An Introduction to Database Management system

Donna R, Phani Kumar Pullela

School of Computer Science Engineering (SOCSE), RV University, RV Vidyanikethan Post, 8th Mile, Mysuru Road, Bengaluru – 560059

Abstract:

A database management system is a piece of software that enables users to define, develop, and maintain databases while also granting them controlled access to the data. It handles all requests, freeing end users and software applications without the need to know where the data is kept physically. The use of DBMS in analysis has promoted more advanced analysis, enhancing corporate solutions. This technique may be applied to many different situations, and its standardized methodology makes it appropriate for monitoring construction claims on active projects. Databases include a significant quantity of structured data, typically in the form of temporary files and metadata, which have been used in internal processes. The data structures are maintained to ensure transaction validity. A locking mechanism designed using different methods and offered by the relational database management system is utilized to secure the centralized database management system. One of the primary functions of web-based databases is collecting, organizing, and displaying all of the data. These databases diversely organize the data. The processing of spatial queries requires a data architecture that is easily matched, giving both spatial and non-spatial components of the data an equal chance to participate in query processing and optimization.

Keywords:

Database Management, Computer Applications, Database management system

Background

A software system known as a database management system (DBMS) offers controlled access to the data and allows users to build, create, and maintain databases. It optimizes the data organization by using a normalized database schema. Additionally, this aids in ensuring the security, consistency, and administration of data. This technology creates a connection between end users or application programs and databases. All requests are handled by the database management system, freeing end users and software applications from needing to know where the data is physically stored. It is a technological solution that aids in archiving

and is in favor of data-driven workflow. Additionally, it offers independent access to both logical and actual data. [1]

By incorporating analytical logic within the database, data analysis enables data processing within databases. The majority of this entails determining the data components required to support a data processing system. Analysis has been encouraged to reach higher levels thanks to the emergence of database management systems, which have improved business solutions. As a result, the design satisfies both current and future needs. Different types of databases can be utilized in DBMS to leverage information. These databases encompass many different types, such as relational, network, and hierarchical ones. On the server, a digital repository is set up to handle and store the data. It can offer a logical perspective on the data-manipulating processes. Additionally, this system has built-in backup and recovery mechanisms. Thus, database management systems are more effective in all fields dealing with big data. [2] **Figure 1** explains that the user, applications, or other database management systems provide the data first. A database system using SQL queries then processes this data. An interface between the user and the database files is provided, in this case, by the database management system. The queries are processed, and the stored data is accessed. Relational, hierarchical, flat file, and object databases are just a few data types stored in database files. Database management solutions have some limitations, such as the high cost of accessing vast amounts of data. Due to the enormous bulk of the data and the consequent reduction in frequency, it might be highly complicated.

