A Historical Review of Data Warehouse Implementation in Higher Education

Eris Zeqo

Edjola Naka

Faculty of Information Technology,
Aleksander Moisiu University,
Durres, Albania

Received: 20 September 2025 / Accepted: 15 November 2025 / Published: 26 November 2025

Doi: 10.56345/ijrdv12n328

Abstract

Academic institutions are entities that offer education, perform research, manage activities, and issue degrees. The majority of educational institutions have provided their services for an extended period; hence, historical data regarding students, professors, and administration must be studied. This publication attempts to evaluate the implementation of data management strategies, specifically data warehouses, in the higher education sector. A systematic literature review using the PRISMA framework conducts an evaluation of data warehouse utilization at universities, emphasizing the application domain, data sources, and architecture of data warehouse schemas, covering the period from 2000 to 2025. We have selected thirty-one manuscripts for this evaluation. The final results present data warehouse solutions mostly related to student evaluation, distant learning, and academic staff's achievement. The analysis closes with final recommendations and observations regarding the significance and advantages of employing data warehouses in higher education.

Keywords: *Data warehouse, higher education, review, schema, data sources*

1. Introduction

A data warehouse is a technology utilized to combine, store, and manage substantial volumes of data. It retrieves data from many sources, transforms it into a standardized format, and subsequently inserts it into a unified and structured system within the data repository. Upon collection, the data are stored for several years, facilitating access to past information. The primary objective of data warehouse technology is to facilitate high-performance data access and analytical capabilities to enhance the decision-making process. A data warehouse is a subject-oriented, integrated, nonvolatile, and time-variant repository of data that facilitates management decision-making (Inmon, 2002). There are two methodologies for creating data warehouses: the Inmon (Inmon, 2002) and the Kimball approaches (Kimball & Ross, 2013). The Inmon methodology is alternatively referred to as the top-down approach. It constructs the enterprise data warehouse initially, followed by the creation of data marts. Data warehouses are not constructed in a single process. Rather, they are constructed and filled incrementally. A conventional procedure for constructing a data warehouse is referred to as the "Day 1-day n" phenomenon. The alternative, bottom-up method commences by establishing small, targeted data marts centered on certain business processes or departments, rather than constructing a comprehensive enterprise-wide data warehouse first. The approach encompasses a four-step dimensional design methodology: selecting the business process to model, declaring granularity, choosing dimensions, and determining the measures that will populate each table.

A University is an institution of higher education that **issues** academic degrees across diverse disciplines. Universities are recognized for offering comprehensive theoretical knowledge, conducting academic research, and frequently incorporating a wide array of subjects or specializing in particular fields. Universities comprise various stakeholders, including governing bodies, executive leadership, academic divisions, and diverse administrative departments, all of which contribute to the institution's general functionality. Universities create substantial data, typically encompassing student applications, enrollment, and transfer; course registrations and academic achievement; faculty workload and research outputs; as well as financial aid, tuition, and university finance.

Research on the use of data warehouses in education has considerably increased in recent years, mostly focused on the design and implementation of data warehouses to facilitate decision-making. Applications of data warehouse technology in educational management include enrollment, student information, course resources, faculty and staff administration, and ultimately, teaching quality assessment (Guorui, 2023). The educational evaluation system utilizes data warehouses for data integration and management, multidimensional assessment, customized learning assistance, and enhancement of educational decision-making (Lv, 2025). Utilizing a multidimensional evaluation technique, an educational institution may thoroughly assess students' academic achievement and overall quality, thereby offering robust support for customized instruction. Conversely, university students and educators are encouraged to possess a fundamental level of digital proficiency. Moreover, higher education institutions are urged to concentrate on enhancing the digital competencies of students and educators, adopt suitable learning methodologies, and use suitable tools to elevate educational quality (Zhao, Llorente, & Gómez, 2021).

