. Thus organizations hold very little information as compared to data [3].

Inmon drafted that for building a DW most organizations starts with an architecture. "Inmon talks about DW that there is still a way long confusion as what it really is". Bill Inmon [3], [4 p.31], said that the description to a DW was and still is today. "A source of data that is subject-oriented, integrated, nonvolatile, and time-variant for the purpose of management's decision processes".

With the thirst and huge need for large blocks of information, DW gain much importance and became an essential strategy component for medium and large organizations. Timely and accurately decision making at management level becomes difficult due to the incapability of traditional databases to handle increasing demands of online information access, retrieval, maintenance and update efficiently which greatly impacts every industry [5]. So companies start seeking the solution for all their problems and adopt DW technology.

With sharp and harder competition, enterprises are targeting in availing fast and pinpoint information to have best decisions. Furthermore, with the thirst for huge chunks of information, enterprises' traditional DB (database) is off no use of smartly managing the increasing needs of online information update, access, maintenance, and retrieval. This lagging impressively effects the efficiently and effectively usage of internal data by the management to hold decision-making in time. As a result, to search for various ways and means to store, access, handle, and utilize the huge chunks of data in an effective manner, is the main concern of every business [5].

Organizations requires a database system for their daily decision making, with better adaptability, top flexibility, and best support. Considering the past decade, the educational (academia) side and the industry side, both have progressively plated different layouts to solve the problems and to present solution to craft an aforementioned system [5]. Adopting the data warehouse technology is one of the solutions to that. DW was defined by Inmon [3, 4] as, ''pooling data from multiple separate sources to construct a main DW". Proper data-analyzing tools can be used by different users to analyze and store required data.

Data Warehouse's purpose is to take large data from heterogeneous sources and furnish them in known formats that helps in understanding and for making smart decisions [6]. The Benefits linked to the DW applications include the region of time saving, with the availability of clean and handful of information, tough and exact decisions making in accordance with the improvement of processes related to business and to help achieving strategic business objectives [2, 4, 5, 6].

Realizing the need after researching literature and for further exploring on this research article, taking in account the importance of the applications of DW in real life and the shortfall of the factual research, we have all the concreate reason to explore the most applications of DW in real life. In this paper we discussed different applications of DW in real life along with available case studies. Its sections as follows; Section 2 presents DW technology. Section 3 presents the applications of data warehousing in different domains. Section 4 provides a tabular and descriptive view of different case studies under the umbrella of government and business categories. Section 5 provides a brief usage analysis of Data Warehouse applications. Finally, conclusion is presented in Section 6.

II. DATA WAREHOUSE TECHNOLOGY

Devlin and Murphy was the pioneer to present the concept of data warehousing [7]. Read-only database that is capable of storing historical datum for operating was suggested. It offers a variety of integration tools. Users can find and query what they want for supporting decision. Time-variant, non-volatile, integrated and subject oriented are the four key attributes of data warehouse defined by Inmon [8]. With the presence of different attributes, datum is encapsulated in "subject oriented" attribute, which is build and is combined in multiple angles. Talking about an example in a traditional system, a datum for point of sale (POS) might be not same as of other sale systems [4, 8]. The data are hidden separately as a one unit, irrespective of what the under used system is. "Subject oriented" entity tells about the datum that it is build and combined through different angles as said by different authors. Taking in account a traditional system, for example, "custom datum viewed from a POS for sure having different angles from other related sale systems (machines)". Whatever system is used, we have single topic from isolated custom data, by usage of DW [5, 8]. Consistency of data will not be present as it is being integrated, converted and/or extracted by different tools, thus getting an integrated data.

Any variation, in the form of result, can be very important, if the focus of system is on a "real-time" attribute, this includes in the characteristics of time variant. The need for related time and portions of time information is needed by the data stored in data warehouse for future querying. The massive past non-volatile data is held by data warehouse, by which we can perform analysis, prediction and discovery with the positivity of effectiveness, reliability and accuracy. Through modification, we ensure the perseverance of best quality, when data are uploaded in data warehouse. The Inmon's [8] definition of data warehouse has modified and/or redefined by many authors in recent span of time [9, 10, 11, 12, 13, 14]. The scope of data warehouse domain has broadened by different definitions, but is still align with Inmon's definition. According to the different definitions, DW could be summed as, "DW pools daily, both externally and internally "transaction-oriented" enterprise data, and then summed, divide in categories and hold (store) massive data from past (historical) for more computation, forecast, analysis, and discovery of data patterns". Obtained data are linked to non-modified, statistics, and stored in DW for longer period. Furthermore, for analyzing and making decisions they are integrated, time-oriented, and effectively used. We can find at least one chapter related to data warehouse in all major books of databases. As the existence of data warehouse exceeds over 20 years, we can get many useful resources of its design and implementation [15, 16].

A. Data warehouse architecture

Figure 1 shows a general view of data warehouse architecture acceptable across all the applications of data warehouse in real life. Every application of data warehousing include extraction of the informatics data from the key system with using as minor resources as it can, transformation of that data by applying a set of rules from source to the target and fetching (loading) the related data into a DW (called ETL process). Some of the areas DW architecture holds it importance are technical related design, data related design, and hardware and software related design [5, 6, 12].

Design domain of DW architecture widely grouped into enterprise DW design and data mart related design. The enterprise DW is the blend of those adoptive data marts [17]. A data mart is considered to be a tinier version linked to a DW but it aimed on specific subjects. Top-down along with bottom-up techniques linked with data design are followed by data marts [17, 18]. The general DW architectures include the presence of enterprise DW, along with "data marts", linked to the "distributed warehouses", and "operational related" data rooms with data marts, or any mixture to those [4, 17, 18, 19, 20].