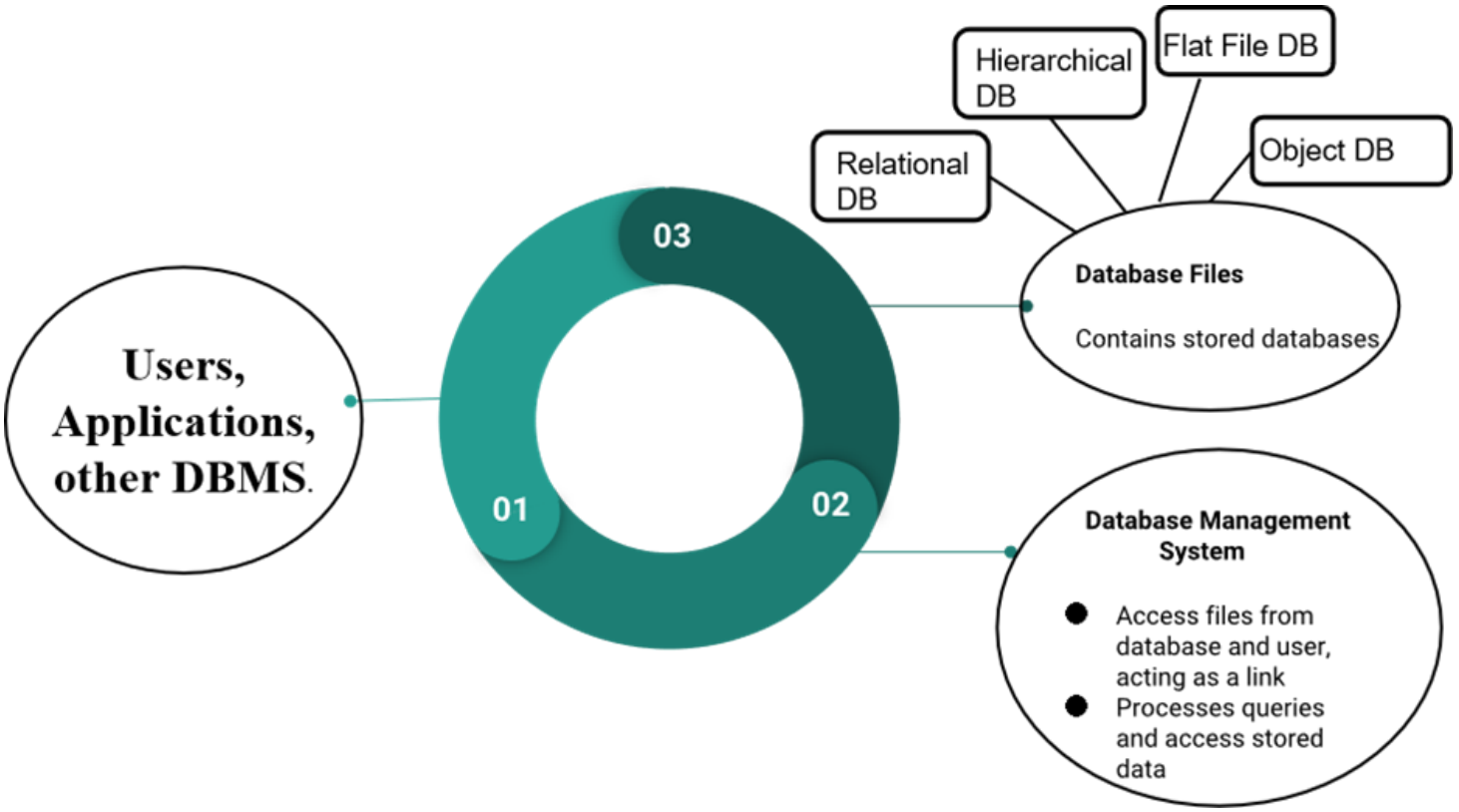

Fig1: Overview of DBMS

Introduction

The paper describes the use of discrete event simulation, content analysis, and business process improvement at the Infranian Information and Documentation Centre. A data flow diagram has been created to begin the process, followed by a Delphi study to determine the range of users. A service model has been developed for dissimilation. As a result of the simulation, several proposals for improving the current systems have been developed. Simulation and Delphi methods were effective in improving processes and could be applied in other organizations. [1] The conventional database systems weren't supporting classes of applications, so the focus shifted to geographic information systems. Numerous architectural techniques have been suggested to address the increasing demand for spatial data management. The data models are not rich enough to represent the spatial dimension of geographic entities because either they perform in the application layer on top of the database management system, or they cross over into the application layer. The study assignment focuses on techniques that facilitate the efficient execution of operations and composite procedures involving the spatial dimensions of geographic entities provided by the database management system. [2]

Due to its complexity and heterogeneity, workload management in database management systems is challenging. In workload management, it is important to predict performance well. As well as prior knowledge of the type of workload, health product, and its performance, tuning the database management system is the same. The study proposed a predictive and adaptive framework named Automatic Workload Performance Prediction (AWPP). Workload management problems are being solved using case-based reasoning. Other machine learning techniques have been compared with this approach. The MySQL database management system is used to execute benchmark workloads for decision support systems (DSS) and online transaction processing to validate the AWPP framework. Thus, it is proposed that the AWPP framework through CBR modeling produces better predictions and adaptive database management system workloads. It is possible to better manage workloads with the algorithms of the database management systems.[3]

The design and development of tool management systems are discussed in this work using a systematic method. This unit of Bharat Heavy Electricals Limited uses it primarily in a value-adding capacity (BHEL). The key aspects of the design and development of the centralized database for tools are explored in the study. These tools include things like parts, machines, and tool flow, as well as tool allocation, storage, and retrieval. The solution that appears to be currently in development is expected to significantly reduce the machine downtime caused by a lack of the right tools.[4] A Microsoft Access database management system has been developed to monitor and analyze the receipt of additional funds and the time needed to finish construction projects. There is now a database to make it easier to identify the root causes, contributing factors, and fundamental management problems of Kuwaiti construction claims. The technology could be utilized for several purposes, and its

uniform methodology would be suitable for keeping track of construction claims on active projects. Testing has been done on the database. This database has undergone testing and is founded on feedback from eight completed construction projects. [5] **Figure 2** explains that the database management system is both a group of connected data and a collection of tools for quickly and effectively accessing that data. This system was created to handle a significant amount of data that might be stored or retrieved. Data management and improved data efficiency are its key goals. The absence of data redundancy, data security, and ease of access are its key benefits. These database management systems are employed in many different fields to improve their effectiveness in the dynamic environment. Telecom, banking, sales, and many more areas are among these categories.

