Enhancing Data Warehouses Security

Muhanad A. Alkhubouli, Hany M. Lala, AbdAllah A. AlHabshy, Kamal A. ElDahshan
Department of Mathematics, Faculty of Science, Al-Azhar University, Cairo, Egypt

***Abstract*—Data Warehouses (DWs) are essential for enterprises, containing valuable business information and thus becoming prime targets for internal and external attacks. Data warehouses are crucial assets for organizations, serving critical purposes in business and decision-making. They consolidate data from diverse sources, making it easier for organizations to analyze and derive insights from their data. However, as data is moved from one source to another, security issues arise. Unfortunately, current data security solutions often fail in DW environments due to resource-intensive processes, increased query response times, and frequent false positive alarms. The structure of the data warehouse is designed to facilitate efficient analysis. Developing and deploying a data warehouse is a difficult process and its security is an even greater concern. This study provides a comprehensive review of existing data security methods, emphasizing their implementation challenges in DW environments. Our analysis highlights the limitations of these solutions, particularly in meeting scalability and performance needs. We conclude that current methods are impractical for DW systems and support for a comprehensive solution tailored to their specific requirements. Our findings underscore the ongoing significance of data warehouse security in industrial projects, necessitating further research to address remaining challenges and unanswered questions.**

Keywords*—*Data warehouse; data security; encryption; security issues; data integrity; privacy; confidentiality

I. Introduction

A data warehouse contains sensitive and confidential information. Since users' access data in the data warehouse at many levels within the organization, protecting this information is crucial. For each of their organizational processes, all organizations collect data and input it into computer systems [1].

The concept of a data warehouse is rooted in storing data in a structured manner for an extended time. This allows the data to be archived and easily accessible for future use. The structure of the data warehouse is designed to facilitate efficient analysis. Data warehouses are among an organization's most important assets and are primarily employed in crucial business and decision-making processes. The data warehouse incorporates data from several sources. As a result, security risks develop when moving data from one location to another [2].

Data warehouse security discusses the methods that may be used to safeguard the data warehouse by preventing access to information by unauthorized users to maintain the data warehouse's reliability [3]. The owner must encrypt critical data before outsourcing to guarantee its secrecy. Developing and deploying a data warehouse is a difficult process and its security is a major concern [4]. The organization does not always benefit from the decoration that emphasizes security considerations. Because of this, it is crucial to assess the security aspect of a data warehouse [5].

The handling of the significant amount of data gathered from the numerous daily transactions is the most crucial responsibility. Data warehouses have seen a growth in data collecting volumes because of the organization's processes becoming more computerized. As a result, more people are now accessing and utilizing the data [6].

Data security encompasses concerns regarding the confidentiality, integrity, and availability of data. These concerns include ensuring privacy, maintaining accuracy, validity, and consistency of data, and ensuring that data is immediately accessible. Confidentiality is the act of safeguarding information from being disclosed without authorization, whether it is by direct access or indirect logical deduction [7].

The rest of this paper is organized as follows. Section II introduces the technology of data warehousing. Section III discusses the security approaches employed in data warehousing. Section IV introduces the research challenges and possibilities. Ultimately, Section V presents our conclusions.

II. Data Warehouse Technology

A. Fundamental Architecture of Data Warehouses

The foundational architecture of data warehouses typically follows a structured and layered approach, which includes the following components [8]:

1) Data sources: Data sources are the systems or applications from which data is extracted and loaded into the data warehouse. These sources can include transactional databases, operational systems, external data feeds, spreadsheets, or other data repositories. Data extraction techniques are employed to gather the required data and prepare it for loading into the data warehouse.

2) Data integration: Data integration involves the process of mixing data from various sources and transforming it into a unified and coherent format appropriate for analysis. This step includes tasks such as data cleaning, data normalization, data aggregation, and data enrichment. The transformed data is then loaded into the data warehouse.

3) Staging area: The staging area acts as an intermediary storage space between the data sources and the data warehouse. It holds the extracted and transformed data temporarily before it is loaded into the main data warehouse. The staging area allows for data validation, error handling,

and data quality checks before the data is moved into the production environment.

4) Data warehouse: The data warehouse is the central repository where the integrated and processed data is stored. It is designed to support efficient querying and analysis. The data warehouse is typically optimized for read-intensive operations and provides a consolidated view of the data from multiple sources. It often employs a relational database management system (RDBMS) or a specialized data warehouse platform.

