

Empowering Librarians with Data: A Comparative Analysis of Data Warehousing Tools

Subham Das^a and Dr. Swapan Khan^b

^aResearch Scholar, Department of Library and Information Science, Jadavpur University, Kolkata-60

Email: realmesdas@gmail.com;

^bLibrarian, Narasinha Dutt College, Howrah -711101

Email: khanswapan@gmail.com;

Received: 22 June 2024; Accepted: 23 December 2024

In the face of exponential data growth, libraries grapple with effectively managing and extracting knowledge from massive datasets. Data warehousing technologies have emerged as indispensable tools to address this challenge. This research investigates the significance of data warehousing within the library domain, focusing on its capacity to streamline data management and analysis processes. A comparative assessment of ten prominent data warehousing platforms is conducted, including Teradata Vantage, OCLC WorldShare Analytics, Snowflake, Amazon Redshift, IBM Db2 Warehouse, Alma, Oracle Autonomous Data Warehouse, SAP Data Warehouse Cloud, Microsoft Azure Synapse Analytics, and Google BigQuery. This evaluation scrutinizes crucial aspects such as architectural design, user-friendliness, applicability within library contexts, security mechanisms, distinctive features, and, crucially, cost-effectiveness. Particular emphasis is placed on identifying solutions that harmonize functionality with the budgetary realities of library operations. The study elucidates the strengths and shortcomings of each platform, providing actionable guidance for libraries to select the most suitable data warehousing solution based on their unique operational requirements and financial limitations. This research underscores the imperative of adopting a suitable data warehousing tool to augment data management and analytical capabilities within libraries, thereby facilitating informed decision-making and enhancing service delivery.

Keywords: Big Data, Data Warehouse Tools, Data Management, Cost-Effectiveness, Comparative Analysis.

1. Introduction

Libraries increasingly rely on solid data management strategies in the face of ever-increasing volumes of information. Data warehousing emerges as an essential solution that provides a framework for storing, organizing, and analyzing large collections of data from various sources¹. This includes gathering data from diverse library systems, patron records, resource usage statistics, and external sources, then consolidating them into a central repository. The primary objective of data warehousing within a library context is to gain an all-inclusive understanding of library operations, patron behavior, and resource utilization². This enriched understanding empowers libraries to make data-driven decisions regarding resource allocation, collection development, service delivery, and outreach programs.

With the progress in data-driven technology, Data Warehousing serves as an important tool for libraries, regardless of their magnitude. Libraries that

consolidate information into a single source can make its decision-making processes easier, improve operational efficiency, and identify insights that can be leveraged to deliver better services and a more satisfying experience to patrons. Given the growing volume of data generated by library systems, data warehousing is an essential strategy for managing and analyzing information to meet the needs of an evolving library community³.

2. Literature Review

According to Laitinen & Saarti (2012)⁴, the data warehouse is necessary for storing historical data in an integrated way, so that data is significantly useful for decision support systems. In the context of libraries, this tool is very useful; such exists various types of statistical data aggregation. Analysing aggregate data of libraries impact on parent organization identifies their operational effectiveness, and provides statistical insights to make this

improvement in future which is crucial area in terms of services quality.

Yang & Shieh (2016)⁵ further highlight the diverse applications of data warehouse tools in libraries. These tools are not limited to a singular function but are versatile in their uses, ranging from performance measurement and resource allocation to enhancing user experience and supporting research activities. The adaptability of data warehouse tools makes them an invaluable asset in the ever-evolving landscape of library services.

However, the implementation of data warehousing is not without its challenges. As discussed by L C & R K (2023)⁶, the concept of data warehousing encompasses numerous complexities that can pose significant hurdles. These challenges include data integration, ensuring data quality, handling large volumes of data, and maintaining the security and privacy of sensitive information. Effective data warehousing requires a deep understanding of these challenges and the development of robust strategies to mitigate them.