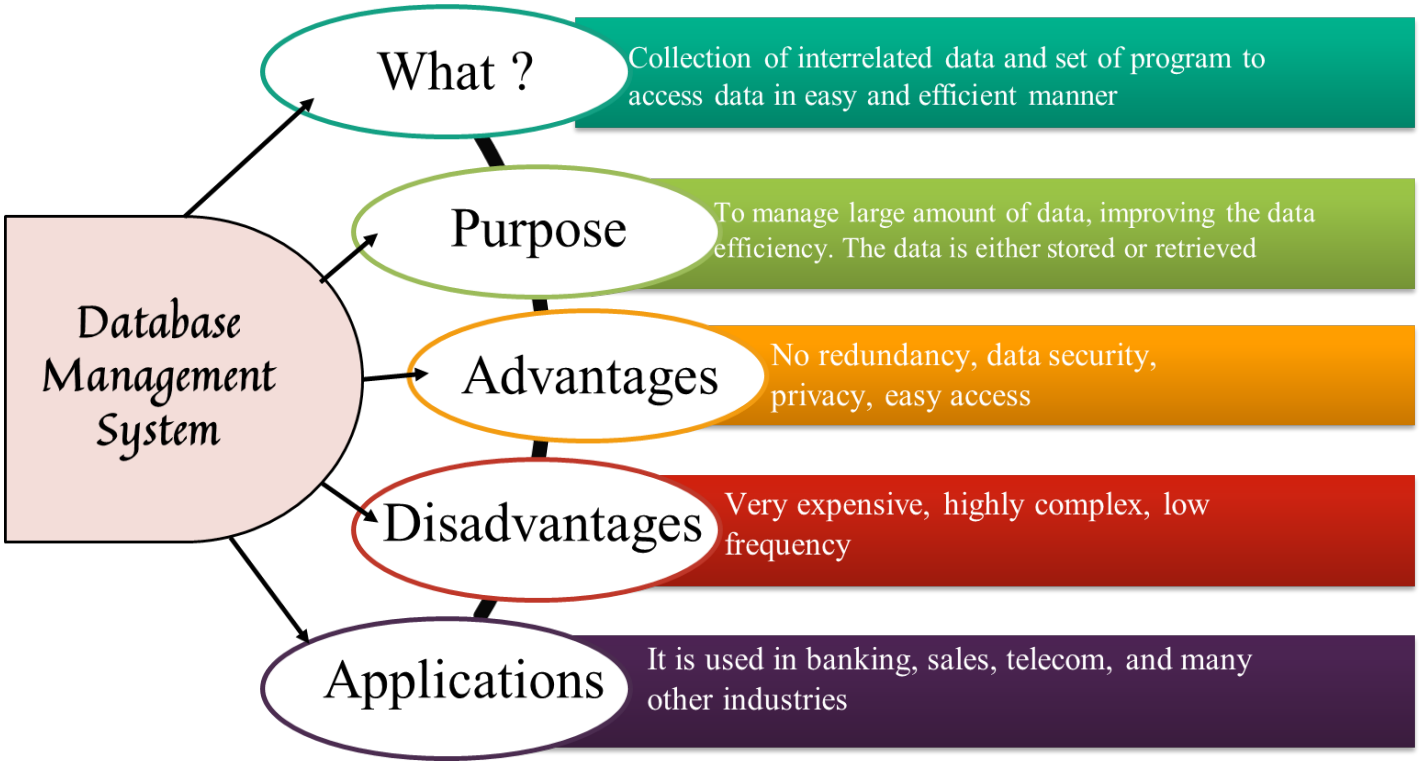

Figure 2: Introduction to DBMS.

Prior Art:

This study outlines several pieces of literature and numerous methodologies that support the method of analysis that patent analysts most frequently employ. Text segmentation, summary extraction, feature selection, term association, cluster construction, topic identification, and information mapping were some of the techniques used in this study. The design of these strategies takes these effectiveness concerns into account. A method to assess the value of segment extracts, a corpus, and a key phrase extraction technique are some of the crucial characteristics it is recommended to have. A successful co-word analysis technique has been

utilized on numerous patents. A general cluster title generator is also developed automatically for easier result interpretation. The outcomes of the procedures that were tested supported the machine generation that was carried out to keep more crucial content terms. The attempt to automate patent sets has been successful in that it has aided in the creation of patent maps, which has enhanced patent analysis. [6] Effective handling of RDF data is required to realize the vision of the semantic web (SW). There are certain problems, primarily with performance and scalability, when implementing SW technology in the real world. This study mainly looked at the causes of the shortcomings of current data management solutions. The RDF data has been vertically partitioned as an alternative remedy for this. Based on queries produced by a web-based RDF browser, this recently created approach is contrasted with the prior art. The outcomes demonstrated that, while having a more straightforward architecture, the vertically partitioned schema performed similarly to that of the property table technique. When employing a column-oriented database management system, performance was seen to increase significantly. A brand-new database management system called SW Store has been actively developed, employing strategies to produce high-performance RDF data. [7]

The amount of multimedia data that is currently being recorded, created, and saved is rising quickly. As a result, the organization and accessibility of data from repositories have stimulated both the commercial and academic sectors. Although the database management system is quite effective and has garnered a lot of attention, multimedia data handling has not been successful. This study introduces Bilvideo, a working prototype of a video database management system. This BilVideo server uses object-relational databases to react to semantic questions and handles spatiotemporal queries. [8] Indigenous people are needed for sustainable use of the natural environment because of their knowledge-based approach to biocultural management. As a result, the use of online geographic databases as a tool to teach land managers about Indigenous biocultural knowledge has increased. This was carried out to raise awareness of the problem, spot knowledge gaps, and then encourage collaboration. The report describes the methods used to construct the online resource and how the Internet and spatial analytic approaches were used to provide an overview of Australian IBK that was recorded and publicly available. The Australian Center for Ecological Analysis and Synthesis, international cooperation, aids AIBK in creating its one-of-a-kind online resources. Additionally, it is projected that as more people become aware of online resources, the Australian Center for Ecological Analysis and Synthesis, through international cooperation, will assist AIBK in creating its one-of-a-kind online resources. Further, it is projected that as people become more aware of online resources, more records will be made available, creating a database that will be essential to future management and the upkeep of natural and cultural resources. [9]

The systematic examination of database management systems has just started, but earlier forensic analysis at the file system level was quite popular. Most of the structured data in databases is temporary files and metadata that have been used in internal processes. Ensure the transaction's legitimacy by ensuring the data structures are kept up to date. These are typically not readable by humans and are also meant to be used by internal system operations. It is said that this study used replication and transactional sources to create a forensic-aware database management system. The investigation's evidence needs to be

rebuilt, so the internal data structures serve as an essential baseline. The database management system's overall benefit would be hampered by not having more logs. Their strategy was constant as well, sharing the veracity of the facts and bolstering custody. [10] **Figure 3** explains that the data is now categorized and transferred to relational database management systems from the data warehouse. The base data leader receives the data from the relational database management system. These facts are now regarded as aggregated facts. Big data and aggregated data are transferred to the MOLAP server. Data is transmitted to the retrieval module of the MOLAP server, where querying, mashing, and indexing are done. The application of the logic module follows. Given rank ratios and forecasts, the database slides in this case. It has been very thoroughly categorised based on the requirements of the user. It produces a presentation module using the categorized data for clients and the web. These all take place on the MOLAP server, which provides services for the client services.