5) Data marts: Data marts are subsets of the data warehouse that are tailored to specific business functions or departments. Data marts contain a subset of the data relevant to a particular area, such as sales, marketing, finance, or operations. They are designed to provide focused and pre-aggregated data for faster and more targeted analysis.

6) Metadata repository: The metadata repository stores information about the structure, semantics, and lineage of the data in the data warehouse. It includes metadata such as data definitions, data mappings, data lineage, business rules, and data transformation rules. The metadata repository helps users understand and interpret the data in the data warehouse, ensuring data consistency and facilitating data governance.

7) Business intelligence tools: Business intelligence (BI) tools are used to access, analyze, and visualize the data stored in the data warehouse. These tools provide end-users with the ability to create reports, dashboards, and perform ad-hoc queries to gain insights from the data. BI tools often offer features like data visualization, data mining, and advanced analytics to support decision-making processes. The foundational architecture of data warehouses provides a structured framework for storing, integrating, and analyzing large volumes of data. It enables organizations to consolidate and transform data from multiple sources into a unified and consistent format, making it easier to extract meaningful insights and support data-driven decision-making.

B. Security Methodologies

Securing data warehouses typically involves a combination of methodologies and best practices. Here are several commonly employed methodologies [9], [10], and [11]:

1) Access control: Access control methodologies focus on managing user access to the data warehouse. This includes implementing strong authentication mechanisms, such as multi-factor authentication, to verify the identity of users. Role-based access control (RBAC) is often employed to assign appropriate privileges and permissions based on user roles and responsibilities. Access control lists (ACLs) and data-level security can be used to restrict access to specific data objects or rows within the warehouse.

2) Encryption: Encryption is a widely adopted methodology for protecting data in transit and at rest within a data warehouse. Transport Layer Security (TLS) or Secure Sockets Layer (SSL) protocols can be used to encrypt data during transmission between components of the data warehouse. Data at rest can be protected using techniques such as full-disk encryption, database-level encryption, or column-level encryption. Encryption keys should be securely managed and stored to prevent unauthorized access.

3) Data masking and anonymization: Data masking and anonymization methodologies involve modifying or obfuscating sensitive data to protect its confidentiality. Techniques like tokenization, pseudonymization, or data substitution can be used to replace sensitive information with fictitious values while preserving the format and structure of the data. Data masking can be applied during data extraction or as part of the data loading process into the data warehouse.

4) Auditing and monitoring: Auditing and monitoring methodologies involve capturing and analyzing activities within the data warehouse to detect and respond to security incidents. Robust logging mechanisms should be implemented to record user activities, system events, data changes, and access attempts. Security Information and Event Management (SIEM) systems can be employed to collect and analyze log data, generate alerts, and facilitate incident response.

5) Data classification and Data Loss Prevention (DLP): Data classification methodologies help identify and categorize sensitive or confidential data within the data warehouse. By classifying data based on its sensitivity, organizations can apply appropriate security controls and data protection measures. Data Loss Prevention (DLP) technologies can be used to monitor and prevent unauthorized data exfiltration or leakage by applying policies and rules to sensitive data.

6) Vulnerability management: Vulnerability management methodologies involve regularly scanning the data warehouse infrastructure, databases, and applications for known vulnerabilities. Vulnerability assessment tools can identify security weaknesses and misconfigurations that could be exploited by attackers. Patch management processes should be implemented to promptly apply security patches and updates to mitigate identified vulnerabilities.

7) Disaster recovery: Incident response and disaster recovery methodologies focus on preparedness and response to security incidents or catastrophic events. Incident response plans should be established, outlining the steps to be taken in the event of a security breach. Disaster recovery strategies should be in place to ensure timely recovery of the data warehouse in case of system failures, cyber-attacks, or natural disasters.

8) Data governance and training: Data governance methodologies establish policies, procedures, and guidelines for managing and protecting data within the warehouse. This includes defining data ownership, accountability, and data lifecycle management practices. Regular training and awareness programs should be conducted to educate users and stakeholders about data security best practices, policies, and their roles in maintaining data warehouse security. These methodologies, when implemented collectively, contribute to the overall security of data warehouses. Organizations should adopt a layered approach, combining multiple security

methodologies, to create a robust security framework that protects critical business information and ensures compliance with relevant regulations.