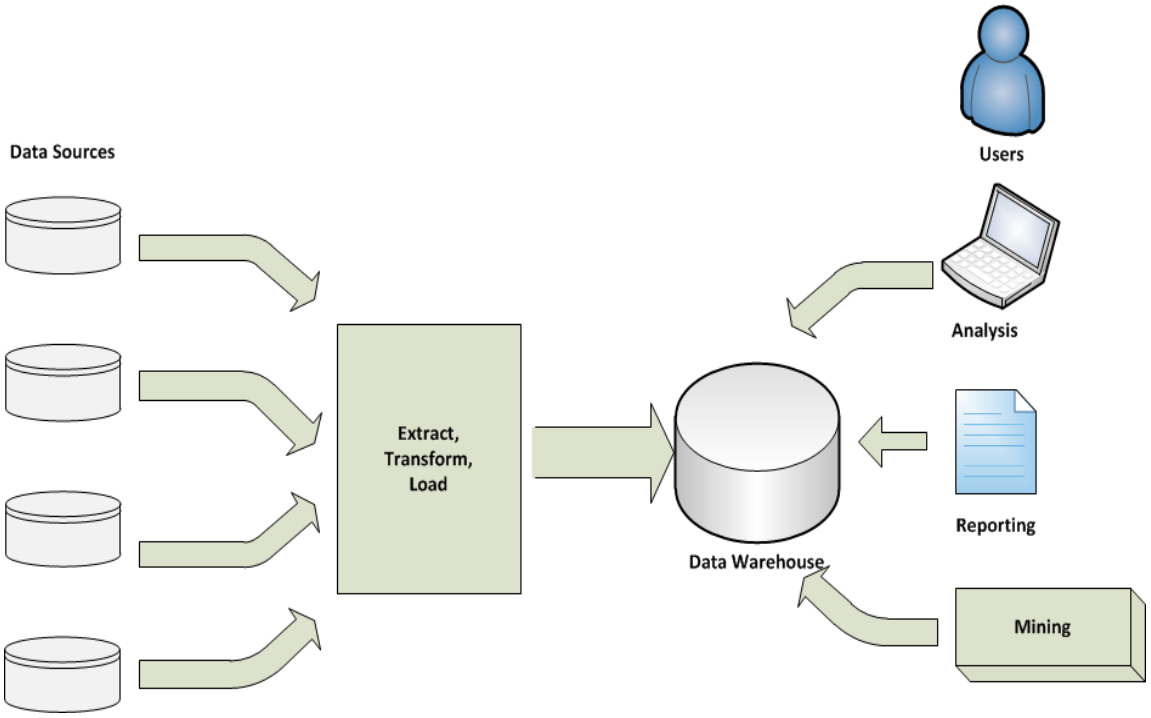

Fig. 1. Data Warehouse Architecture

Figure 1 deeply shows a standard DW architecture. There are many sayings on which architecture best suits the design and implementation. Authors [3, 4, 8, 11, 17] consider Inmon and Kimball as the top of every other, taking in account Sen and Sinha pushed 15 separate methodologies to DW architecture [20]. Figure 1 shows a color print of a general DW architecture. Data are propagated from "operational DBMS" and it is processed by the process called, "extraction, transformation and loading (ETL)" into the DW or data marts. The process or body of the ETL gives a unique data room for decision-making so we always have one unit for it. ETL is said to be the most difficult process of DW construction. Up-to-date and many powerful tools are available to assist this area, but along with artificial tools real human administration is important and for that we require front panels to assist human administrators. Once all the aforementioned processes are completed and the data gathers in DWs or data marts, then we came up with the tools called "online analytical processing (OLAP)". OLAP provides the data into graphical, and in multidimensional prints to help users to query, dig or mine and analyze the data [6, 20, 21].

State of the art research papers have also been published stating the overview, frameworks and up to date practices [22, 23]. Failures parts are also handled by many researchers [24]. The most important thing in making a DW is selecting the best architecture. Extraction from relational database, moving to Transformation, and at the end loading (ETL process), include in the data warehousing environment. It also includes Online Analytical Processing (OLAP) plus the client analysis tools [5, 23].

The process of data warehousing starts from propagation of data from main (original) format passed to a "dimensional data" region for storages purpose, it handles a huge amount of work, clock and money. Implementation and designing of a DW demands cost and is quite critical, for handling those critical tasks, tons of tools related to data extraction, data cleaning and load utilities are present to aide in. Data integration is considered to be the top and most useful part of the DW [1, 5, 6].

III. Applications of Data Warehouse in Real Life

Importance of DW cannot be denied due to its benefits because decisions at management level will no longer need to be taken on the limited and inaccurate data and it also helps the companies to avoid different challenges. So it becomes the need of every individual company to implement data warehouse.

It is estimated that by 2020 around 200% more devices will join the Internet and share data. DW strongly depends upon devices and inter linked data. The more interlinked devices are, the more powerful and useful DW. According to the forecast by many organization [25, 26] by 2016 around 6.4 billion connected peers will join the room globally, an increase of 30% from 2015. Cisco and other research agencies [25, 26] think that approximately 20 - 50 billion devices will be connected by 2020, (see Figure 2) [25, 26].

Other side of the picture is that cost will increase too. If we talk about spending on hardware, the applications related to consumer will hit to $546 billion by the end of 2016; apart from that the usage of connected items in the organization will be somewhere around $868 billion by the end of 2016 (refer to Figure 3) [25, 26].

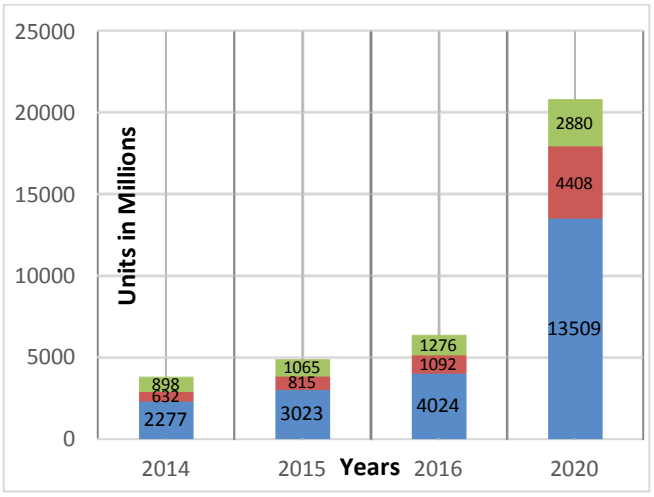

Fig. 2. Number of Units

Talking about relevance of DW, it is said that few of the application areas holds the presence and integration of data throughout the enterprise, furthermore a fast decisions on live and previous (historical) data, give specific information for those systems that are defined loosely. Figure 4 shows the cycle of real life applications of data warehouse in different fields and how they are interrelated according to user preference.