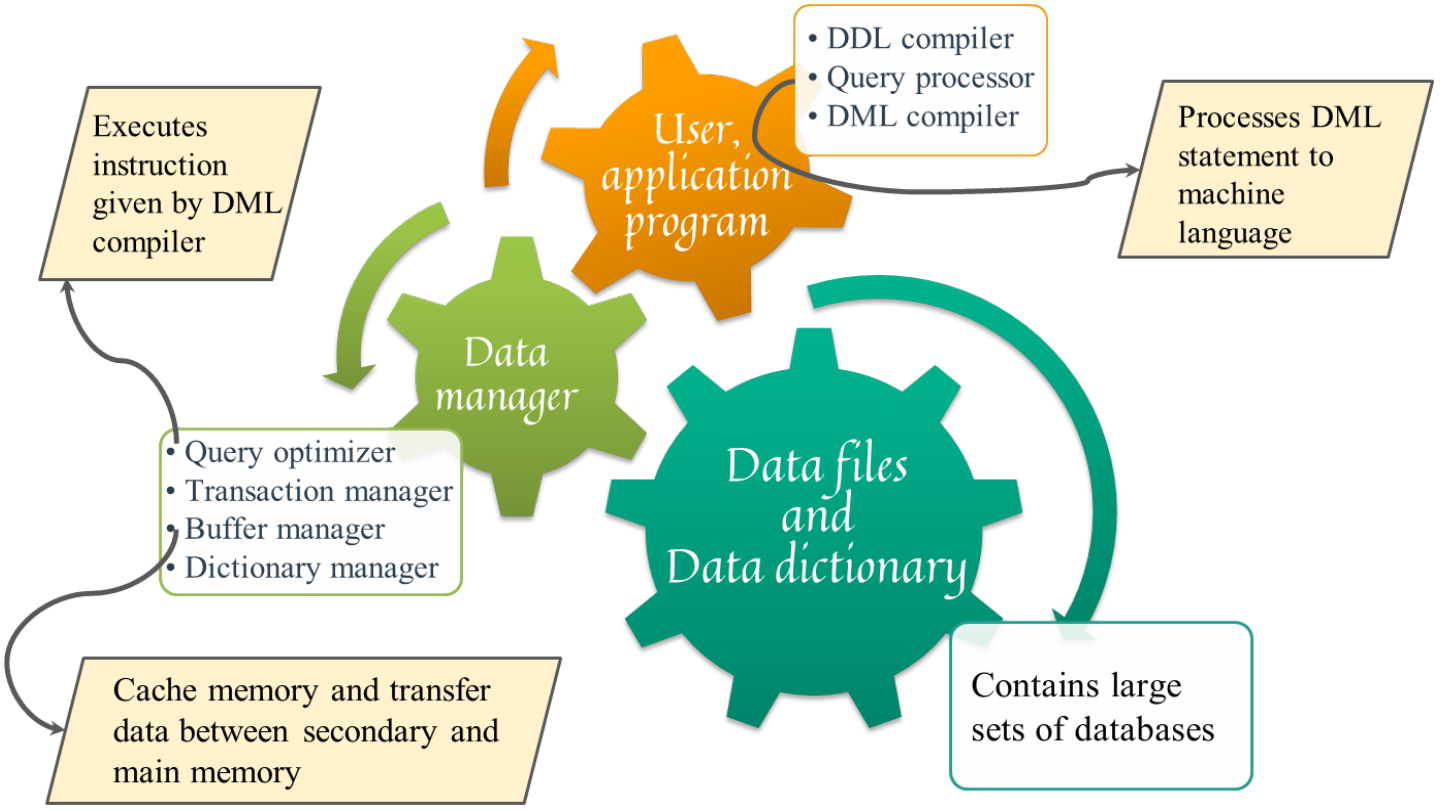

Figure 3: Prior Art of Database Management System

Theory:

The study suggests that GEDBLOG systems enable the development of programs that handle graphic elements using a declarative defining approach. Additionally, this method enabled the constant development and creation of graphic applications. The main method used for this was incremental development, which ensures that only applications that fit its specialization are used. Graphics serve as a distinguishing feature of GEDBLOG applications on their own. One could accomplish this by adding a graphical data language to the logical database

management system EDBLOG, enabling it to handle both graphical and non-graphical information declaratively. [11] Research on the digitized capture of spoken language data is being carried out now in many different domains. Before the development of this program, data access was provided on an as-needed basis using tools and labelling standards. To get around this, the study concentrated on the Emu system's design and provided a thorough assessment of the system's supporting framework. In this situation, the densely nested annotations were accommodated using a comprehensive speech database management system called Emu. This system supported overlaying extra annotation with inner token relations on existing labelled files, as well as reading a variety of common label and data file formats. Additionally, the Emu system offers a graphical label tool for use in specialty displays. [12]

HOLMES is a database system that is based on object relationships. Here the generalized schema idea is put forth in the system. The system has a uniform language that could express both formulas of a generalized schema and user queries. It also has a high-level database that allows defining the database application as a "theory," thus making it more efficient in various fields. [13] The impact of two-phase locking is being assessed on the database management system in this study, mainly on how concurrency control strategies relate to the typical transaction response time. Regarding the static degree of interference, the foundation of this technique is two-phase locking. Additionally, the transaction that was abandoned is measured. The effect of variables like data replication, time out, etc. on the concurrency control mechanism performance of distributed database management systems is another metric. The SIMULA programming language and simulation have both been employed in this work. [14]

Concurrency control is a crucial component of the database management system that safeguards database consistency. An approach for securing centralized database management systems is the use of locking mechanisms. With the help of the relational database management system built into the locking mechanism, the developer is provided with a wide range of possible alternatives. Here, a simulation model that produced findings for the tests is described. These outcomes show how well it performs when dealing with various locking mechanisms and transaction behaviour patterns. [15] **Figure 4** depicts compiler activity. The compiler converts the DML statement into machine language using a DML compiler. The query processor then interprets the client's request. This information has now been entered into the database management system. In this case, authorization is regulated, and the data dictionary is managed. The data subsequently goes to the query optimizer after passing through the command processor. The DML compiler's interaction is carried out here. The data manager, which is also the buffer manager, processes the additional database. Prior to moving on to data files and dictionaries, the buffer manager clears cache memory and moves data between the secondary and main memory.