Besides, the architecture impacts the following security aspects:

1) Network security: The architecture influences network security considerations, particularly in distributed data warehouse environments. It determines how data flows between different components of the data warehouse, including data sources, staging area, data warehouse, and data marts. Secure network architecture includes measures like network segmentation, firewalls, encryption, and intrusion detection systems to protect data during transmission and prevent unauthorized access.

2) Scalability and performance: Architecture considerations impact security implementations concerning scalability and performance. A scalable architecture can handle increasing data volumes, user loads, and concurrent queries without compromising security. It should accommodate security measures without significantly impacting system performance, ensuring that security controls do not hinder data warehouse operations.

C. Applications of Data Warehouse in Real Life

The significance of a data warehouse is undeniable due to its numerous advantages. It eliminates the need for management choices to be based on limited and inaccurate data, while also assisting firms in avoiding various issues. Therefore, it is imperative for any organization to have a data warehouse. When discussing the importance of data warehousing (DW), it is noted that certain application areas require the presence and integration of data across the entire organization. Additionally, the ability to make quick decisions based on both real-time and historical data provides specialized information for loosely defined systems.

1) Business: The primary motivations for implementing a data warehouse in a firm are to enhance decision-making and improve organizational performance [12]. The utilization of data warehouses in various applications is influenced by the importance of business. All other non-governmental and partially non-governmental organizations fall under its authority. A data warehouse employs a unified repository to conveniently store data that is retrieved from various databases [13]. This data repository offers forecasting services that assist business professionals and managers. This comprehensive process is utilized to facilitate the identification of business requirements and the formulation of a business strategy [14]. The impact of several disciplines on data warehousing in business, ranging from significant to trivial, is examined.

a) Social media websites: Social media serves as a prime illustration of data warehousing. The social media industry is growing, and as a result, there is a growing demand to deploy data warehousing in this sector. Several characteristics seen on Facebook, Twitter, and other social media platforms are derived from the examination of extensive datasets. The system collects many types of data, such as groups, likes, friends, and geographical mapping, and saves it in a unified central repository. While several databases keep this information separately, the most important and meaningful data is saved in a centralized aggregated database [15].

b) Construction (material-based industries): The utilization of a data warehouse in the construction sector proves to be effective in facilitating decision-making processes. This strategy equips construction managers with comprehensive access to both internal and external data, enabling them to assess and oversee construction performance. The implementation of data warehousing in the construction industry demonstrates the ability of construction managers to effectively assess the remaining stock, track inventory trends associated with materials, and determine the quantity and cost of each material. To ensure the proper allocation of resources, it is important to consider the necessary services, maintenance, and operation of the systems, as well as the allocation of financial budgets. Additionally, good management of long-term investment plans and identification of potential hazards are crucial [16], [17].

c) Manufacturing industry: Data warehouse plays a crucial part in the maintenance of household and industrial operations. The manufacturing industry encompasses activities such as product and process design, scheduling, planning, production, maintenance, and substantial investments in equipment, labor, and heavy machinery. The actions taken in this situation will have significant impacts on both profitability and long-term strategic considerations. Several industries are seeking to transform themselves, and it is advisable for many of them to embrace data warehousing (DW) technology instead of relying on traditional decision-making methods. By implementing a data warehouse, organizations can collect, standardize, and store data from different applications. This enables them to streamline processes and enhance efficiency, as analyzing data across multiple applications can be a time-consuming task. During this phase, manufacturing and construction companies frequently employ transaction processing systems that are regularly updated to facilitate their ordinary business operations [18].

d) Banking: The banking industry is classified as one of the most information-intensive sectors in the business world. The relevance of business intelligence (BI) in banking operations has significantly increased due to advancements in the information technology industry [19]. The rapid pace of corporate growth and intensifying competition has underscored the critical importance of banking intelligence. Bank intelligence refers to the capacity to collect, oversee, and scrutinize a substantial volume of data pertaining to bank clientele, products, operations, services, suppliers, partners, and transactions. As the volume of data grows, the process of collecting, managing, and converting it into valuable insights becomes increasingly challenging. Data warehousing (DW) offers a solution to this challenge. Several data warehouse

variants are specifically tailored to cater to the needs of the banking industry [20].