Significant projects are also being executed in higher education, as shown by the United Kingdom. A significant active data infrastructure initiative investigates the sociotechnical networks of organizations, software applications, standards, dashboards, and visual analytics technologies that form the infrastructure, as well as the integration of these technologies with governmental guidelines for education reform (Williamson, 2018). Each higher education institution must thoroughly assess every factor related to the storage, transformation, and secure loading of historical data into repositories for research purposes. As university data transitions to big data over the years, it is essential to assess and compare big data technologies to identify the most effective option. This study is beneficial for comprehending how big data might enhance a data warehouse in higher education, all the necessary architectural components, and the associated challenges (Santoso & Yulia, 2017). Typically, there exists a synergy between Hadoop and relational database management system (RDBMS), as the latter is prevalent in academic data. Hadoop can complement RDBMS as a data ingestion and staging tool, as well as a data management and visualization platform.

The design and implementation of a data warehouse for an online learning platform using three essential technologies—Microsoft SQL Server, MongoDB, and Apache Hive is described from (Rudniy, 2022). The design, construction, and evaluation of a DW were examined to identify improved ways of storing, organizing, and extracting information. It was found that Apache Hive demonstrated the best processing time, followed by SQL Server and MongoDB. In terms of analytical queries, MS SQL Server demonstrated the best processing time, followed by MongoDB and Hive. The purpose of this study is to establish a historical foundation for identifying existing data warehouse infrastructures, the fields in which these infrastructures are implemented, the sequential trends in research within this area, and the potential gaps when applied to higher education systems. The systematic literature review will concentrate on the primary questions outlined below, which will be assessed through the collected data:

RQ1. How much are used the data warehouses in storing and analyzing data in universities?

RQ2. Which areas or services of the universities are more employed by data warehouse designers?

RQ3. How effective have been their usage?

RQ4. Which are some future recommendations that can be given?

This analysis consists of the subsequent sections. Section 2 presents information regarding data warehouses, universities, their relationship, and the methodology employed. Each publication selected from the literature is detailed in section 3, organized chronologically according to three distinct time periods. Part 4 summarizes and discusses the data presented in the literature from section 3. The final findings and conclusions of the review are summarized in section 5.

2. Methods

2.1 An overview about data warehouses

A data warehouse integrates data from multiple sources into a centrally located database intended for querying and analysis. Data warehousing systems can process immense amounts of data from many source systems, including

operational databases, transactional systems, and customer relationship management (CRM) platforms. Schemas in a data warehouse indicate the organization of data. These schemas are dimensional data models intended to enhance data retrieval times in OLAP systems. Dimensional models can enhance redundancy, facilitating information retrieval for reporting and optimizing query efficiency . Data modeling typically follows either a star schema or a snowflake schema. These schemas contain fact tables and dimension tables. Fact tables are typically very large tables and contain information about specific occurrences in the business. Quantitative data is stored in the fact table's characteristics and can be aggregated. Dimension attributes are used to store descriptive information for fact tables. Schemas in data warehouses can be employed by data analysis tools, including OLAP tools, reporting tools, and data mining tools. A simple structure of a data warehouse is given in Figure 1.

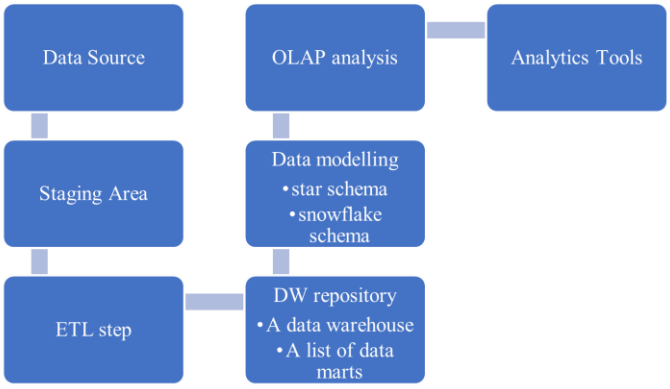

Figure 1. Basic structure of a data warehouse

The data flow and organization process typically commence with the extraction of relevant data from the source system. They are temporarily relocated to the staging area, where they are prepared for conversion in the Extract, Transform, and Load (ETL) process. Depending on the design methodology, a data warehouse or a collection of data marts is established. Star schemas or snowflake schemas should be developed. A data analysis environment is established, enabling the application of Online Analytical Processing (OLAP) and machine learning techniques for analytics.