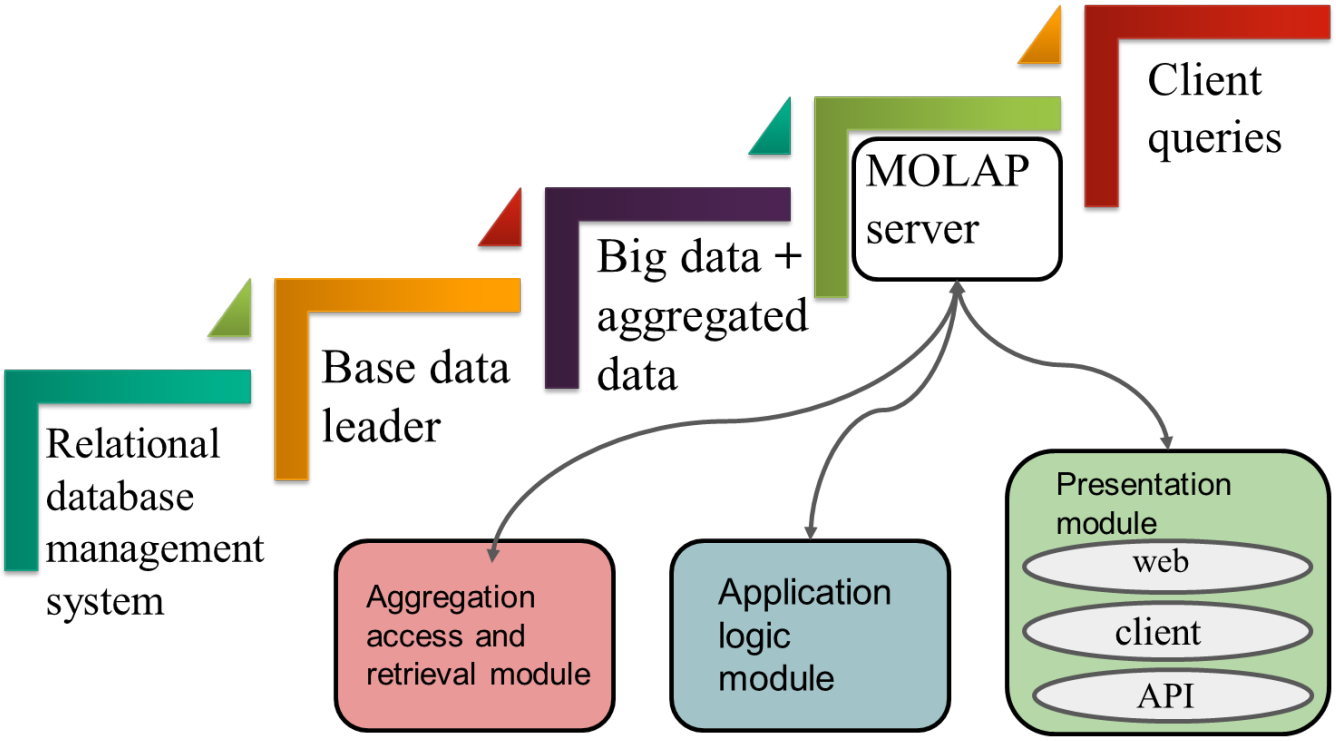

Figure 4: Detailed overview of Database Management System

Current status:

Numerous studies on allergenic chemicals and proteins have been conducted throughout the last few years. One of the primary functions of web-based databases is to gather, arrange, and display all of the data. These databases use various methods to organize the data. This technique has protected the databases from links to one another and modification by other biomedical areas. This study provided information on the techniques and database conditions for allergens; analysis was also carried out using a bioinformatics program. The platform allergome, together with its modules and tools, has been used as an example of interconnectivity and data integration. Hence, it serves as a platform for the fusion of all experimental, clinical, and epidemiological data. [16] As the world's population grows, there are numerous academic and professional initiatives being made to build intelligent wearable systems (SWS). The principal causes of the quick rise in demand could be attributed to rising healthcare expenditures and technological advancements, particularly micro- and nanotechnologies. There have also been advancements in SWS, smart fabrics, and many other areas. The database is gathered to build a model that will serve the SWS's present healthcare demands. This makes it possible to control a patient's health individually and continuously. The primary goal is to provide low-cost wearable timepieces with a system that supports sophisticated healthcare applications, tracking health, activity, mental state, and much more. [17]

A technique for making decisions is conditional-based maintenance. The primary foundation of this methodology is the real-time identification of potential malfunctions and the forecast for future medical equipment. This proactive method necessitates the development of a more statistical model that might set off an alarm for maintenance needs. Many different technologies and algorithms that are thought of as steps for prognostic maintenance are the subjects of extensive study and development. Taking these actions is crucial for managing operational reliability and making decisions. Prognostic models come in four varieties: the physical model, the knowledge model, the data-driven model, and the combination model. The study mentioned above summarises a growing trend in the field of machine prognostication. [18] Over the past couple of years, database management systems have displaced traditional file processing systems. This is mostly due to database management systems' ability to circumvent many of the drawbacks of traditional file systems. This study emphasizes the use of federated databases. As a result, users and programs can access and modify data from several heterogeneous databases while remaining independent. Here, the most important problems, difficulties with establishing systems, and needs have been disregarded. From the perspectives of transaction management, system architecture, and schema integration, there are three ways to approach design difficulties. Additionally, this strategy adheres to a five-step integration technique. [19]

Pattern mining is a more narrowly focused area of study in data mining. The study required a significant quantity of data collection, ranging from several frontiers to efficient and scalable algorithms in transaction databases. Sequential pattern mining, correlation mining, frequent pattern-based grouping, etc. are examples of research horizons. Frequent pattern mining research is currently in a very advanced state. Therefore, expanding the scope of data analysis has a significant long-term impact on data mining methodology and application. [20]

Figure 5 explains the database then moves on to cloud services in this location, where organizational priority is consistent, achieving the objectives. Then comes big data, which has the capacity to process all data types. Both unstructured and structured data are possible. Business intelligence comes next. It has methods and instruments for business strategic planning. The database in bridge (SQL or no SQL) comes next. This method brings out the greatest aspects of both. The following trend emphasizes security that eliminates potential flaws. The automatic methods utilized in this approach include project overflow, upgrades, patching, and streamlined maintenance.