e) Education: Data warehousing (DW) is gaining increasing popularity in the realm of education. The utilization of Data Warehousing (DW) in the educational sector offers numerous advantages in facilitating informed decision-making and timely data evaluation, which are the primary objectives of the DW process. DW offers a comprehensive and unified perspective of an institute. Most of the relevant departments utilize a data warehouse as a primary source of information regarding teachers and students. DW facilitates expedient access to students' results and notes from a web-based database via a student portal. Additionally, it aids in decision-making by offering both current and historical information pertaining to the institute [21].

f) Finance: The progress of technology, particularly in the IT industry, has introduced innovative approaches to managing financial processes in business. The government and business sectors play equally significant roles in the field of finance. Financial systems encompass many institutions such as banks, post offices, insurance firms, income tax departments, and other tax agencies. The use of a data warehouse in the financial industry offers numerous advantages, such as enhancing transparency in account opening and transactions. Likewise, the government has the authority to make decisions to address any financial crises. These systems possess sufficient intelligence to detect individuals who have failed to meet their obligations and may respond accordingly based on the circumstances. Efficient decision-making can be easily achieved in this case due to the maintenance of data warehousing [22].

2) Government: The government can employ data warehousing techniques in various sectors, such as looking for terrorist profiles and conducting threat assessments, improving agricultural practices, enhancing educational systems, optimizing financial operations, streamlining medical departments, and detecting fraudulent activities. The telecommunication and banking industries are plagued by numerous difficulties pertaining to user fraud [23].

a) Medical: The medical sector is currently leading in the implementation of data warehousing technology. The importance of data quality and the need for high-quality medical services has significantly increased in the field of health care. The complexity and diversity of medical and clinical data resulted in a slower adoption of data warehouses in the healthcare industry compared to other sectors. In recent years, there has been a significant increase in the utilization of data warehouses in both administrative and clinical domains. Data warehouses have the potential to enhance the quality of care provided to individual patients. These healthcare organizations are implementing data warehousing as a tool to help strategic decision-making. It offers the means to obtain medical data, extract pertinent information from that data, and disseminate this knowledge to all relevant individuals. The administrative data stored in a data warehouse can be utilized to obtain information regarding the required competent staff for a specific treatment. This information is then used for scheduling treatments and providing support to medical personnel in the field of human resources [24].

b) Fraud and threat detection: Governments are actively engaged in detecting and mitigating threats and fraudulent activities perpetrated by individuals with malicious intentions. Regrettably, there is a scarcity of available known implementations of data warehouses. Government entities have access to data warehouses; nevertheless, they require a comprehensive data warehouse system that is interconnected to all areas to effectively monitor threats and terrorists [23].

III. Security Approaches for DWH

A data warehouse is a crucial component of an organization, providing users with the ability to access comprehensive information about the whole business process. As stated in reference [25], ensuring security is a crucial necessity in all stages of data warehouse construction, including requirements, implementation, and maintenance. The security measures implemented for online transactional processing (OLTP) systems are not suitable for data warehouses [26]. In OLTP, security controls are applied at the level of rows, columns, or tables. However, data warehouses require access by varying numbers of users for different content due to their multidimensional nature, which is a fundamental principle of a data warehouse. Prior to loading the data into the data warehouse, the processes of data extraction, transformation, cleansing, and preparation have all been completed. Security considerations must be considered at every level of a data warehouse system. Furthermore, it is imperative to address the security of the underlying operating system and network to maintain data warehouse security [27]. The data warehouse literature has presented several security solutions, which can be classed based on how they meet fundamental security concerns, including Confidentiality, Integrity, and Availability.

A. DWH Security Approaches for Confidentiality Issues

The emphasis of confidentiality is on preventing information from being improperly discovered, either directly or through logical inference [28]. Numerous access control-related strategies have been put out to resolve concerns about data warehouse confidentiality. The administration and invocation of the source databases and the data warehouse are both under the supervision of access-control mechanisms. In a data warehouse setting, authentication and auditing systems are likewise categorized as access control and need to be set up. In [29], the author introduced a role-based authorization model and distinguished between two types of roles: the operations role, which initiates the associated procedures, and the developer role, which oversees extracting, integrating, and transforming data scripts. These positions just need to execute trustworthy procedures; they do not require direct access to data. Permissions to access data are assigned based on roles. Additional rights may be issued as needed to access more data in the event of failures or issues, but audits must keep an eye on these permissions. Traditionally, high-level users like business analysts and upper management have had access to data warehouses. As a result, serious problems with access control also surface at the data warehouse's front end. Since it