A data warehouse can yield significant benefits when deployed effectively, with several advantages outlined as follows (Bhatia, 2019):

- ✓ Potential high Return on Investment
- ✓ Unbeatable competitive advantage
- ✓ High productivity in corporate decision making and business intelligence
- ✓ Cost effective
- ✓ Enhanced customer service

2.2 Higher education institutions, and relation with data warehouses

The global landscape has increasingly evolved into a competitive and dynamic environment for higher education institutions. In response to national and global economic, political, and social transformations, they must compete with other higher education institutions to improve their functionalities and to provide better education and research policies. One type of higher education institution is universities, which contribute to society by educating students in critical thinking and analytical skills. Figure 2 illustrates an example of a university's organizational structure. This framework offers a comprehensive overview of the principal responsibilities undertaken by a university and does not identify any specific institution. They are governed by an academic board responsible for supporting the study programs, students, academic personnel, and representing the university in collaborations with other institutions. Data are supplied concerning learning, research, and projects. The administrative board supervises services related to employment, legal affairs, financial management, information technology, and student administration. For each division, data are collected from various sources, requiring storage and analysis.

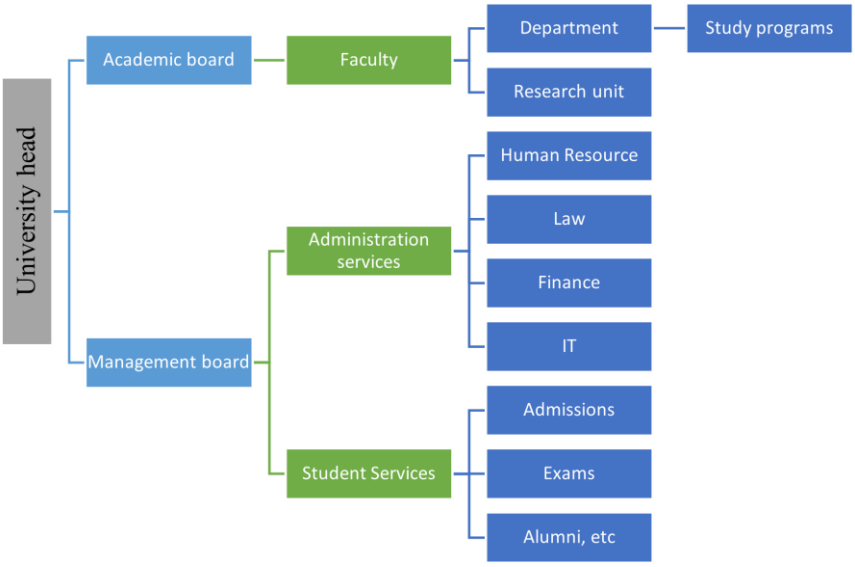

Figure 2. An illustration of an academic institution's organizational structure

Designing a centralized system that gathers, combines, and saves data from various sources to facilitate analytics, reporting, and decision-making is the process of building data warehouses for academic institutions. Data warehouses typically use a variety of input data sources, such as databases from alumni/CRM programs, learning management systems (LMS), human resource systems (HRIS), financial management systems, library systems, and student information systems (SIS). ETL tools can be used to transform data for the data warehouse layer, which is mostly a relational layer.

In a recent study, (Moscoso-Zea, Paredes-Gualtor, & LujáN-Mora, 2018) conducted an extensive review of empirical research publications on the subject of data warehouses in education from 2008 to 2018. In order to have a comprehensive understanding of data warehouses in education, a qualitative and quantitative methodology based on a four-stage research process was used for this work. Thirty-four appropriate articles were found and thoroughly examined after the suggested criteria were applied and excluded. For instance, Kimball's technique is the most often used methodology for data warehouse design in education, according to the study's noteworthy findings. Based on this data, they found that the most prevalent approach is the star schema.

Education institutions need data warehouses, and they should employ them to aid in the administration of the establishments (Yu, 2021). Some advantages of using data warehouse technology in education management are (Guorui, 2023):

- ✓ It can perform data integration and centralized management;
- ✓ It can provide data decision support
- ✓ It can support accurate forecasting and planning.

When data warehouses are used in education evaluation systems, the accuracy and thoroughness of assessments are greatly increased, and educators, students, and management have a more objective and scientific foundation for decision-making (Lv, 2025).