Sureddy & Yallamula (2020)⁷ emphasize the importance of selecting the appropriate data warehouse tool for an enterprise. They outline various techniques and criteria that should be explored to make an informed choice. These techniques include evaluating the specific needs of the enterprise, considering the scalability and flexibility of the tool, assessing the cost-benefit ratio, and ensuring the tool's compatibility with existing systems and workflows. By thoroughly exploring these factors, organizations can select the most suitable data warehouse tool to meet their unique requirements and drive their data management and analysis efforts forward.

3. Objectives of the Study

- Conduct a comparative analysis of various data warehousing tools to determine their suitability for meeting the specific data management and analysis needs of libraries.
- Evaluate the cost-effectiveness of different data warehousing tool types, considering factors such as licensing fees, maintenance costs, and scalability.
- Identify and prioritize data warehousing solutions that best align with the unique operational requirements and budgetary constraints of libraries

4. Research Methodology

This study employs a mixed sampling approach, integrating purposive, convenience, and stratified sampling techniques to ensure comprehensive coverage of relevant tools. The sample size has been expanded to include ten tools, selected to represent a mix of:

- Library-specific tools: Chosen for their direct relevance to libraries.
- Established market leaders: Selected based on their widespread adoption and strong reputation.
- Emerging and niche tools: Included to account for innovative features and potential benefits.

This selection ensures a balanced evaluation of both widely used and emerging tools, providing insights that cater to diverse library needs, from small institutions to large-scale operations.

4.1 Sampling

The ten data warehouse tools selected to be analysed in this study are as follows:

- i. Amazon Redshift
- ii. Google BigQuery
- iii. Microsoft Azure Synapse Analytics
- iv. Snowflake
- v. Alma
- vi. OCLC WorldShare Analytics
- vii. IBM Db2 Warehouse
- viii. Teradata Vantage
- ix. Oracle Autonomous Data Warehouse
- x. SAP Data Warehouse Cloud

Among these, Alma and OCLC WorldShare Analytics are specialized tools designed for libraries, while the others are general-purpose tools that libraries can adopt based on specific needs. The inclusion of Oracle Autonomous Data Warehouse and SAP Data Warehouse Cloud ensures representation of tools with cutting-edge automation and integration features, catering to libraries seeking scalability and efficiency.

4.2 Points of Comparison

This study would compares ten tools frequently used by libraries for data warehousing, analytics, and management, focusing on the following critical aspects which are chosen after literature search of other related studies of data warehousing⁸:

- i. **Year of Establishment:** The foundation year of each tool, highlighting its maturity and longevity in the market.

- ii. **Architecture:** The underlying structural framework of the tool, such as cloud-based, columnar, or massively parallel processing, catering to the diverse needs of libraries.
- iii. **Special Features:** Unique functionalities that differentiate each tool, emphasizing capabilities such as automation, AI integration, scalability, or data sharing, which libraries can leverage for operational efficiency.
- iv. **Type of Clients:** The primary clientele served by the tool, including libraries, healthcare, education, retail, and finance sectors, showcasing its versatility and applicability in library operations.
- v. **Recommended for:** Specific use cases where the tool excels, tailored to the requirements of libraries such as integrated library systems, complex data analytics, or resource management.
- vi. **Pricing or Cost of Acquisition:** Information on the pricing models, from pay-as-you-go to custom pricing, enabling libraries to evaluate cost-effectiveness based on their budgets.
- vii. **Key Benefits:** The standout advantages of each tool, such as scalability, advanced reporting, or integration capabilities, that address critical library needs.
- viii. **Why Libraries Should Opt for This Tool:** Insight into how each tool aligns with library-specific demands, helping libraries make informed decisions for data management and analytics.
- ix. **After-Sales and Support:** Evaluation of the customer support services provided, including technical assistance, user communities, and training resources, ensuring smooth implementation and usage for libraries.
- x. This structured comparison allows libraries to identify the most suitable tools based on their unique operational needs and strategic objectives.