impedes the discovery of analytical information, the majority of data warehouse or OLAP suppliers believe that fine-grained access-control functionality for a data warehouse front end is unnecessary. This assumption is incorrect, though, as many users have access to analytical tools that allow them to query the data warehouse. Applications for front-end data warehouses can offer both dynamic and static reporting. Because access control may be specified report-by-report, it is not problematic to impose it on static reports. It is challenging to implement suitable access-control measures for dynamic reporting, such as data-mining queries. This brings up the issue of data inference; for instance, a person could be able to access specific information through an aggregated query even when they are not permitted to do so [30].

B. DWH Security Approaches for Integrity

Integrity refers to safeguarding data from unauthorized or malevolent modifications, including the insertion, infection, or deletion of false data [31]. One drawback of access-control techniques is that, in the event of an aggregated OLAP query, they are unable to capture conclusions about the data. Data inferences result in integrity problems. Inference-control techniques have been researched in statistics and census databases for over thirty years [32], [33], and [34]. The suggested methods fall into two categories: perturbation-based and restriction-based methods. To stop malicious inference, restriction-based inference control systems merely refuse unsafe queries. Perturbation approaches can dynamically apply data modification to each query in addition to adding noise, swapping, or altering the original data. The methods put out to address the integrity problem can be further categorized as outlined below.

1) Restriction-based approaches: The greatest number of values aggregated by distinct questions, the minimum number of values aggregated by a query, and the highest rank of the matrix expressing answered queries are used in restriction-based inference-control techniques to establish the safety of a query. Sensitive data can also be protected by partitioning and cell suppression. Cells with low COUNT values can have suppression applied to them to identify inference in the data. Techniques based on linear programming can be used to eliminate inferences. This kind of detection technique only functions with two-dimensional tables; three- or higher-dimensional tables are not compatible with it [35].

2) Combined access-and inference-control approaches: To effectively eliminate security risks, the combination of access control and inference control can offer a robust solution. Preserving the security of data warehouse and OLAP systems should not compromise their functionality. The author in [36] suggested three-tier security architecture for a data warehouse. Statistical databases typically consist of two tiers: sensitive data and aggregate queries. The two-tier design mentioned above has certain inherent limitations. One problem is that doing inference checking during run-time query processing might lead to undesirable delays. Additionally, under this architecture, inference-control techniques are unable to take advantage of the unique features of OLAP. To address these limitations, the study has established a three-tier framework to facilitate access control between the first and second layers, as well as inference control between the second and third tiers. The suggested design mitigates superfluous delays caused by inference checking through various means. Implementing these techniques can decrease the size of the inputs to inference control systems, hence reducing complexity.

3) Modeling-based approaches to DWH security: In their publication [37], the author introduced a conceptual-level Access and Audit Control (ACA) model for data-warehouse modeling. This model is founded on data classification. The document outlined three security regulations: permission regulations for people and objects, assignment regulations for sensitive information that establish multilevel security procedures, and audit regulations that examine user actions at all stages and points at the conceptual level. The ACA model is incorporated in multi-dimensional modeling to enhance UML skills in the design of secure data warehouse systems.

4) Data masking and perturbation-based security approaches: In their publication [38], the author introduced a data-masking technique specifically designed for data warehouses that exclusively contain numerical values. The proposed methodology relied on mathematical modulus operators, including division, remainder, and two basic arithmetic operations. These operators can be implemented without modifying the source code of the database management system (DBMS) or user applications. According to their assertion, the suggested formula necessitated minimal computational resources. Consequently, the additional time required for query response was insignificant, while maintaining an adequate level of security.

C. DWH Security Approaches for Availability Issues

Ensuring the availability of data is crucial in every data warehouse system. This entails the retrieval of data that has been affected by immediate corruption. Data replication is implemented to facilitate the restoration of corrupted data using various suggested methods. By using this approach, it is possible to prevent database downtime caused by maintenance interventions and distribute query-processing efforts to prevent data-access hot spots. Familiar RAID architectures can be employed to mirror data [39], [40] in systems where centralized servers house the database. Nevertheless, corporations have started deploying their data warehouses on inexpensive processors to achieve cost efficiency. RAID technology is unsuitable for this situation due to the presence of only one disk drive, which is normally the case.