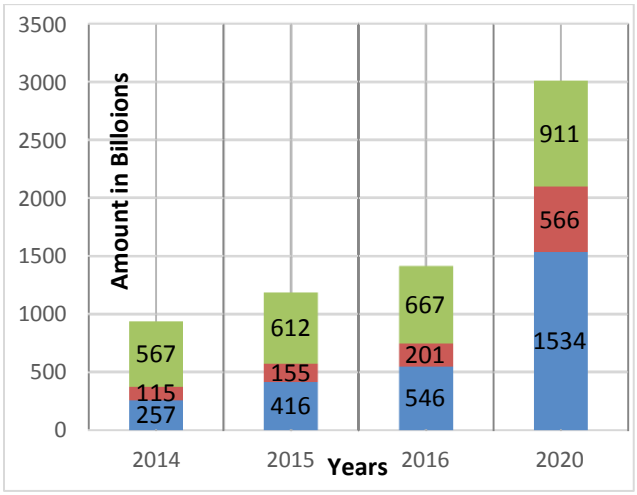

Fig. 3. Cost in Billions

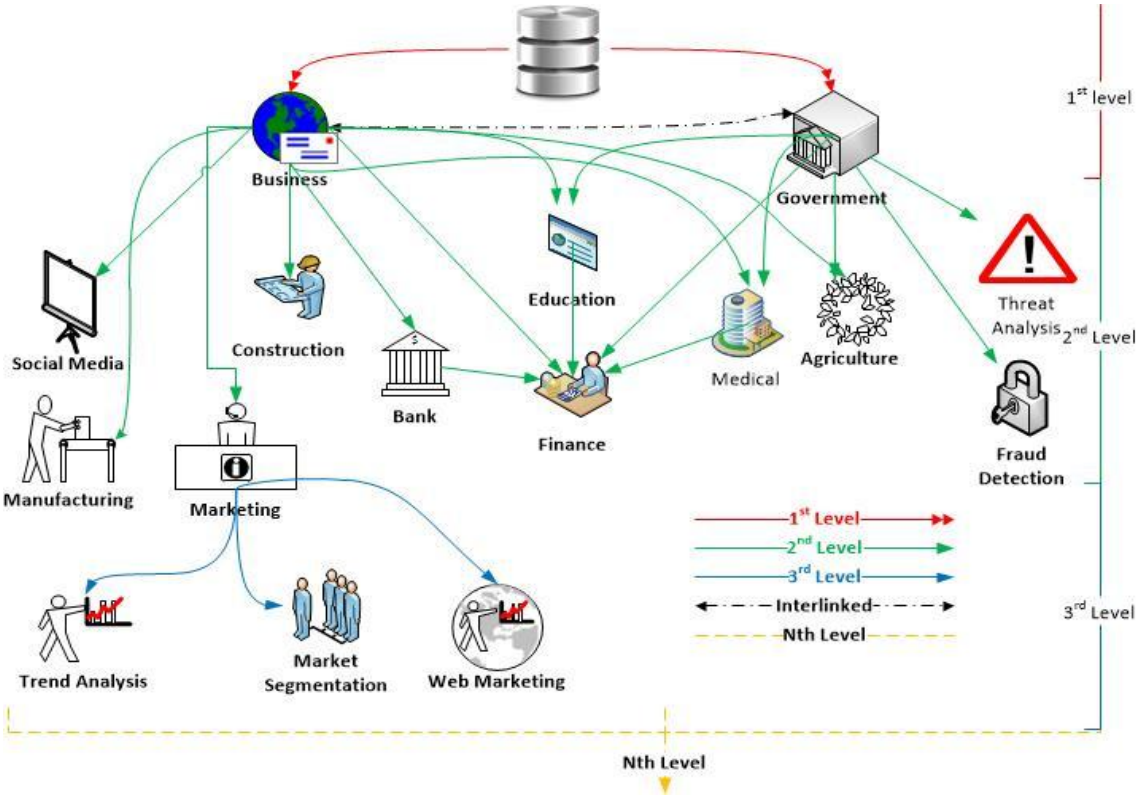

Fig. 4. Applications of Data Warehouse in Real Life

We have suggested a generic layout of interlinked applications of data warehouse (DW). As we can see that different levels are defined. These levels are associated with the hierarchy such that first level is the core component. The first level is always be a central DW (core system(s), hardware system(s)). Furthermore, 2nd level is associated with one of the world's top domains (Root level, business and Government). The reason behind selecting Business and Government as top of hierarchy is a handful of literature, and all other domains are encapsulated under them. With the presence of 2nd level all other sublevel gets populated. The 2nd level serves as the only pillar that supports all other domains. 2nd level is said to be a specific level. 3rd level domains are the more general than specific. The Nth level is the most general level that holds all minor to major domains. Figure 5 shows the flow diagram, which moves from specific to general.

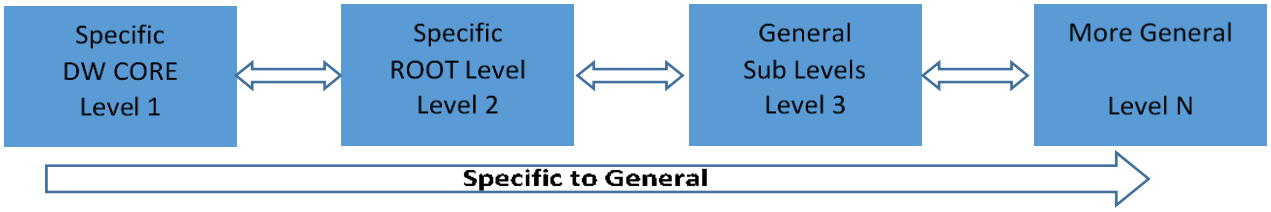

Fig. 5. Specific to General Flow of DW

A. Buisness

Improvement related to decision making and increasing organizational performances are the basic reasons to adopt DW in business [27]. Business holds a key location in applications of data warehouse. All other private and semi-private organizations come under its umbrella.

In DW, for easiness a single repository is used to store data, which is extracted from different databases. This data repository provides forecasting which helps the business personals and business managers. This complete cycle is used to help in identifying the requirements for business and to draft a plan for business [28]. Some of the major to minor fields effecting data warehousing in business are discussed further as shown in Figure 6.

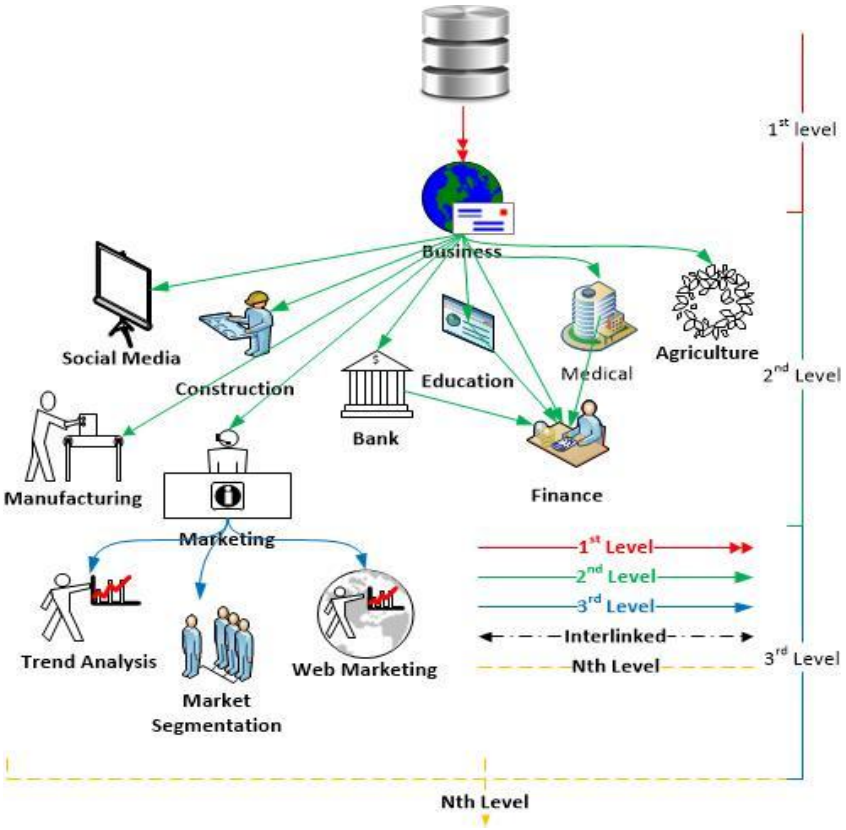

Fig. 6. Business (Application of Data Warehouse)

1) Social media websites

Social media is a great example of data warehousing. Social media industry is emerging and so is the need to implement DW in it. A number of features from Facebook, Twitter and other social media sites are also based on analyzing large data sets [29]. It gathers all data like groups, likes, friends, location mapping etc. and stores it in a single central repository. Although all this information is stored in separated databases but the most relevant and significant information is stored in a central aggregated database [28].