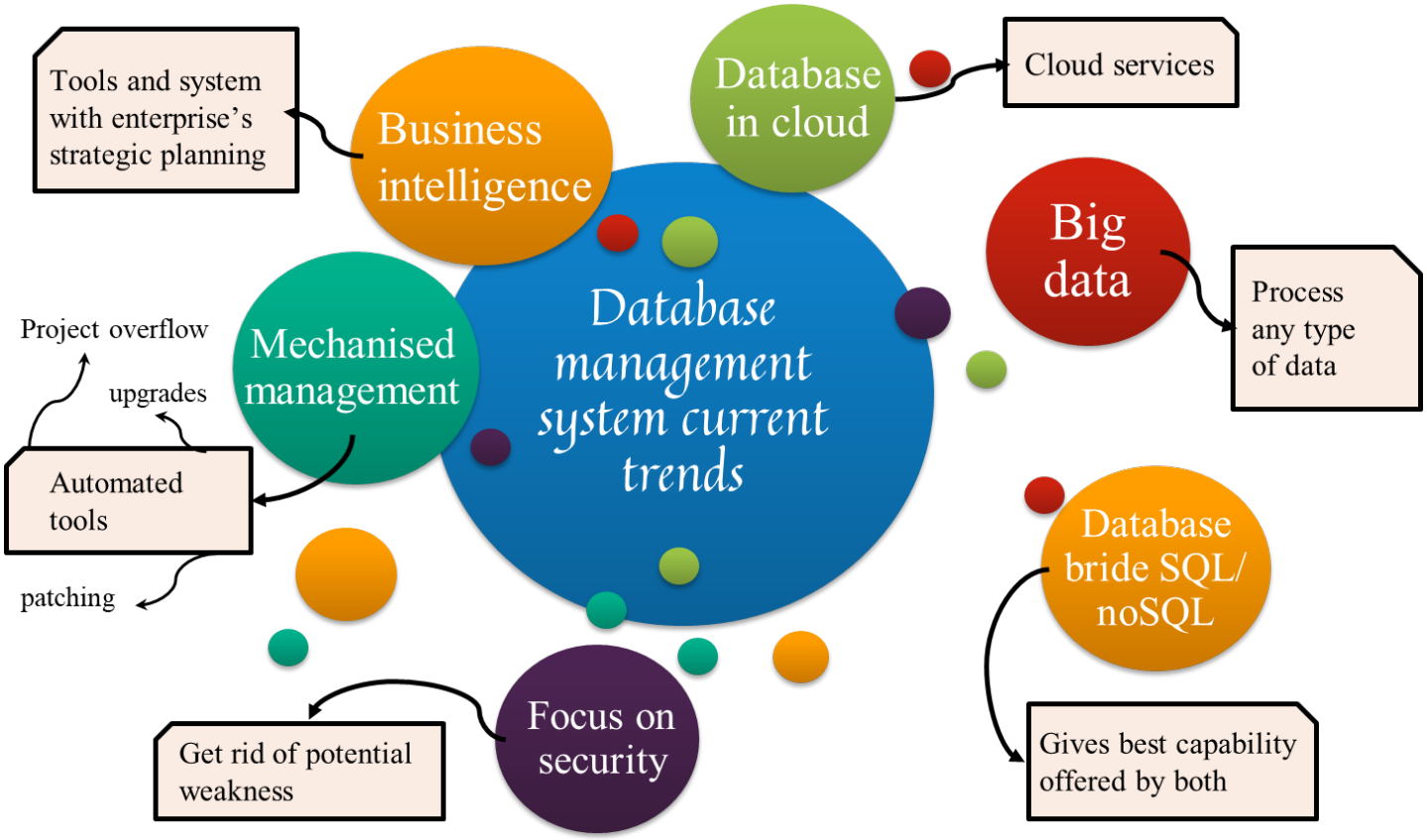

Figure 5: Current trends with Database Management System

Future Scope:

The 4Ps, which are fresh on the market in the 21st century, are the focus of new marketing and marketing communication. A structured model was developed for three markets based on how information technology is handled. When information technology went from one market to another, it does appear that this technique guaranteed changes in communication development. The idea of integrated marketing communication resulted from this evolution. This idea serves as the foundation for the level transaction process, which calls for the company to move from one stage of integrated marketing communication to another. Contrarily, this was predicated on the organization's capacity to seize and control information technology. [21] Today's dominant style is the adoption of service-oriented decision support systems (DSS). The conceptual foundation of DSS and how it has been handled in the product oriented DSS environment were both discussed at length, along with an outline of the challenges faced by DSS in the cloud. Data, information, and analytics are well defined in this system. Organizations considered the need to value service level and quality in addition to price and time. The DSS in the cloud enabled scale, scope, and greater speed. Applying information technology strategy to database and design viewpoints does result in new knowledge. [22]