3. Methodology

To guarantee that our systematic review remains focused, transparent, and reproducible, we established clear rules for which studies to include or exclude using PRISMA 2020 framework recommendations. The PICOT framework is applied:

- ✓ **P (Population):** Higher education institutions (universities, colleges, etc.)
- ✓ **I (Intervention):** Implementation of data warehouse systems
- ✓ **C (Comparison):** No data warehouse system or traditional data management systems (e.g., spreadsheets,

flat databases).

- ✓ **O (Outcome):** Improvements in data management, decision-making processes, efficiency, or student success outcomes.
- ✓ **T (Time):** Within the last 25 years

Information about eligibility criteria, sources, search strategy, data analysis, and PRISMA flow diagram, are given in below paragraphs.

3.1 Eligibility criteria's

Some inclusion criteria are:

- ✓ Universities using data warehouses.
- ✓ Studies discussing the use, implementation, or outcomes of data warehouse systems in higher education.
- ✓ Measurable impacts on operational efficiency, data accuracy, decision-making processes, or academic performance indicators.
- ✓ Open access conference proceedings, journals, reports, and specific universities case studies.
- ✓ Literature in English language.

Some exclusion criteria are:

- ✓ Literatures that focus on applying data warehouses on industries outside higher education.
- ✓ Studies not assessing the actual outcomes of using data warehouses.
- ✓ Non-English language manuscripts.
- ✓ Review papers, books, or dissertation thesis.

3.2 Data sources and search strategy

On July 25, 2025, a search was conducted exclusively using the Google Scholar database, as it provides an extensive resource related to the relevant topic. The search strings were: ("data warehouse") AND ("university"). The search was limited to papers published within the last 25 years, starting with the year 2000.

In the identification phase, database searches in Google Scholar yielded around 240,000 records. Before screening, only records published between 2000 and 2025 were removed. No duplicates or ineligible records were found during the screening. A total of 18,000 records were identified related to the search topic. Since the articles are sorted by relevance, a manual selection of the first 300 records was included for the reviewing process. For the 300 records, inclusion and exclusion criteria were applied, screening them from the title and abstract. The full text of the remaining 31 articles was then retrieved and assessed for eligibility. Figure 3 illustrates the literature screening process using the PRISMA 2020 diagram.

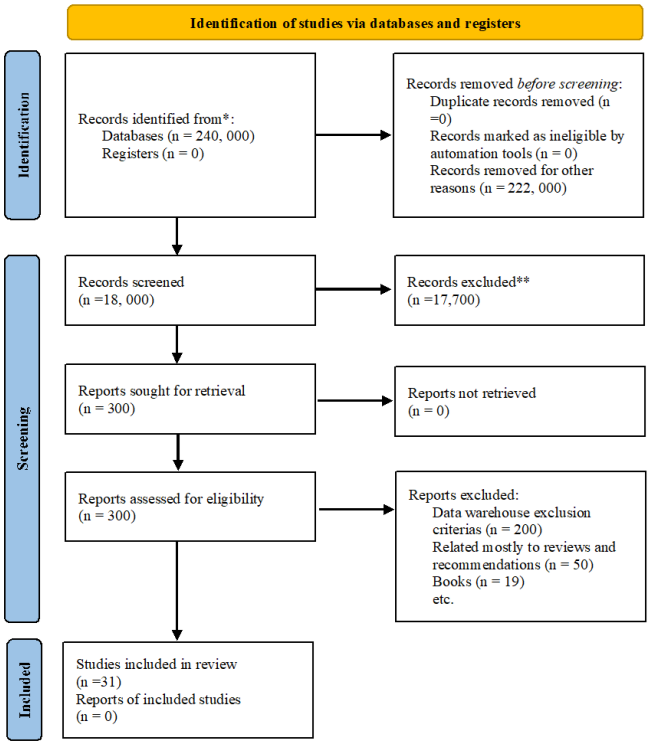

Figure 3. The PRISMA flow diagram for the literature screening.