2) Construction (material based industries)

Data warehouse approach in construction industry seems to be efficient in decision making as it provides construction managers the complete internal and external knowledge about available data so that they can measure and monitor the construction performance.

Application of DW in construction industry clearly shows that construction bosses can smartly judge the stock remaining, inventory related trend linked to the materials, the amount and quantity of each material and also the price of all materials [30,

56]. It would also be helpful in reasonable resource allocation to fulfill the required services, maintenance and operation of the systems, allocation of financial budgets, effective managing of investment related long term plans and identification of potential risks [31].

3) Manufacturing Industry

DW plays a vital role in daily house to industrial hold things. Manufacturing industry includes product and process design, scheduling, planning, production, maintenance and huge investments in equipment, manpower and heavy machinery. In this scenario, decisions taken will have wide-ranging effects in terms of profitability and long-term strategic issues. Many industries are trying to convert themselves and many should adopt DW technology rather than traditional decision making so that a warehouse gathers, standardizes and stores data from various applications for improvement in processes and increasing its efficiency as analyzing the data in separate applications is time-consuming. At this stage, some transaction processing systems, which are updated timely, are often hired to propagate the routine business of manufacturing and construction companies [56, 57].

4) Marketing

Every business is not successful without proper marketing and marketing is not successful without knowing the latest trends and demands. Shown in Figure 7 is a general lay out of marketing and its sub domains. Relationship marketing is a new terminology linked with how different businesses handle their customers and the relationships in between that are assets for them and how they can be improved for long-term profitability. DW in marketing is used to examine the patterns of customer's behavior and use this customer information for implementing relationship marketing. They play a vital role in identifying and targeting the profitable customers [32].

Uses of Data Warehousing in marketing area shown in Figure 7 are further categorized as:

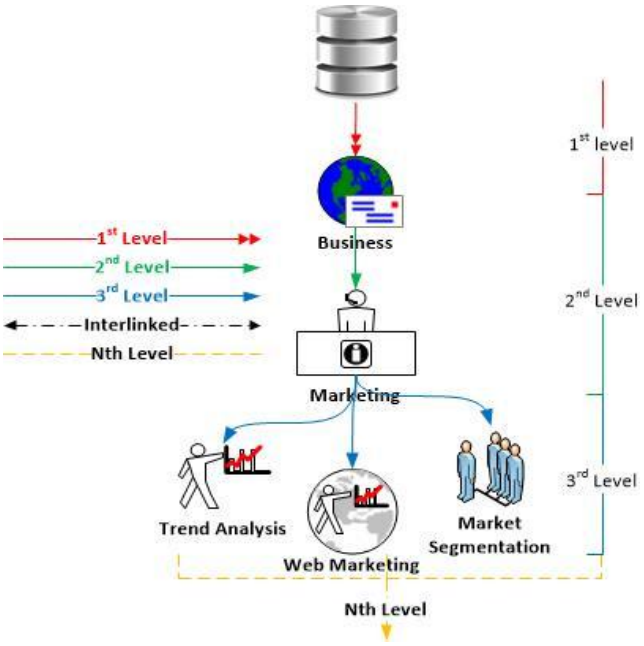

Fig. 7. Marketing (Application of Data Warehouse)

a) Trend Analysis

It is a technique that is used to predict future outcomes from historical results or information. Different medium to large scale enterprises are converting to this. In trend analysis, DW can be used to examine the behaviors of the customer by using historical records over consecutive months.

b) Web Marketing

Web is a hub of billions of devices and around 20 – 50 billion devices till 2020. It refers to a category of advertising that includes any marketing activity conducted online. Facebook, google, and many major to minor such like sites uses web marketing and are relying on latest updated data warehouse.

c) Market Segmentation

Behavior identification is the top most priority of any organization. Market segmentation is the identification of the customer's behavior and common characteristics related to the purchases made against that product of related company. Many organizations are focusing on integrating data warehouse to get best behavior analysis.

5) Banking

The banking industry is categorized as one of the highest information demanding industry in the business world. With the advancement in information technology sector, the role of business intelligence (BI) increases with great number in the process of banking operations [54]. The increased business speed and growing competition has shown the need of banking intelligence dramatically. Bank intelligence is the ability to gather, manage, and analyze a large amount of data on bank customers, products, operations, services, suppliers, partners and all the transactions. As data increases, it becomes difficult to collect, handle and transform it into useful knowledge and DW solves this problem. Many data warehouse flavors are designed for the support of banking industry.

6) Education

DW in education field is becoming popular day by day. Use of DW in educational field presents several potential benefits in making appropriate decisions and for evaluating data in time which is the basic target of DW process. DW provides an integrated and total view of an institute [33]. Most of the related departments use data warehouse as a source of information about faculty and students. DW helps the students in getting their results and notes from a web enabled database quickly through a student portal and last but not the least it helps in decision making by providing current and historical information of the institute.

On a large scale, a DW can integrate the information of different institutes into a single central repository for analysis and strategic decision making.

7) Finance

With the advancement in technology, especially IT industry has opened the doors to the new ways of handling business considering financial systems. Government and Business domain holds equal part in finance. Financial systems may include banks, post offices, insurance companies, income tax and all other tax departments etc. Implementation of data

warehouse in financial industry has several benefits e.g. it can maintain transparency in account opening and transactions. Similarly, government can take decisions against any financial crises. These systems are intelligent enough to spot the defaulters and may act according to the situation. As data warehousing is maintained in this scenario so efficient decision making process can easily be performed. These data warehouses in finance applications can also be used for the analyzation and to have forecasting of different aspects of business, stock and bond performance analysis [34, 58, 60].

B. Government

Amongst the two major sub-divisions of DW industry, government holds equal division. Government can use data warehousing technique in different fields e.g. for searching terrorist profile and threat assessments, in agriculture, in educational industry, in financing department, medical departments and for fraud detection. The telecommunication industry and Banking industry holds many issues related to user frauds. Figure 8 shows application of data warehouse in government departments.