IV. Research Challenges and Opportunities

While typical encryption methods can offer robust data privacy and are present in today's main DBMS, their influence on database speed renders their use in data warehouses impractical. As previously demonstrated, the computational overhead incurred by methods such as AES and 3DES significantly affects performance. Options that can provide a high degree of privacy while reducing the overhead in query

response time are required. Given bitwise operations' simplicity and speed, bit-based encryption algorithms might offer a means of achieving novel, workable solutions. Naturally, the degree of privacy will decrease if the encryption procedure is simplified to increase database speed. It is necessary to create a tradeoff compromise that minimizes the impact on performance while maintaining the desired level of privacy. A further option would be the creation of query engines that could handle queries on encrypted data directly, i.e., without the need to first decrypt the data[41].

V. CONCLUSION

This study has conducted a comprehensive analysis of the security solutions available for data warehouses, examining their limitations and the effects they have on the scalability and performance needs of these warehouses. The suggested methods are impractical or ineffective for implementation in data warehouse systems. A data warehouse necessitates specific capabilities that must adhere to strict scalability and performance criteria. Hence, a comprehensive solution is required to effectively tackle these directives. Data warehouse security is a pertinent area of ongoing research that holds significance for all industrial projects. Additional investigation into data warehouse security is necessary to tackle the difficulties, as there are other variables that still need to be considered and some unanswered questions.

REFERENCES

[1] Keshta, I. and A. Odeh., Security and privacy of electronic health records: Concerns and challenges. Egyptian Informative Journal, 2021. 22(2): p. 177–183.

[2] Wixom, B.H. and H.J. Watson., An Empirical Investigation of the Factors Affecting Data Warehousing Success,. MIS quarterly, 2001. 25: p. 17-41.

[3] Santos, R.J., J. Bernardino, and M. Vieira, A survey on data security in data warehousing: Issues, challenges and opportunities. 2011 IEEE EUROCON - International Conference on Computer as a Tool,, 2011: p. 1-4.

[4] Lincke, S., Attending to Information Privacy, in Information Security Planning: A Practical Approach. 2024, Springer. p. 185-200.

[5] Samarati, P. and S.D.C.d. Vimercati. Data protection in outsourcing scenarios. in Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security - ASPICS '2010.

[6] Dash, S., et al., Big data in healthcare: management, analysis and future prospects. Journal of Big Data, 2019. 6.

[7] Santos, R.J., J. Bernardino, and M. Vieira, A survey on data security in data warehousing: Issues, challenges and opportunities. 2011 IEEE EUROCON - International Conference on Computer as a Tool,, Apr. 2011.

[8] Ariyachandra, T. and H. Watson, Key organizational factors in data warehouse architecture selection. Decision Support Systems, 2010. 49.

[9] Santos, R.J., J. Bernardino, and M. Vieira. A data masking technique for data warehouses. In Proceedings of the 1st Symposium on International Database Engineering & Applications (IDEAS '11). in Association for Computing Machinery. 2011. New York, NY, USA.

[10] Santos, R.J., et al. A Specific Encryption Solution for Data Warehouses. in Database Systems for Advanced Applications: 18th International Conference, DASFAA 2013. 2013. China: Springer

[11] Fernández-Medina, E., R.V. J. Trujillo, and M. Piattini, Access control and audit model for the multidimensional modeling of data warehouses. Decision Support Systems, 2006. 42(3): p. 1270–1289.

[12] Watson, H.J., D.L. Goodhue, and B.H. Wixom, The benefits of data warehousing: why some organizations realize exceptional payoffs. Information & Management, 2002. 39(6): p. 491-502.

[13] Dinesh, L. and K.G. Devi, An efficient hybrid optimization of ETL process in data warehouse of cloud architecture. Journal of Cloud Computing, 2024. 13(1): p. 12.

[14] Joseph, M.V., Significance of Data Warehousing and Data Mining in Business Applications. International Journal of Soft Computing and Engineering, 2013. 3(1): p. 2231-2307.

[15] Thusoo, A., et al. Data warehousing and analytics infrastructure at face. . in Proceedings of the 2010 ACM SIGMOD International Conference on Management of data (SIGMOD '10). 2010. . New York, NY, USA: Association for Computing Machinery.

[16] R. Chowdhury, et al., Implementation of Central Dogma Based Cryptographic Algorithm in Data Warehouse Architecture for Performance Enhancement. International Journal of Advanced Computer Science and Applications, 2015. 6.