5. Scope and Limitations

This study's scope is confined to a curated sample of ten tools, which, while diverse, may not represent all available options in the rapidly evolving data warehousing market. The focus on tools with strong market presence or emerging relevance could exclude lesser-known alternatives. The comparative analysis is limited to the specified parameters, which may not cover all aspects critical to every library. Furthermore, findings are primarily tailored to libraries venturing into data warehousing for the first time, with limited application to advanced or highly customized use cases.

Additionally, the inclusion of tools offering general-purpose functionality may overlook niche tools developed for highly specific library environments.

6. An Overview of Selected Data Warehousing Tools

6.1 Amazon Redshift

Amazon Redshift, part of Amazon Web Services(AWS) is a next generation cloud data warehousing service designed for executing large petabyte-scale databases efficiently. Designed for patterns in data, it boasts a columnar storage architecture that manages structured, semi-structured, and unstructured data. Easily integrates with other AWS services like S3, EMR, and Data Pipeline for loading data into Redshift and analysing⁹.

It provides features like automatic backups, snapshots, and encryption that ensure the maximum security of data, making it an ideal selection for working with sensitive data. It provides flexible pricing models, such as on-demand and reserved instances, allowing businesses to achieve cost efficiency. Moreover, Redshift works with a host of BI tools and SQL interfaces, while also providing fast query processing power. This makes Redshift a very scalable, economical solution for data analytical purposes.

6.2 Google BigQuery

Google BigQuery, a comprehensive cloud-based data warehousing and analytics solution from Google Cloud, efficiently handles structured, semi-structured, and unstructured data utilizing a columnar storage format for optimal query performance. It offers features such as automatic scaling, real-time data streaming ingestion, and seamless integration with other Google Cloud services like Cloud Storage and Cloud Dataflow, streamlining the processing of large-scale data¹⁰.

BigQuery prioritizes security with multi-layered controls and encryption while offering flexible pricing options to accommodate organizations of varying sizes. Its compatibility with various programming languages and analytics tools ensures that it remains a versatile, scalable, and cost-effective option for managing and analyzing extensive datasets.

6.3 Microsoft Azure Synapse Analytics

Azure Synapse Analytics by Microsoft is a cloud-based service that integrates big data analytics, data integration, and business intelligence to handle large-

scale data processing. Its columnar storage architecture enables fast query processing and seamless integration with Azure services like Azure Data Lake Storage and Azure Data Factory. Azure Synapse Analytics offers robust security features, including encryption and access controls, to ensure compliance with regulatory standards. The platform's flexible pricing options and collaborative workspace make it suitable for complex analytics projects. It is ideal for organizations seeking a scalable and efficient solution for deriving actionable insights¹¹.

6.4 Snowflake

Snowflake is a cloud-based data warehousing platform with a unique architecture that separates storage, compute, and services, enabling on-demand scalability. Unlike traditional shared infrastructures, Snowflake allows organizations to scale resources independently, ensuring flexibility and cost efficiency. It seamlessly integrates with various third-party tools, such as data integration and business intelligence platforms, for enhanced customization. Snowflake's advanced security features, including encryption, multi-factor authentication, and data masking, ensure data privacy and compliance with regulatory standards. Its scalability, integration capabilities, and robust security features position Snowflake as a leader in cloud-based data warehousing.¹²

6.5 IBM Db2 Warehouse

IBM Db2 Warehouse is an on-premises and cloud-based solution tailored for enterprise-scale analytics. It supports hybrid deployments, allowing organizations to seamlessly manage and analyze data across diverse environments. The platform's in-memory data processing delivers high-speed query execution, while built-in machine learning tools enable advanced analytics. Db2 Warehouse's integration with IBM's AI and cognitive computing tools provides added value for predictive and prescriptive analytics. Its focus on enterprise-grade security, compliance, and hybrid flexibility makes it a preferred choice for organizations requiring robust data warehousing solutions¹³.