3.3 Data extracting, and analysis

A total of 31 manuscripts were selected. All of them were carefully investigated by abstract, applied methods, implemented data warehouse schemas, and main conclusions. We worked independently to pull out all the important details in six parts:

- ✓ The name of the university.
- ✓ The domain of higher education is where the data warehouse is applied.
- ✓ Structure of the data warehouse.
- ✓ Source systems, if provided.
- ✓ Software used, mostly related with the implementation of the data warehouse.
- ✓ Short conclusions regarding their effectiveness.

The following section will extract and summarize the key findings from each selected literature source.

4. Results

Generally, the examination and application of data warehousing in educational systems are infrequently investigated. This section summarizes the principal findings from each selected publication. This is a chronological review split into three-year intervals: 2000-2010, 2011-2020, and 2021-2025. The reason for beginning the analysis in the year 2000 is to review a time when information technology related to data warehouses had not yet reached its current level of advancement and also to enable a longer period of analysis. Each paragraph encompasses the information for each time period.

4.1 A literature analysis covering the years 2000 to 2010

Earlier papers initiated the design and analysis of data warehouse utilization. Their information is captured in Table 1.

Table 1. A summary of the selected papers for the years 2000-2010

No.	University	Field	DW structure	Source systems	Software	Conclusion	Reference
1	Polytechnic University, member of National Science Foundation Gateway Coalition project.	Student outcomes	Not provided.	Data from legacy/departmental systems, external reports, and student information systems were collected, cleaned, and integrated into the warehouse over a ten-year period.	Apache Webserver, Perl 5m and MySQL.	A data warehousing system was proposed to consolidate institutional data to support the assessment process.	
2	University of Latvia, Riga, Latvia.	E-learning	Three-star schemas: Structure, Usage, and Activity	WebCT logs, Management Information system	Relational schema, but not information about a specific one	A viable and valuable way to assess how e-learning influences university processes.	
3.	University of Pretoria, South Africa	E-learning	Student data mart, click-stream data mart	Learning management system	Not provided	LMS tracking + effectiveness metrics support continuous improvement in teaching.	
4.	Aristotle university of Thessaloniki, Greece	Grade evaluation / graduation procedure	Two-star schemas	Department historical data	Not provided	Allows the longitudinal analysis of student performance within the department and the investigation of factors that may influence their performance.	
5.	University of Novi Sad, Serbia	Course evaluation	Star schemas; Two fact tables, seven dimensions, two materialized views	Course data: assignments, attendance, grades etc. in one semester	Not provided	Student progress, duties, attendance evaluation. With this data warehouse, the authors applied data mining techniques in another article.	
6.	University of Rijeka, Croatia	E-learning	Star schema, fact_User_Activity table, related with 5 dimensions.	Learning Management System Moodle-based named Mudri	Microsoft SQL Server 2008, MySQL.	Reports on system workload distribution, usage of e-learning tools, and change in user activity by years are created.	

4.2 A literature analysis covering the years 2011 to 2020

This subsection proceeds with the primary information for the years 2011 to 2020, summarized in Table 2.

Table 2. A summary of the selected papers for the years 2011-2020

No.	University	Field	DW structure	Source systems	Software	Conclusions	Ref
1.	Not specified.	Students' evaluation.	Snowflake schema with TranscriptFactTable, and dimensions: achievements, student, registered activities, instructor, alumni, account, and receipt.	Fourteen relational tables related to: registration, accounting, course management, alumni, and asset systems.	Ms Access 2010.	A practical small-scale implementation of operational databases on a structure optimized for decision support.	(Bassil, 2012)
2.	R V College of Engineering, Bangalore, Karnataka, India.	Employees, student, and assets information.	Not provided.	Smart campus, asset management server, .csv files	Not provided, just open-source ones.	Decision makers can extract information regarding employees, students and infrastructure.	(Manjunath, Ravindra, Umesh, & Ravikumar, 2012)
3.	A consortium of universities in Iraqi.	E-governance.	Not provided.	Information, documents and reports have collect from universities' data sources.	Not provided.	To increase interaction among universities in higher education in Iraq. Similar approach for Malaysia	(Mohammed, Hasson, Shawkat, & Al-khafaji, 2012)
4.	University of Bari, Italy.	Analysis of scientific publications and the students'	Hybrid top-down, and bottom-up design; Research and Didactics data marts.	Six source systems related with students, financial, legal, scientific research.	Not provided.	Business applications are developed and deployed using a web-based OLAP	(Di Tria, Lefons, & Tangorra, 2015)