In contemporary database design, the issue of spatial operation with database processes is more significant. To resolve this, a data architecture that can readily match the speedy

processing of spatial queries is required. Consequently, it offers an equal opportunity for both geographic and non-spatial components of data to participate in query processing and optimization. The enhanced operators give the database environment a suitable interface for integrating multimedia data. These operations redefine the terms insertion, deletion, relation-based selection, and spatial-based selection. The previously stated data architecture is built on an extensible database management system. As a result, it is used to gauge and confirm the effectiveness of supporting spatial data. [23] Big data analytics has gained importance in e-commerce recently, but it not been thoroughly examined. This study provides an interpretive framework for examining the definition of big data analytics' characteristics, economic value, and challenges in the context of e-commerce. Presenting a general review of the difficulties encountered in both the theory and practice of the same. The result is a variety of big data analytical principles, including the definition of big data, its types, its nature as commercial value, and much more. Providing both deeper insights and cross-cutting analytics as a result. [24]

A literature evaluation and classification of the articles were used in this study to build knowledge management. Since the study was based on a poll, there were primarily seven categories. The KM framework, knowledge-based systems, data mining, information and communication technology, artificial intelligence, database technology, and modelling with an emphasis on various problem domains are some of these. The outcomes pointed to the continued advancement of KM technology. According to the findings, KM technology may eventually move toward expert orientation and problem-oriented domains. The KM technologies have required the application of statistical methodologies; therefore, the addition of qualitative and quantitative approaches could enhance and widen the KM technologies' frontiers. [25]

Figure 6 provides a peek into database management systems' potential. In the long run, database management systems are quite effective in a variety of ways. This is achieved through managing big data and evaluating existing data for potential usage in the future. By combining several technologies, the new generation is made accessible. Polyglot persistence is a cloud-based technology that works with complex databases remarkably quickly. For real-time processing, more database connections are used. Data destruction maintains rapid analytics.

.

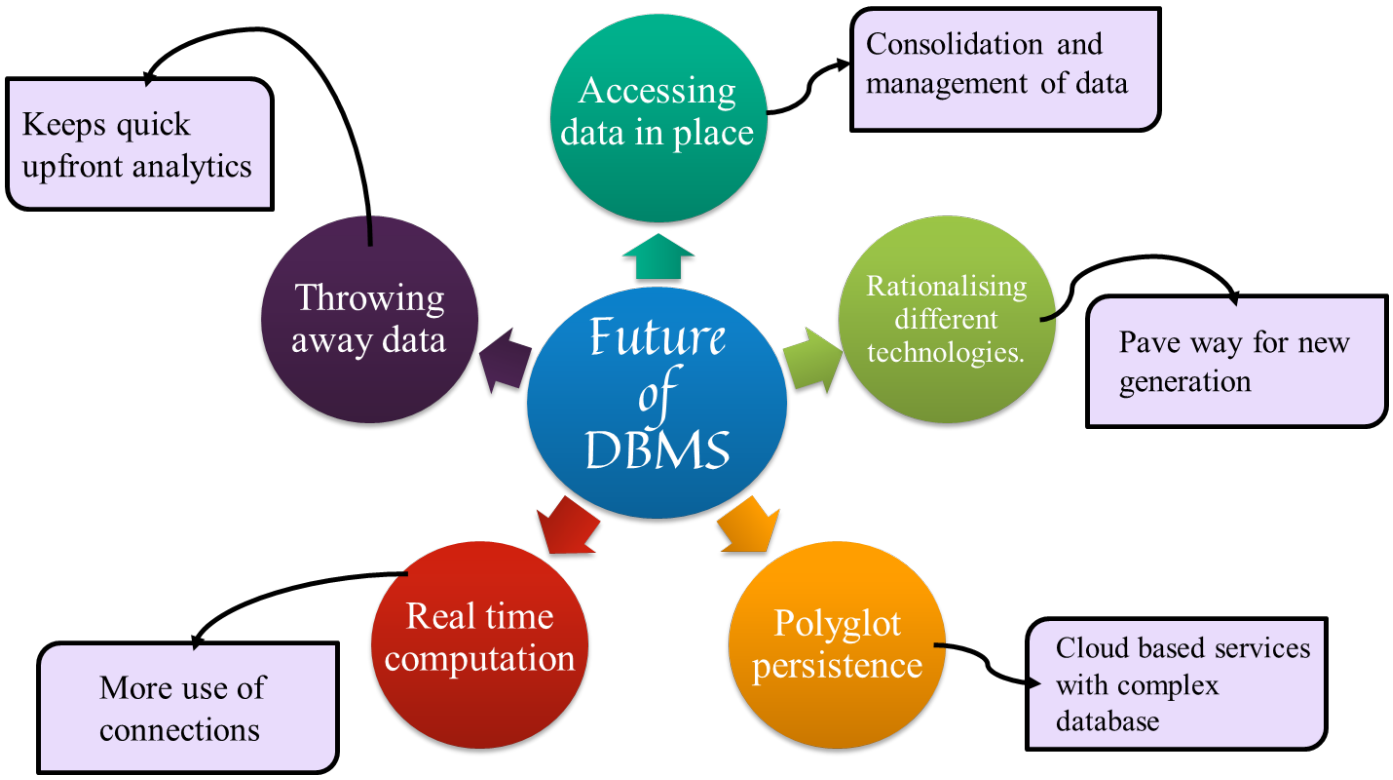

Figure 6: Future of Database Management System

Conclusion:

It aids in supplying data security, fusion, and uniform data administration practices. It offers separate access to both logical and physical data. The use of database management systems (DBMS) in analysis has promoted greater levels of analysis that have improved business solutions and involved the identification of data elements required to support data processing systems. According to estimates, the system now under development will greatly reduce the amount of time that machines are unavailable due to a lack of the necessary tools. A new database management system called Semantic Web Store has been actively developed to achieve high performance of RDF data. Here, a broad speech DBMS called Emu enabled complicated multi-level annotations. Due to the present state of frequent pattern mining research, the scope of data analysis has been significantly elevated. This has had a significant long-term influence on data mining applications and methodologies. The KM technology's scope might likely expand to include expert orientation and problem-oriented domains, which include the use of statistical methodologies in addition to quantitative and qualitative approaches. As a result, the DBMS excels in many areas and has a broad range of applications in these rapidly developing technologies.