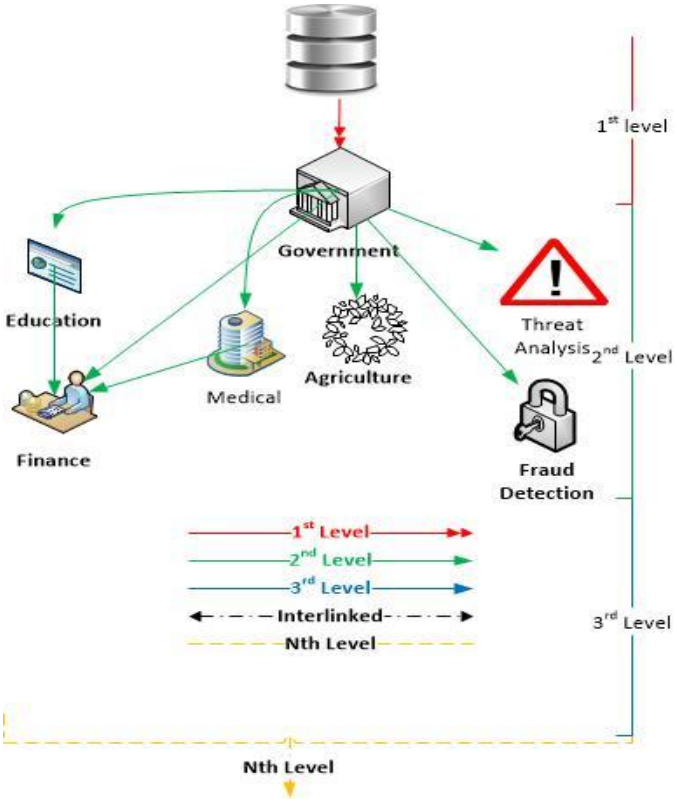

Fig. 8. Government (Application of Data Warehouse)

1) Medical

Medical sector is emerging as the highest DW implementer industry. In health-care, data quality and demand for quality medical services has become increasingly important [55, 59]. Due to the intricacy and variety of medical cum clinical data, the adoption of data warehouses by health care was slow as compared to other fields. Over the past few years it was reported that the usage of DW increased by the administrative and clinical areas. Data warehouses can help in improving the care of specific patients. These health-care institutions are adopting data warehousing for strategic decision making as a decision supporting tool. It provides the tools for acquiring medical data, for extracting the relevant information from that data and finally making this knowledge available to all the concerned persons. Administrative data in data warehouse can help in providing the information about skilled staff needed for a particular treatment and this information further used for the treatment scheduling and to help supporting medical personals in human resources area [36].

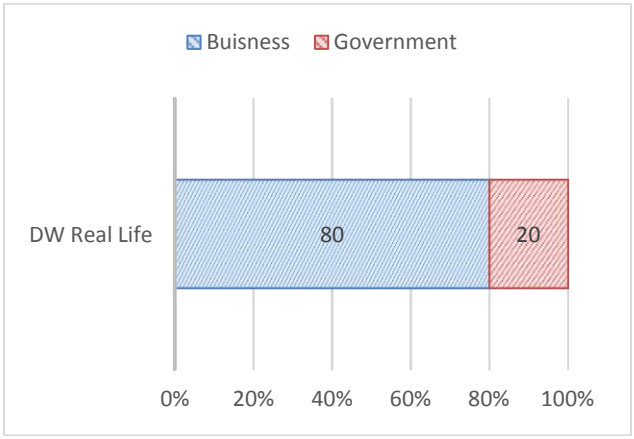

Fig. 9. Percentage wise Contribution of Data warehouse

2) Fraud and Threat detection

Governments are playing their part to detect any threat and fraud caused by ill-minded people. Unfortunately, almost no specific data warehouse implementation that is known is available. Data warehouse access to governments are there, but they need a data warehouse system that is linked with every corner so that threats and terrorists will be monitored.

IV. Case Studies

In this section few case studies are discussed. As discussed earlier data warehouse world is a blend of two parts i.e. business side and government side. Both sides have their own further divisions and any other increment will be added under them. A graphical view is presented in the Figure 9, which is related to the contribution made by business and government domains to DW. It is clearly observed by the survey that 80% of Business and 20% Government related organizations are contributing in the progress of data warehouse.

A. Business

DW in business is now emerging like a hurricane. Around 80% of data warehouse implementation is captured by business. Following are few case studies related to business implementation of data warehouse.

1) Finance

Financial services company (FSC) is considered to be the leading marketer of investment besides banking for products. They implemented DW named as VISION. The user of VISION consists of financial and marketing analysts, managers. It was developed with substantial business and technical goals that can gave a factual and precise picture of best customers of banks and also about most important products [27].

2) Medical

This case study is based on generation of evidence-based guidelines performed by University Health Network (Toronto) which clearly showed that it is authentic, influential and user-friendly to have a DW related to clinic for best strategic decision making. Without this IT support, it would not be imaginable to look for evidence-based medicine as it is difficult for clinicians to gather data for a specific disease [36].

3) Banking

Their research problem is based on the factors that banking industry should consider before and during the adoption of DW technology. Their results revealed the number of banks in Taiwan that adopted this technology and also the architectures that these banks implemented [5].

4) Manufacturing

Large Manufacturing Company (LMC) is making its way to top for production of home related appliances. LMC implemented data warehouse technology as there is a great need to improve the technical infrastructure of the company. Before this, data was scattered in different formats throughout the company and this makes normal and basic functioning difficult for business units. This warehouse provides support to marketing, manufacturing and logistic applications by providing data to dependent data marts [27].

B. Government

Data warehouse in government plays a vital and critical role. Around 20-35% of data warehouse industry is captured by government. Many developing countries are now transferring to the use of data warehouse. Few case studies related to government and usages of data warehouse are as follows.

1) Medical

In Utah and Idaho, Intermountain Healthcare implemented EDW. This healthcare system operates 22 hospitals, 179 clinics, physician offices. This case study is about venous thrombosis patients. Datasets consists of: records of Inpatients, columns of outpatient, financial data linked to or from patient's accounts, data from laboratories related to clinics for the process of imaging and surgery [35] etc. Their DW is updated each night that includes: Large Metadata Repository, Security and auditing infrastructure and Master Reference Data. By using latest information from data warehouse patients with high risk are identified and their reports were sent at every hospital or clinic [35].

2) Finance

Internal Revenue Service is the agency of U.S. that is responsible for tax collection and tax laws imposition. They implemented data warehouse CRIS as there is no way to recoup entity with convinced attribute and perform some analysis on these marked entities. This implemented DW consisted of five domains: business entity, tax returns entity, related to taxpayer transactions entity, peoples' income sources entity and tax payments details entity [27].

C. Tabular view of case studies

Table 1 is the tabular view of all aforementioned case studies.