6.6 Oracle Autonomous Data Warehouse

Oracle Autonomous Data Warehouse leverages machine learning to automate tasks such as indexing, tuning, and patching, reducing administrative overhead. It supports structured and semi-structured data, with features like elastic scaling and automated performance optimization. The platform's integration

with Oracle Analytics Cloud and third-party tools enhances its utility for complex data analytics. Advanced security measures, including always-on encryption and data masking, ensure compliance and data protection. Oracle's self-optimization capabilities make it a suitable choice for organizations looking to reduce operational complexities while achieving high performance¹⁴.

6.7 SAP Data Warehouse Cloud

SAP Data Warehouse Cloud is a unified platform designed for data management and analytics. Its integration with SAP applications, such as SAP HANA and SAP Analytics Cloud, enables seamless workflows for enterprise systems. The platform supports real-time analytics and offers self-service capabilities, allowing non-technical users to explore data independently. SAP Data Warehouse Cloud's pricing flexibility and focus on integration make it particularly appealing to organizations leveraging SAP's broader ecosystem¹⁵.

6.8 Alma

Alma is a cloud-based integrated library system (ILS) offering comprehensive data warehousing capabilities tailored for libraries. It consolidates data from multiple sources into a unified platform, supporting structured, semi-structured, and unstructured data. Advanced analytics and reporting tools allow libraries to derive actionable insights, while customizable reports cater to specific institutional needs. Alma integrates with various third-party systems, including discovery tools and learning management platforms, ensuring seamless resource management. Its robust data warehousing and analytics features make it a valuable asset for libraries aiming to enhance data-driven decision-making.¹⁶

6.9 OCLC WorldShare Analytics

OCLC WorldShare Analytics is a cloud-based platform designed specifically for libraries, providing robust data warehousing and analytics capabilities. It consolidates data from integrated library systems (ILS), circulation records, and electronic resources into a centralized platform. The tool offers advanced reporting, trend analysis, and customizable dashboards, empowering libraries to make data-driven decisions. WorldShare Analytics also includes data governance features to ensure data accuracy, security, and compliance with industry standards. By enabling libraries to manage and analyze data effectively, WorldShare Analytics enhances service quality and operational efficiency¹⁷.

6.10 Teradata Vantage

Teradata Vantage is an enterprise-grade data analytics and warehousing platform designed for high-performance analytics. It supports multi-cloud and hybrid deployments, offering flexibility for organizations with diverse infrastructure needs. Vantage's advanced analytics capabilities, including predictive and prescriptive modeling, cater to complex data scenarios. Its scalability and support for multi-structured data make it suitable for handling large datasets. With features like workload management and integration with AI/ML tools, Teradata Vantage is ideal for organizations seeking robust analytics with enterprise-level scalability.¹⁸

This comprehensive overview highlights a diverse range of tools, offering libraries and organizations options that cater to a wide spectrum of needs, from specialized library solutions to enterprise-grade analytics platforms. A comprehensive table 1 presented below illustrates how these tools vary across key parameters, enabling libraries to identify the most suitable option based on their unique requirements and objectives.

7. Comparative Analysis of Data Analytics Tools for Libraries

This analysis compares various data analytics tools for libraries, focusing on cost-benefit and use case scenarios. The comparison considers architecture, features, target clients, pricing, benefits, suitability for libraries, and after-sales support.

7.1. Amazon Redshift (2012)

- **Cost-Benefit:** Starting at \$0.25 per hour, Redshift offers a cost-effective solution for large-scale data storage and complex analytics. Its pay-as-you-go model allows libraries to scale resources based on demand, optimizing costs.
- **Use Case:** Ideal for large libraries with complex data needs, requiring high-performance query processing and integration with AWS services. Research libraries with extensive datasets and complex analytical requirements would particularly benefit.
- **Key Benefit:** High scalability and integration with the broader AWS ecosystem.
- **Considerations:** Requires familiarity with AWS services for optimal utilization.

7.2. Google BigQuery (2010)

- **Cost-Benefit:** With a pay-as-you-go model starting at \$0.01 per query, BigQuery is highly

cost-effective for libraries with variable usage patterns. The serverless architecture eliminates infrastructure management costs.