Acknowledgement:

This research is funded by student research grant from RV University through Technical Writing course.

References:

[1] C. a. L. B. a. O. R. Mohan, "Transaction Management in the R* Distributed Database Management System," *ACM Transcations Database Systems,* vol. 11, no. 4, p. 378–396, 1986.

[2] J. P. a. S. E. H. Fry, "Evolution of Data-Base Management Systems," *ACM Computing Surveys,* vol. 8, no. 1, p. 7–42, 1976.

[3] O. Ö. B. D. I. K. N. S. A. K. Fatma Pakdil, "Simulation for business process re-engineering: case study of a database management system," *Journal of the Operational Research Society,* vol. 52, no. 12, pp. 1327-1337, 2017.

[4] E. S. &. T. Sellis, "Enhancing Operations with Spatial Access Methods in a Database Management System for GIS," *Cartography and Geographic Information Systems,* vol. 25, no. 1, pp. 384-388, 2013.

[5] Y. J. K. A. K. M. A. A. M. F. Basit Raza, "Performance prediction and adaptation for database management system workload using Case-Based Reasoning approach," *Information System,* vol. 76, pp. 46-58, 2018.

[6] ,. A. G. ,. S. A. &. P. R. G. Subrahmanyam, "Development of a Tool Database Management System," *The International Journal of Advanced Manufacturing Technology,* vol. 15, no. 8, p. 562–565, 2013.

[7] S. M. F. D. J. H. Sheikh Sabah Jaber Ali Al-Sabah, "A database management system to document and analyse construction claims," *Advances in engineering software,* vol. 34, no. 8, pp. 477-491, 2003.

[8] C.-J. L. Y.-I. L. Yuen-Hsien Tseng, "Text mining techniques for patent analysis," *Information Processing & Management,* vol. 43, no. 5, pp. 1216-1247, 2007.

[9] D. J. A. M. S. R. H. K. Abadi, "SW-Store: a vertically partitioned DBMS for Semantic Web data management," *The VLDB journal,* vol. 18, no. 2, pp. 385-406, 2009.

[10] M. E. Ş. E. A. U. U. Ö. G. U. Dönderler, "BilVideo: Design and Implementation of a Video Database Management System," *Multimedia tools and application,* vol. 27, no. 1, pp. 79-104, 2005.

[11] E. J. E. J. L. P. A. C. J. M. P. G. T. Petina L. Pert, "An online spatial database of Australian Indigenous Biocultural Knowledge for contemporary natural and cultural resource management," *Science of the total environment,* vol. 534, pp. 110-121, 2015.

[12] P. K. K. K. E. W. Peter Frühwirt, "Towards a forensic-aware database solution: Using a secured database replication protocol and transaction management for digital investigations," *Digital investigation,* vol. 11, no. 4, pp. 336-348, 2014.

[13] D. D. G. P. I. F. N. P. Asirelli, "Graphics by a logic database management system," *Journal of Visual Languages & Computing,* vol. 5, no. 4, pp. 365-388, 1994.

[14] J. H. Steve Cassidy, "Multi-level annotation in the Emu speech database management system," *Speech communication,* vol. 33, no. 1-2, pp. 61-77, 2001.

[15] H. R. Janusz Getta, "A deduction augmented database management system," *Information Systems,* vol. 9, no. 2, pp. 167-179, 1984.

[16] E. B. C. C. C Thanos, "The effects of two-phase locking on the performance of a distributed database management system," *Performance Evaluation,* vol. 8, no. 2, pp. 129-157, 1988.

[17] C. C. Cory Devor, "Structural locking mechanisms and their effect on database management system performance," *Information systems,* vol. 7, no. 4, pp. 345-358, 1982.

[18] A. R. C. P. P. S. E. Mari, "Allergen databases: Current status and perspectives," *Current allergy and asthma reports,* vol. 9, no. 5, pp. 376-383, 2009.

[19] D. E. J.-Y. F. C. E. E. C. Marie Chan, "Smart wearable systems: Current status and future challenges," *Artificial intelligence in medicine,* vol. 56, no. 3, pp. 137-156, 2012.

[20] Y. D. M. Z. M. J. Peng, "Current status of machine prognostics in condition-based maintenance: a review," *The International Journal of Advanced Manufacturing Technology,* vol. 50, no. 1, pp. 297-313, 2010.

[21] N. N. K. Magdi N Kamel, "Federated database management system: Requirements, issues and solution," *Computer Communication,* vol. 15, no. 4, pp. 270-278, 1992.

[22] J. C. H. X. D. Y. X. Han, "Frequent pattern mining: current status and future directions," *Data Mining and Knowledge Discovery,* vol. 15, no. 1, pp. 55-86, 2007.

[23] D. E. ,. H. F. Schultz, "Transitioning marketing communication into the twenty-first century," *Journal of Marketing Communication,* vol. 4, no. 1, pp. 009-026, 1998.

[24] D. D. Haluk Demirkan, "Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud," *Decision Support Systems,* vol. 55, no. 1, pp. 412-421, 2013.

[25] W. G. S. H. Aref, "Extending a DBMS with spatial operations," *Advances in Spatial Databases,* vol. 525, pp. 297-318, 1991.

[26] S. W. S. F. Akter, "Big data analytics in E-commerce: a systematic review and agenda for future research," *Electronic Markets,* vol. 26, no. 2, pp. 173-194, 2016.

[27] S.-h. Liao, "Knowledge management technologies and applications—literature review from 1995 to 2002," *Expert System with Applications,* vol. 25, no. 2, pp. 155-164, 2003.