No.	University	Field	DW structure	Source systems	Software	Conclusions	Ref
		performance.				layer. In a previous work, the data warehouse was presented .	
5.	International Ataturk Aalto University, Kyrgyzstan	Students' evaluation.	Snowflake schema, study_indicator fact table, and 4 dimensions.	Fifty-nine relational tables.	MySQL, Apache Hive, Apache Sqoop.	A university decision support system.	(Bondarev, Zakirov, & Zakirov, 2015)
6.	Public University in Kenya, Africa.	Fee payment.	Star schema, named, fee payment with dimensions: students, courses, employees, and time.	Student Information System, external files and Spreadsheets.	Excel, MS SQL Server, MS Access.	The results showed an increase in revenue over the years.	(Moturi & Emurugat, 2015)
7.	STMIK Pelita Nusantara university, Indonesia.	Accreditation.	Kimball approach, three-star schemas, for registration, studying, and lecturer.	Data from the management System of the university.	Pentaho Data Integration / Kettle.	Designed only for fulfil the standard3 (students and alumni) and standard4 (Human resource).	(Sinaga & Girsang, 2016)
8.	Faculty of science and technology UIN Jakarta, Indonesia.	Lecturer performance.	Star schema.	Lecturer attendance, courses taught, student assessment.	SQLyog and Pentaho Data Integration/Spoon.	Management can take decisions that will improve the performance and competence of the lecturers.	(Kumaladewi & Sugiarti, 2016)
9.	University of Basra, Iraq, Al-Iraq University	Course management.	Star schema.	Data related to students, courses and lecturers.	Ms SQL Server, SQL Server Data Tools, SQL Server Reporting Services, Excel.	The approach might be implemented by the lecturers and heads of the departments and decision makers.	(Abdullah & Obaid, 2016)
10.	Campus XYZ	Borrowing process in university library.	Snowflake schema.	Library transactional data from year 2003 until January 2013.	Not provided.	Best seller books, authors of best sellers, busiest days, etc	(Dahlan & Wibowo, 2017)
11.	Not provided	Accreditation.	Nine-star schemas.	Student and Graduate Profile, Student Service Profile, Graduation Evaluation Profile and Alumni Association.	Not provided.	Helps administrators, and executives to quickly access accreditation data.	(Budiarta, Wijaya, & Indrapartha, 2017)
12.	University Utara Malaysia	E-learning.	Star schema, with four dimensions: staff, school, course and session.	Personnel Information System, Graduate Affairs Information System, Academic Student Information System, and UUM Learning System.	MySQL, Excel, Talend Open Studio, PhP, R programming.	The usage of blended learning can be monitored successfully by lecturers and university management.	(Ta'a, Bakar, Chit, & Soid, 2017)
13.	Not provided	Analyzing active students.	Kimball's approach, star schemas, Registration Fact, Assessment Fact, Leave Request Fact, Withdrawal Fact, and Graduation Fact.	Multiple database sources, not specified.	Pentaho Data Integration.	Produce analytic data and dashboard for decision-making.	(Sutedja, Yudha, Khotimah, & Vasthi, 2018)
14.	Aalto University, Finland.	Graduation probability.	Not provided.	Student Information System, Current Research Information System, Financial Information Systems, Human Resources Information System and User Account Database.	Not provided.	They could successfully produce predictions of graduation probability and time-to-degree.	(Pesonen, Fomkin, & Jokipii, 2018)
15.	Mindanao State University-Iligan Institute of Technology, Philippines.	Financial analysis.	Kimball's approach, Star schema, fund amounts fact table with dimensions: finance unit, account, time, fund dimensions.	Financial Management System.	PostgreSQL, Pentaho Data Integration (Kettle).	A REST API was created to allow external applications to interact with the warehouse.	(Lapura, Fernandez, Pagatpat, & Dinawanao, 2018)

4.3 A literature analysis covering the years 2021 to 2025

This last subsection continues with presenting the main information for the next following years in Table 3.