- **Use Case:** Best suited for libraries already using Google Cloud Platform (GCP) and requiring serverless, scalable storage and analytics. Libraries needing real-time analytics and seamless integration with other Google services would find this tool particularly useful.
- **Key Benefit:** Serverless architecture reduces management overhead, and integration with GCP simplifies workflows.
- **Considerations:** Primarily beneficial for organizations already invested in the Google Cloud ecosystem.

7.3. Microsoft Azure Synapse Analytics (2019)

- **Cost-Benefit:** Starting at \$1.10 per hour, Synapse Analytics offers a robust solution for libraries already within the Microsoft Azure environment. The cost is justified by its integration with other Microsoft services and enhanced security features.
- **Use Case:** Perfect for libraries heavily reliant on Microsoft tools and services. Its advanced security features and seamless integration with Azure are key advantages for libraries handling sensitive data.
- **Key Benefit:** Strong integration with the Microsoft ecosystem and robust security features.
- **Considerations:** Most cost-effective for organizations already invested in the Microsoft Azure ecosystem.

7.4. Snowflake (2012)

- **Cost-Benefit:** Starting at \$40/TB/month, Snowflake is suitable for libraries with complex data needs, including semi-structured data. Its high scalability and performance justify the cost for organizations with growing data volumes.
- **Use Case:** Best for libraries needing flexibility with semi-structured data and real-time analytics. Its cross-cloud capabilities are beneficial for libraries with diverse IT environments.
- **Key Benefit:** Flexibility in handling diverse data formats and high scalability.
- **Considerations:** The cost may be higher for smaller libraries with limited data volumes.

7.5. Alma (2011)

- **Cost-Benefit:** Pricing is available on request, reflecting its tailored approach to library

Table 1 — Data Collection cum Comparison chart of Selected Data Warehousing Tools