TABLE I. TABULAR VIEW OF CASE STUDIES

Ref. No	Domain	Architecture	Methodology	Dataset	Method Description	Strength	Limitation
[35]	Healthcare	Enterprise Data Warehouse	Questionnaire	• 22 hospitals, • 179 clinics, • physician offices, home healthcare in Utah and Idaho	• A computer program was for monitoring. • Patient's identification according to score. • Update EDW. • Evaluation.	• Proposed framework can be reused easily for new applications.	• No enterprise database with daily updated patient lists.
[36]	Healthcare	Data Warehouse based approach (for integration of data sources) + Data mining techniques	From published clinical evidence i.e. books, magazines, journals, healthcare, protocols, clinical trials.	University Health Network (Toronto)	• Generation of treatment rules based on clinical evidence. • Data loading to DWH. • Trends identification by data mining techniques. • Rules examination and approval. • Judgments about recommendations and for improving patients care.	• Reliable. • Powerful. • User-friendly platform.	• External evidence-based knowledge is not enough but needs to be adjusted according to patient's health and preferences
[5]	Banking	Data Warehouse	Questionnaire	Banks of Taiwan i.e. from 50 banks and 30 valid responses with response rate 60%.	• Questionnaire with six sections. • Analysis about banks that adopted, in process of adopting or abandoned DWH technology. • Analysis about	• Identification of factors that can affect DWH adoption. • Facilitate implementation in global or overseas branches.	• Limited to domestic banks. • Approach is restricted to banking industry only. • Limited samples.

					architecture of DWH adopted by banks.		
[27]	Finance	VISION data warehouse	Interviewing of employees, examining documents and video tapes of key events.	Financial services company (FSC, US)	• In first phase, top revenue producing customers are identified. • Second phase provided, profitable information for all bank's customers and products.	• Gives more clear and accurate picture of most important customers and products.	• Limited to Critical Financial data • Limited data samples
[27]	Finance	Compliance Research IS (CRIS)	Interviewing of employees, examining documents and video tapes of key events.	Internal revenue service (IRS, US government	• A query processing front-end in CRIS for automatic weighting. • Business rules for facilitating queries.	• Improvement in accessing and generating the reports on taxpayers that was time-consuming without DWH. • Increase in revenue.	Not Defined
[27]	Manufact-uring	Data warehouse	Interviewing of employees, examining documents and video tapes of key events.	Large Manufacturing Company	• Transfer of data from 100 mainframes and 6 external data sources to DWH. • From DWH data is transferred to dependent data marts.	• Helpful in making better decisions and creating better information • Quality information access. • Performance and failure for all parts can be measured. • Reason of failure detection becomes easier.	• Limited data samples

TABLE II. COMPARISON OF DIFFERENT CROSS DOMAIN AREAS AFFECTING DATA WAREHOUSE

Ref. No	Domain	Areas of Usage	% age Used	Cross domain
[37,38,40,39,35,41,42,36,28,12]	Medical	Hospitals, Clinics, Physician offices	23.3%	Government/Business
[34,45,28,27]	Finance	Tax departments	6.2%	Government/Business
[28,5,46]	Banking	Baking industry all around the world	6.2%	Business
[27]	Manufacturing	Home appliances	1.9%	Business
[33]	Education	Schools, colleges, universities	3.8%	Government/Business
[28, 47, 48, 49, 50,32, 52]	Marketing	Customer relationship management, trend analysis and information system	16%	Business
[29]	Social Media	Facebook, Twitter, others.	6.2%	Business

[43,44,31]	Construction	Infrastructure management	8.7%	Business
[6]	Agriculture	Agricultural production department	3.8%	Government/Business
[37,38,40,39,35,41,42,36,28,12]	Fraud Detection	Airports, Crime Agencies	1.9%	Government
[34,45,28,27]	Threat Analysis	Airports, Crime Agencies	1.9%	Government
[28,5,46]	Others	Others	20%	All

V. Analysis and Results

In this section we will see the areas, cross domains and usage of data warehouse around the world and the graphical view of inter related data effecting data warehouse.

A. *Comparison of different cross domain areas affecting data warehouse*

Table 2 shows the comparison of different cross domain areas and their interlinked data.

B. *Graphical representation of Survey*

Following graph shows percentage captured by different areas in DW around the world. As we can see from the Figure 10, medical holds top position in using DW technology.

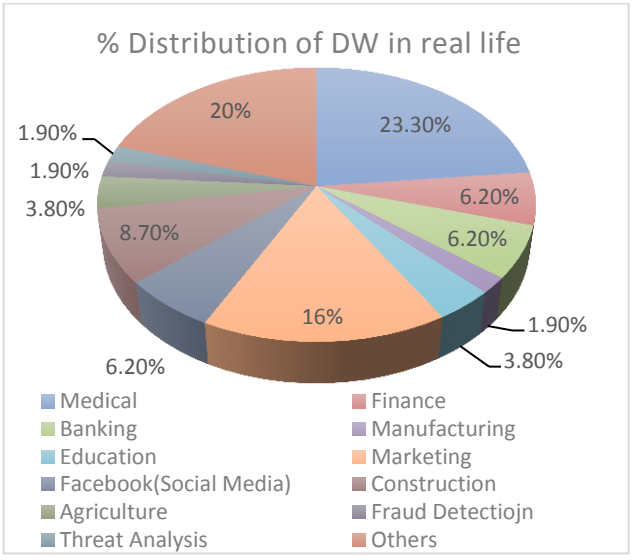

Fig. 10. Percentage Distribution of DW in Real Life

The Figure 11 shows the domain wise importance of DW, we can see clearly business domain holds top position. If we further drill down and look into specific business domain, we see from below the Figure 12 that banking and construction organizations are on top and competing each other with very less margin.

At the end if we take government domain we see that it holds a minor part in data warehouse. Fraud and threat detection are the only region effecting data warehouse through government as shown in the Figure 13.