[17] Park, T. and H. Kim, A data warehouse-based decision support system for sewer infrastructure management. Automation in Construction 2013. 30: p. 37–49.

[18] Sarda, N.L. Temporal issues in data warehouse systems. in Proceedings 1999 International Symposium on Database Applications in Non-Traditional Environments (DANTE'99). 1999. Japan.

[19] Bany Mohammed, A., et al., Towards an understanding of business intelligence and analytics usage: Evidence from the banking industry. International Journal of Information Management Data Insights, 2024. 4(1): p. 100215.

[20] Sarkar, A., Data Warehouse Requirements Analysis Framework: Business-Object Based Approach. International Journal of Advanced Computer Science and Applications, 2012. 3.

[21] Goyal, M. and R. Vohra, Applications of data mining in higher education. International Journal of Computer Science Issues (IJCSI), 2012. 9(2).

[22] Chau, K.W., et al., Application of data warehouse and Decision Support System in construction management. Automation in construction, 2003. 12(2): p. 213–224.

[23] Bilal, M., et al., Application of Data Warehouse in Real Life: State-of-the-art Survey from User Preferences' Perspective. International Journal of Advanced Computer Science and Applications, 2016. 7.

[24] Stolba, N. and A.M. 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, 2006. 3.

[25] Devanbu, P.T. and S. Stubblebine. Software engineering for security: a roadmap. In Proceedings of the Conference on The Future of Software Engineering (ICSE '00). in Association for Computing Machinery. 2000.

[26] Hoi, L.M., W. Ke, and S.K. Im, Manipulating Data Lakes Intelligently With Java Annotations. IEEE Access, 2024. 12: p. 34903-34917.

[27] Bellatreche, L., ed., Security in Data Warehouses, Data Warehousing Design and Advanced Engineering Applications: Methods for Complex Construction. 2010: IGI Global.

[28] Farkas, C. and S. Jajodia, The inference problem: a survey. ACM SIGKDD Explorations Newsletter 2002. 4(2): p. 6-11.

[29] Doshi, V., S. Jajoda, and A. Rosenthal, A programmatic approach to access control in the Data Warehouse. Personal notes, 1999.

[30] Aleem, S., Luiz Fernando Capretz and F. Ahmed, Data security approaches and solutions for data warehouse. International Journal of Computers, 2015. 9: p. 91-97.

[31] Georgiev, A. and V. Valkanov, Custom data quality mechanism in Data Warehouse facilitated by data integrity checks. Mathematics and Education in Mathematics, 2024. 53: p. 67-75.

[32] Adam, N.R. and J.C. Worthmann, Security-control methods for statistical databases: a comparative study. ACM Computing Surveys (CSUR), 1989. 21(4): p. 515–556.

[33] Denning, D.E. and J. Schlorer, Inference Controls for Statistical Databases. Computer, 1983. 16(7): p. 69–82.

[34] Willenborg, L. and T.D. Waal., Statistical disclosure control in practice Springer Science & Business Media, T1996. 111.

[35] Cox, L.H., On properties of multi-dimensional statistical tables. Journal of Statistical Planning and Inference, 2003. 117(2): p. 251–273.

[36] Jajodia, L.W.a.S., Security in Data Warehouses and OLAP systems. Handbook of Database Security: Applications and Trends 2008: p. 191-212.

[37] Fernández-Medina, E., et al., Access control and audit model for the multidimensional modeling of data warehouses. Decision Support Systems, 2006 42(3): p. 1270-1289.

[38] Santos, R.V., J. Bernardino, and M. Vieira. A data masking technique for data warehouses. in Proceedings of the 15th Symposium on International Database Engineering & Applications. 2011.

[39] Wu, X., et al. RAID-Aware SSD: Improving the Write Performance and Lifespan of SSD in SSD-Based RAID-5 System. in IEEE Fourth International Conference on Big Data and Cloud Computing. 2014. Australia.

[40] Liu, W., et al. Understanding the SWD-based RAID System. in International Conference on Cloud Computing and Big Data. 2014. China.

[41] Santos, R.J., J. Bernardino, and M. Vieira, A survey on data security in data warehousing: Issues, challenges and opportunities. 2011 IEEE EUROCON - International Conference on Computer as a Tool, 2011.