Tool	Year of Establishment	Architecture	Special Features	Type of Clients	Recommended for	Pricing or Cost of Acquisition	Key Benefits	Why Libraries Should Opt for This Tool	After-Sales and Support
Amazon Redshift	2012	Columnar	Automatic backups, snapshots, encryption, flexible pricing	Libraries, Healthcare, Retail, Finance	Large libraries with complex data needs	Starts at \$0.25 per hour	High scalability, cost-effective, integration with AWS services	Ideal for large-scale data storage and complex data analytics needs in libraries. Works well for libraries that require fast query processing and integration with AWS resources.	Comprehensive support from AWS including 24/7 customer service, technical assistance, and documentation
Google BigQuery	2010	Columnar	Serverless, scalable, integrates with Google Cloud services	Libraries, E-commerce, Healthcare, Finance	Libraries using Google Cloud Platform	Pay-as-you-go, starts at \$0.01 per query	Scalable, fast performance, cost-effective, serverless	Great for libraries already using Google Cloud services or requiring serverless, scalable storage and analytics.	Extensive Google Cloud support, including 24/7 technical support, user community forums, and tutorials
Microsoft Azure Synapse Analytics	2019	Massively Parallel Processing	Integration with Azure, built-in security, serverless	Libraries, Healthcare, Finance, Retail	Libraries using Microsoft Azure	Starts at \$1.10 per hour	Integration with Microsoft services, security features	Perfect for libraries using Microsoft tools and services. Supports advanced security and integration with Azure.	Microsoft support services, including documentation, forums, and premium support plans
Snowflake	2012	Cloud-based, Columnar Storage	Zero-copy cloning, data sharing, cross-cloud capabilities	Libraries, Healthcare, Finance, E-commerce	Libraries with complex data needs, including semi-structured data	Starts at \$40/TB/month	High scalability, cost-effective, fast performance	Best for libraries needing flexibility with semi-structured data and real-time analytics. Highly scalable for growing data needs.	Snowflake offers 24/7 support with an active user community and documentation
Alma	2011	Cloud-based	Automation, analytics, open platform	Libraries, Higher Education, Research	Integrated library systems and resource management	Pricing available on request	Integrates with Ex Libris services, customizable workflows	Suitable for libraries seeking an integrated library management system (ILS) with powerful data warehousing and analytics tools tailored for library needs.	Ex Libris provides robust customer service, online help, and user community for Alma
OCLC World Share Analytics	2013	Cloud-based, Data Warehousing	Consolidation of library data, advanced reporting and analysis tools	Libraries, Higher Education, Research	Libraries needing OCLC integration and library-specific data warehousing	Custom pricing based on library size and usage	Advanced reporting tools, trend analysis, data governance	Best for libraries using OCLC products, with strong support for data consolidation, reporting, and data governance.	OCLC offers dedicated support for WorldShare Analytics, including helpdesk and training resources
IBM Db2 Warehouse	1983	Columnar	Advanced AI integration, scalable processing	Libraries, Retail, Healthcare, Finance	Libraries looking for AI and scalable storage solutions	Custom pricing	AI-powered, scalable, automated data management	Good for libraries seeking advanced AI capabilities and scalable data warehousing solutions. Offers an automated, self-driving management system.	IBM provides extensive customer support, including knowledge base, forums, and technical assistance
Oracle Autonomous Data Warehouse	2018	Cloud-based, Automated	Self-driving, autonomous data management	Libraries, Healthcare, Finance, E-commerce	Libraries needing an autonomous, self-managing data solution	Custom pricing	Self-driving, highly autonomous with advanced security	Ideal for libraries needing a fully autonomous, self-managing data warehouse solution with minimal administrative overhead.	Oracle offers 24/7 customer support with a wide range of training resources
SAP Data Warehouse Cloud	2020	Cloud-based, Integrated	AI-powered analytics, flexible storage	Libraries, Higher Education, Research	Libraries needing flexible, collaborative, and automated data workflows	Custom pricing	Seamless integration with SAP applications	Well-suited for libraries looking for deep integration with SAP for business intelligence and data management.	SAP provides premium support with dedicated teams and extensive training materials
Teradata Vantage	1983	Cloud-based, Scalable	End-to-end analytics, cloud-native solution	Libraries, E-commerce, Healthcare	Libraries seeking full data lifecycle management with scalability	Custom pricing	End-to-end analytics with cloud-native solutions	Perfect for libraries needing a full analytics solution with seamless integration, high scalability, and end-to-end data lifecycle management.	Teradata provides comprehensive support with a strong user community, technical assistance, and dedicated customer service

requirements. The cost is justified by its integrated nature and focus on library-specific workflows.

- **Use Case:** Designed specifically for integrated library systems (ILS) and resource management. Suitable for academic and research libraries seeking a comprehensive library management solution with built-in analytics.
- **Key Benefit:** Tailored workflows for library-centric tasks like cataloging and acquisitions.
- **Considerations:** Primarily focused on core

library management functions rather than general-purpose data analytics.

7.6. OCLC WorldShare Analytics (2013)

- **Cost-Benefit:** Custom pricing based on library size and usage ensures flexibility. The cost is justified by its focus on consolidating library data and providing advanced reporting tools.
- **Use Case:** Best for libraries already using OCLC products. It offers seamless integration with other

OCLC services and provides advanced reporting and data governance features.

- **Key Benefit:** Advanced reporting tools and seamless integration with OCLC services.
- **Considerations:** Primarily beneficial for libraries within the OCLC ecosystem.

7.7. IBM Db2 Warehouse (1983)

- **Cost-Benefit:** Custom pricing positions it for libraries seeking scalable, AI-driven solutions. The cost reflects its advanced AI capabilities and robust data management features.
- **Use Case:** Good for libraries seeking advanced AI capabilities and scalable data warehousing solutions. Its automated, self-driving management system reduces administrative overhead.
- **Key Benefit:** AI-powered analytics and scalable data management.
- **Considerations:** May be more complex to implement and manage compared to simpler cloud-based solutions.