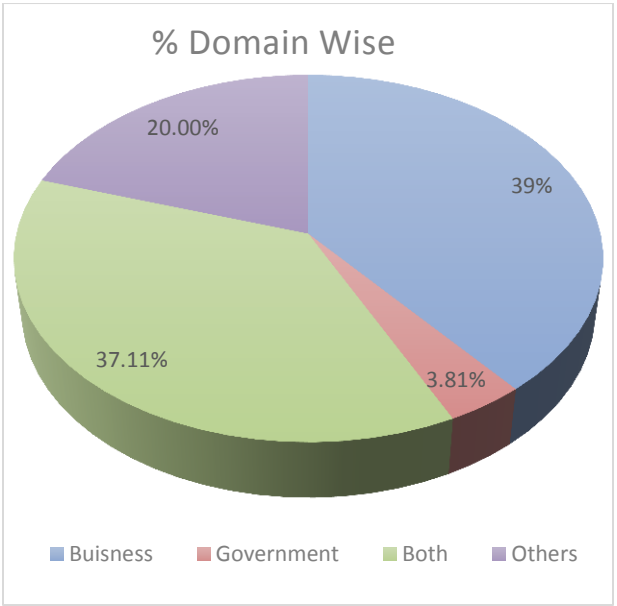

Fig. 11. Percentage Domain Wise

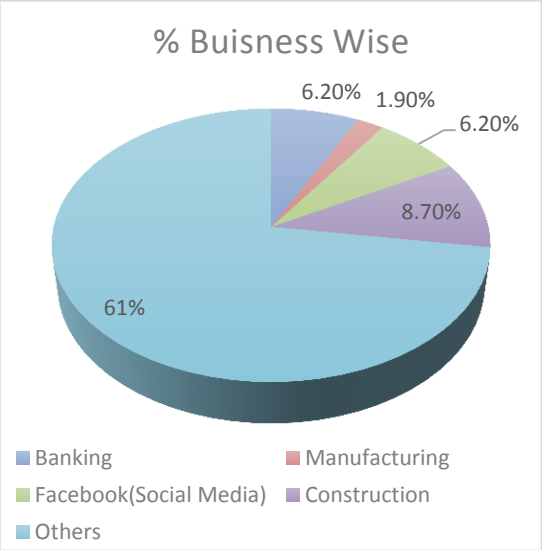

Fig. 12. Percentage business wise

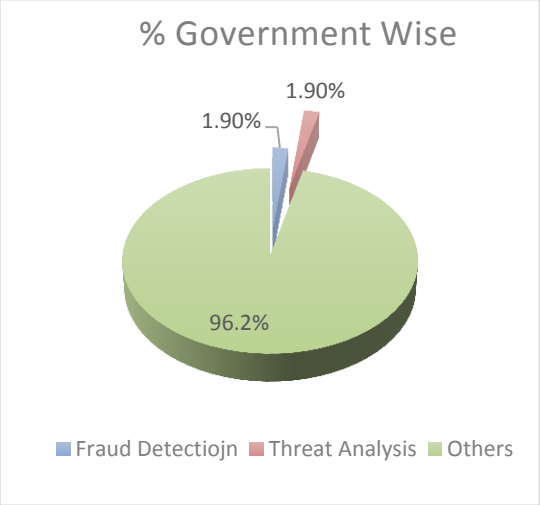

Fig. 13. Percentage government wise

VI. Discussion and Conclusion

This research survey describes the applications of data warehouse in various domains including government and non-government organizations. Our analysis is based on the literature review and case studies provided in this survey. The analysis of this study shows that the non-governmental organizations use data warehouse technology much more than the government organizations. The governments mostly use data warehouse for controlling the crime and fraud. Non-governmental organizations mostly use DW for data analysis, prediction and making decisions. Case studies are shown in the Table 1 that describe the importance of data warehouse in four domains; Healthcare, Banking, Finance and Manufacturing. The details of these case studies and their use of data warehouse have been discussed in the Section 4. The analysis of the Table 2 shows that data warehouse is being used in many application domains. The Figure 10 clearly depicts the areas that are using data warehouse. It shows that medical and marketing areas are using data warehouse much more than the other domains, whereas manufacturing, agriculture, education, and government sector are rarely using data warehouse. The areas such as social media, construction, and finance are moderately using data warehouse technologies. The Figure 12 shows business-wise comparison and the Figure 13 shows the government-wise comparison of data warehouse usage.

The analysis shows that data warehouse technology have been adopted in business as well as in government organizations for managing their huge data and for decision making. Still many organizations have not gone for the adoption of DW technology. Either they do not realize its importance or there may be difficulties in its adoption. The reasons for ignoring the importance of implementing DW technology have been discussed in literature that include quite large investment in terms of capital, more time utilization, looking for intangible benefits are difficult, the last but not the least problems holding with recent data management systems' infrastructure etc.

References

[1] T. Ariyachandra, H. J. Watson, "Key organizational factors in data warehouse architecture selection", Decision Support Systems 49 (2010) 200–212.

[2] T. R. Sahama, P. R. Croll, "A Data Warehouse Architecture for Clinical Data Warehousing", in Roddick, J. F. and Warren, J. R., Eds. Proceedings Australasian Workshop on Health Knowledge Management and Discovery (HKMD 2007) CRPIT, 68, pages pp. 227-232, Ballarat, Victoria.

[3] W.H. Inmon., "DW 2.0 Architecture for the Next Generation of Data Warehousing", DM Review, Apr 2006, Vol. 16 Issue 4, p.8-25.

[4] W.H. Inmon, "Building the Data Warehouse", Third Edition, York: John Wiley & Sons, 2002.

[5] Hwang, Hsin-Ginn, et al. "Critical factors influencing the adoption of data warehouse technology: a study of the banking industry in Taiwan." *Decision Support Systems* 37.1 (2004): 1-21.

[6] Nilakanta, Sree, Kevin Scheibe, and Anil Rai. "Dimensional issues in agricultural data warehouse designs." *Computers and electronics in agriculture* 60.2 (2008): 263-278.

[7] B.A. Devlin, P.T. Murphy, An architecture for a business and information system, IBM Systems Journal 27 (1) (1988) 60 – 80

[8] W.H. Inmon, Building the Data Warehouse, Wiley, New York, 1996.

[9] S.R. Gardner, Building the data warehouse, Communications of the ACM 41 (9) (1998) 52 – 60.

[10] J.V.D. Hoven, Data warehousing: bringing it all together, Information Systems Management (1998 Spring) 92 – 96.

[11] R. Kimball, The Data Warehouse Toolkit, Wiley, New York, 1996.

[12] R.M.T. Lu, K.A. Mazouz, A conceptual model of data warehousing for medical device manufacturers, Proc. of the 22nd Annual EMBS International Conference 2000 (July).

[13] D. Powell, To outsource or not to outsource? Netwo rking Management (1993) 56 – 59.

[14] Y. Yao, H. He, Data warehousing and the Internet's impact on ERP, IT Professional (2000 March) 37–41.

[15] Rob, P., Coronel, C., 2006. Database Systems: Design, Implementation, and Management. Course Technology.

[16] Sen, A., Sinha, A.P., 2005. A comparison of data warehousing methodologies. Commun. ACM 48 (3), 79–84

[17] Kimball, R., 2002. The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling. John Wiley & Sons, Inc.

[18] Alsquor, M., Matouk, K., Owoc, M. L., A survey of data warehouse architectures:preliminary results. Proceedings of the Federated Conference on Computer Science*and Information Systems*,Wroclaw,2012,Sivut 1121-1126.

[19] Hackney, D., 2002. Architectures and Approaches for Successful Data Warehouses, Oracle White Paper.

[20] CHAKIR, Aziza, Hicham MEDROMI, and Adil SAYOUTI. "Actions for data warehouse success." Editorial Preface 4.8 (2013).

[21] Chaudhuri, S., Dayal, U., 1997. An Overview of Data Warehousing and OLAP Technology. ACM SIGMOD Record 26 (1), 64–74.

[22] Thakur, Garima, and Anjana Gosain. "A Comprehensive Analysis of Materialized Views in a Data Warehouse Environment." IJACSA) International Journal of Advanced Computer Science and Applications 2.5 (2011).

[23] Watson, H.J., Haley, B.J., 1997. Data warehousing: a framework and survey of current practices. J. Data Warehousing 2 (1), 10–17.

[24] Watson, H.J., Gerard, J.G., Gonzalez, L.E., Haywood, M.E., Fenton, D., 1999. Data warehousing failures: case studies and findings. J. Data Warehousing 4 (1), 44–55

[25] www.informationweek.com/mobile/mobile-devices/gartner-21-billion-iot-devices-to-invade-by-2020/d/d-id/1323081

[26] www.gartner.com/newsroom/id/3165317

[27] Watson, Hugh J., Dale L. Goodhue, and Barbara H. Wixom. "The benefits of data warehousing: why some organizations realize exceptional payoffs." *Information & Management* 39.6 (2002): 491-502.

[28] Joseph, Madhuri V. "Significance of Data Warehousing and Data Mining in Business Applications." *International Journal of Soft Computing and Engineering (IJSCE) ISSN* (2013): 2231-2307.

[29] Thusoo, Ashish, et al. "Data warehousing and analytics infrastructure at facebook." *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data*. ACM, 2010.

[30] Chowdhury, Rajdeep, et al. "Implementation of Central Dogma Based Cryptographic Algorithm in Data Warehouse Architecture for Performance Enhancement." International Journal of Advanced Computer Science & Applications 1.6: 29-34.

[31] Park, Taeil, and Hyoungkwan Kim. "A data warehouse-based decision support system for sewer infrastructure management." *Automation in Construction* 30 (2013): 37-49.

[32] Ryals, Lynette, and Adrian Payne. "Customer relationship management in financial services: towards information-enabled relationship marketing." Journal of strategic marketing 9.1 (2001): 3-27.

[33] Goyal, Monika, and Rajan Vohra. "Applications of data mining in higher education." International journal of computer science 9.2 (2012): 113.

[34] Bhedi, Vaibhav R., Shrinivas P. Deshpande, and Ujwal A. Lanjewar. "Data Warehouse Architecture for Financial Institutes to Become Robust Integrated Core Financial System using BUID." International Journal of Advanced Research in Computer and Communication Engineering 3.3 (2014): 2278-102.

[35] Evans, R. Scott, James F. Lloyd, and Lee A. Pierce. "Clinical use of an enterprise data warehouse." AMIA Annual Symposium Proceedings. Vol. 2012. American Medical Informatics Association, 2012.

[36] Stolba, Nevena, and A. Min Tjoa. "The relevance of data warehousing and data mining in the field of evidence-based medicine to support healthcare decision making." International Journal of Computer Systems Science and Engineering 3.3 (2006): 143-148.

[37] Schubart, Jane R., and Jonathan S. Einbinder. "Evaluation of a data warehouse in an academic health sciences center." International journal of medical informatics 60.3 (2000): 319-333.

[38] Liu, Baoyan, et al. "Data processing and analysis in real‐world traditional Chinese medicine clinical data: challenges and approaches." Statistics in medicine 31.7 (2012): 653-660.

[39] Leitheiser, Robert L. "Data quality in health care data warehouse environments." System Sciences, 2001. Proceedings of the 34th Annual Hawaii International Conference on. IEEE, 2001.

[40] Yoo, Sooyoung, et al. "Electronically implemented clinical indicators based on a data warehouse in a tertiary hospital: its clinical benefit and effectiveness." International journal of medical informatics 83.7 (2014): 507-516.

[41] Prokosch, Hans-Ulrich, and T. Ganslandt. "Perspectives for medical informatics." Methods Inf Med 48.1 (2009): 38-44.

[42] Adlassnig, Klaus-Peter, et al. "Fuzziness in healthcare-associated infection monitoring and surveillance." Norbert Wiener in the 21st Century (21CW), 2014 IEEE Conference on. IEEE, 2014.

[43] Chong, Heap Yih, Rosli Mohamad Zin, and Siong Choy Chong. "Employing data warehousing for contract administration: e-dispute resolution prototype." Journal of Construction Engineering and Management 139.6 (2012): 611-619.

[44] Chau, Kwok-Wing, et al. "Application of data warehouse and decision support system in construction management." Automation in construction 12.2 (2003): 213-224.

[45] Chen, Wenzhe. "The Application of Data Warehouse Technology in Modern Finance." 2015 International Conference on Advances in Mechanical Engineering and Industrial Informatics. Atlantis Press, 2015.

[46] Lin, Zhonglin, et al. "Banking intelligence: application of data warehouse in bank operations." Service Operations and Logistics, and Informatics, 2008. IEEE/SOLI 2008. IEEE International Conference on. Vol. 1. IEEE, 2008.

[47] Shaw, Michael J., et al. "Knowledge management and data mining for marketing." Decision support systems 31.1 (2001): 127-137.

[48] Ngai, Eric WT, Li Xiu, and Dorothy CK Chau. "Application of data mining techniques in customer relationship management: A literature review and classification." Expert systems with applications 36.2 (2009): 2592-2602.

[49] Nedeva, Veselina Ivanova. "ANALYSIS OF MARKETING INFORMATION SYSTEMS AND CONCEPTION OF AN INTEGRATED MARKETING INFORMATION SYSTEM." International Journal of Computing 3.2 (2014): 127-133.

[50] Payton, Fay, and Debra Zahay. "Why doesn't marketing use the corporate data warehouse? The role of trust and quality in adoption of data-warehousing technology for CRM applications." Journal of Business & Industrial Marketing 20.4/5 (2005): 237-244.

[51] Thomas, Davenport, et al. "Data to Knowledge to Results, Building an Analytic Capability." California Management Review 43.2 (2001).

[52] Cunningham, Colleen, Il-Yeol Song, and Peter P. Chen. "Data warehouse design to support customer relationship management analyses." Proceedings of the 7th ACM international workshop on Data warehousing and OLAP. ACM, 2004.

[53] Watson, Hugh J., Celia Fuller, and Thilini Ariyachandra. "Data warehouse governance: best practices at Blue Cross and Blue Shield of North Carolina." Decision Support Systems 38.3 (2004): 435-450.

[54] Sarkar, Anirban. "Data Warehouse Requirements Analysis Framework: Business-Object Based Approach." International Journal 3 (2012).

[55] Diana, Nova Eka, and Aan Kardiana. "Comprehensive Centralized-Data Warehouse for Managing Malaria Cases." International Journal of Advanced Computer Science & Applications 1.6: 40-46.

[56] N.L. Sarda, Temporal issues in data warehouse systems, Database Applications in Non-Traditional Environments ’99, TheProceedings of the 1999 International Symposium on Database Application in Non-traditional Environments (DANTE ’99), IEEE Computer Society, Los Alamitos, 1999, pp. 27–34

[57] J.-B. Yang, N.-J. Yau, Application of case-based reasoning in construction engineering and management, Proceedings of the Third Congress held in conjunction with A/E/C Systems 1996, Computing in Civil Engineering, American Society of Civil Engineers, New York, 1996, pp. 663–669.

[58] K.W. Chau, Y. Cao, M. Anson, J.P. Zhang, Application of data warehouse and decision support system in construction management, Automation in Construction 12 (2) (2002) 213–224.

[59] J. Vanegas, P. Chinowsky, Computing in Civil Engineering, American Society of Civil Engineers, New York, 1996.

[60] J. Dyche, e-Data Turning Data into Information with Data Warehousing, Addison-Wesley, Reading, 2000