7.8. Oracle Autonomous Data Warehouse (2018)

- **Cost-Benefit:** Custom pricing caters to libraries needing high automation with minimal administrative overhead. The cost is justified by its self-driving capabilities and reduced need for manual intervention.
- **Use Case:** Ideal for libraries with limited technical resources but requiring robust analytics. Its autonomous nature minimizes administrative burden.
- **Key Benefit:** Highly autonomous data management with minimal manual intervention.
- **Considerations:** May require specialized expertise for initial setup and configuration.

7.9. SAP Data Warehouse Cloud (2020)

- **Cost-Benefit:** Custom pricing supports libraries needing collaborative and automated data workflows. The cost is justified by its seamless integration with SAP applications and collaborative features.
- **Use Case:** Well-suited for libraries looking for deep integration with SAP for business intelligence and data management. Its collaborative capabilities enhance inter-departmental communication.
- **Key Benefit:** Seamless integration with SAP applications and collaborative features.
- **Considerations:** Primarily beneficial for organizations already using SAP products.

7.10. Teradata Vantage (1983)

- **Cost-Benefit:** Custom pricing makes it viable for libraries requiring comprehensive analytics solutions.
- **Use Case:** Perfect for libraries needing a full analytics solution with seamless integration, high scalability, and end-to-end data lifecycle management.
- **Key Benefit:** End-to-end analytics with cloud-native solutions.
- **Considerations:** May be more complex to implement and manage compared to simpler cloud-based solutions.

8. Conclusion

The optimal choice depends on the library's specific needs, budget, existing IT infrastructure, and technical expertise. Libraries should carefully evaluate their requirements and prioritize features based on their strategic goals. Smaller libraries with limited budgets and technical expertise might benefit from serverless solutions like Google BigQuery. Larger research libraries with complex data needs and existing AWS or Microsoft Azure investments might find Redshift or Synapse Analytics more suitable. Libraries requiring specialized library management functions should consider Alma or OCLC WorldShare Analytics. Finally, libraries prioritizing automation and minimal administrative overhead should consider Oracle Autonomous Data Warehouse.

Declaration:

“This is to declare that this article is developed by us, based on our personal study and research or, on primary/secondary published data and that We have duly acknowledged the use of all such previously published data in preparation of this report in a conventional manner and it is free from plagiarism. Further, We declare that this report has not previously been submitted for assessment / publication in any other journal.” – Subham Das

Authors Contribution

Dr. Swapan Khan and Subham Das both worked simultaneously in this article where the former one give the idea about the topic the later one executed the whole study.

Consent of authors: Consent of each author towards publication of the article is to be submitted separately along with the manuscript in the following format –

Declaration of Consent

“This is to declare that I have full consent in publishing the article entitled “EMPOWERING LIBRARIANS WITH DATA: A COMPARATIVE ANALYSIS OF DATA WAREHOUSING TOOLS” along with Dr. Swapan Khan as co-author. I do not have any conflict of interest in publication of the said article.” – Subham Das

“This is to declare that I have full consent in publishing the article entitled “EMPOWERING LIBRARIANS WITH DATA: A COMPARATIVE ANALYSIS OF DATA WAREHOUSING TOOLS” along with Subham Das as co-author. I do not have any conflict of interest in publication of the said article.” – Dr. Swapan Khan

References:

- 1 Chandra, P., & Gupta, M. K. (2018). Comprehensive survey on data warehousing research. *International Journal of Information Technology*, 10(2), 217–224. <https://doi.org/10.1007/s41870-017-0067-y>
- 2 March, S. T., & Hevner, A. R. (2007). Integrated decision support systems: A data warehousing perspective. *Decision Support Systems*, 43(3), 1031–1043. <https://doi.org/10.1016/j.dss.2005.05.029>
- 3 Redman, T. C. (2013). Data quality management past, present, and future: Towards a management system for data. In *Handbook of Data Quality: Research and Practice* (pp. 15–40). Berlin, Heidelberg: Springer Berlin Heidelberg.
- 4 Laitinen, M., & Saarti, J. (2012). A model for a library-management toolbox: Data warehousing as a tool for filtering and analyzing statistical information from multiple sources. *Library Management*, 33, 253–260. <https://doi.org/10.1108/01435121211242290>
- 5 Yang, Y.-T., & Shieh, J.-C. (2016). *Data Warehouse Applications in Libraries—The Development of Library Management Reports*. 88–91. <https://doi.org/10.1109/IIAI-AAI.2016.129>
- 6 L C, M., & R K, S. (2023). A Review on Data Warehousing Concepts, Challenges and Applications. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 25–31. <https://doi.org/10.32628/CSEIT239015>
- 7 Sureddy, M. R., & Yallamula, P. (2020). *Approach to help choose right data warehousing tool for an enterprise*. 6, 579–583.
- 8 Dhaouadi, A., Bousselmi, K., Gammoudi, M.M., Monnet, S., & Hammoudi, S. (2022). Data Warehousing Process Modeling from Classical Approaches to New Trends: Main Features and Comparisons. *Data*, 7, 113.
- 9 Cloud Data Warehouse – Amazon Redshift – Amazon Web Services. (n.d.). Amazon Web Services, Inc. Retrieved April 30, 2023, from <https://aws.amazon.com/redshift/>
- 10 BigQuery enterprise data warehouse | Google Cloud. (n.d.). Retrieved December 22, 2024, from https://cloud.google.com/bigquery/?utm_source=bing&utm_medium=cpc&utm_campaign=japac-IN-all-en-dr-bkws-all-all-trial-e-dr-1009882&utm_content=text-ad-none-none-DEV_c-CRE_-ADGP_Hybrid+%7C+BKWS+-+EXA+%7C+Txt+~+Data+Analytics_BigQuery_gcp+bigquery_main-KWID_43700079460740557-kwd-71950090695949:loc-90&userloc_116075-network_o&utm_term=KW_gogle%20big%20query&gclid=d53db1dd16751e376526d6ba0de44f21&gclsrc=3p.ds&
- 11 Azure Synapse Analytics | Microsoft Azure. (n.d.). Retrieved April 30, 2023, from <https://azure.microsoft.com/en-in/products/synapse-analytics>
- 12 Snowflake Cloud Data Platform | Snowflake AI Data Cloud. (n.d.). Retrieved December 22, 2024, from <https://www.snowflake.com/content/snowflake-site/global/en/data-cloud/platform>
- 13 IBM Db2 Warehouse. (n.d.). Retrieved December 22, 2024, from <https://www.ibm.com/products/db2-warehouse>
- 14 Automate Your Data Warehouse. (n.d.). Retrieved December 22, 2024, from <https://www.oracle.com/autonomous-database/autonomous-data-warehouse/>
- 15 SAP Data Warehouse Cloud—Try it out! - SAP Community. (n.d.). Retrieved December 22, 2024, from <https://community.sap.com/t5/technology-blogs-by-sap/sap-data-warehouse-cloud-try-it-out/ba-p/13518500>
- 16 Alma Overview—Ex Libris Knowledge Center. (n.d.). Retrieved December 22, 2024, from [https://knowledge.exlibrisgroup.com/Alma/Product_Documentation/010Alma_Online_Help_\(English\)/010Getting_Started/010Alma_Introduction/010Alma_Overview](https://knowledge.exlibrisgroup.com/Alma/Product_Documentation/010Alma_Online_Help_(English)/010Getting_Started/010Alma_Introduction/010Alma_Overview)
- 17 WorldShare: Enable shared efficiencies and innovation. (2021, November 3). OCLC. <https://www.oclc.org/en/worldshare.html>
- 18 What is Teradata Vantage? | Video. (2024, May 20). <https://www.teradata.com/resources/videos/what-is-vantage>