Table 3. A summary of the selected papers for the years 2021-2025

No.	University	Field	DW structure	Source systems	Software	Conclusions	Ref
1.	Wasit University, Iraq.	Finance, research, human resource, and students' data marts.	An architectural way, not schema-level detail presented.	University departments, structured, and unstructured data.	Hadoop, and relational database.	To introduce a data warehouse in order to support decision-making.	(Neamah, 2021)
2.	A private higher education institute.	Student registration.	Kimball approach, star schema, fact_student_registration, and dimensions: scholarship, student, batch, course, Date.	Inquiry Handling System, Human Resource Information System and Management Information System, .csv files, .xls files, and manual records.	Microsoft SQL Server.	The data warehouse prototype has shown usefulness in bringing together data, enabling historical trends, and better report generation.	(Serasinghe, Jayakody, Dayananda, & Asanka, 2021)
3.	University of Novi Sad, Serbia.	Students' accomplishment of course requirements on traditional and distance learning.	Bottom-up approach, star schema, Fact table: completed assignment, and attendance, and dimensions: teaching staff, student, course, date.	Student exams, teaching data, curriculum data etc.	Microsoft SQL Server.	Creation of a model that can guide students in their choice of university.	(Gladić & Petrovački, 2021)
4.	Private National university, Jakarta, Indonesia	Student performance	Three-star schemas: new students, students value, and student academic notes.	Academic Information System, Student Academic Services, New Student Registration, Visual Scheduling System.	PostgreSQL 8.1.	To assess academic activities, from registration through graduation.	(Edastama, Dudhat, & Maulani, 2021)
5.	A private university, not clearly provided.	Accreditation procedure.	Star schema, Kimball's approach, not schema design provided.	Twenty-four relational tables.	Not provided.	A prototype for measuring the performance evaluation for academic programs.	(Sihombing, 2022)
6.	Romblon State University, Philippines.	Accreditation evaluation.	Not specified	Internal data, not specified.	PhpMyAdmin, MySQL.	A data warehouse designed to store all documents required for accreditation.	(Mindoro-Mesana, Mesana, & Mariño, 2022)
7.	Six major universities in South Kivu Province, DR Congo	Student academic career.	Kimball's approach, a star schema developed.	IT information systems form some universities.	Pentaho Data Integration.	Creation of a model that can guide South Kivu students in their choice of university.	(Philippe, et al., 2023)
8.	Anadolu University, Turkey.	E-learning.	Star schema, many data marts are created, but are not specified.	Learning Management System, Student Information System, Material management system, flat file sources.	MS SQL Server, .NET environment, C#.	Various analytical reports about students' study styles, preferred e-learning resources, and the time spent in the system.	(Güneş & Birgin, 2023)
9.	Gombe State University, Nigeria	Prediction system for staff promotion.	Entity-relational schema mentioned.	A two hundred (200) curriculum vitae of academic staff was selected with options qualified for promotion or not.	Oracle Data Warehouse Builder.	The system offers high degree of accuracy in predicting the case of promotion, staff-student's ratio as well as the budget projection.	(Ahmed, Muhammad, Adamu, & Umar, 2023)
10.	South East European University, North Macedonia	Identification of students at-risk of academic failure.	Star schema, Fact_StudentPerformance, and dimensions: Student, Course, Teacher, Time, Financial and Attendance,	University Management System.	Not provided.	They have identified various predictive models from literature.	(Ismaili & Besimi, 2024)

5. Discussions

This section analyzes the findings from the reviewed literature. It is noted that the majority of the selected manuscripts originated in the second time period, 2011–2020, totaling 15 instances, followed by 10 instances for the years 2021–2025 and 6 instances for the years 2000–2010. This study confirms the expansion of data warehouse applications over the

years. A discussion will be provided on each topic, including institution, higher education field, data warehouse schema design, source systems, software utilized, and comments and conclusions regarding effectiveness.

Regarding the university, the majority of the authors who conceived and designed the data warehouse focus on their specific institution of employment and concentrate on the individual institution where they were employed or studied. In some instances, the authors did not specify the university for which they developed the data warehouse, despite their research benefiting their own university. A consortium of universities was proposed in two instances to provide solutions for students' university selections for country assessments.

The second subject relates to identifying the most often designed scenarios in universities. Figure 4 indicates the distribution of university areas utilizing the data warehouse.

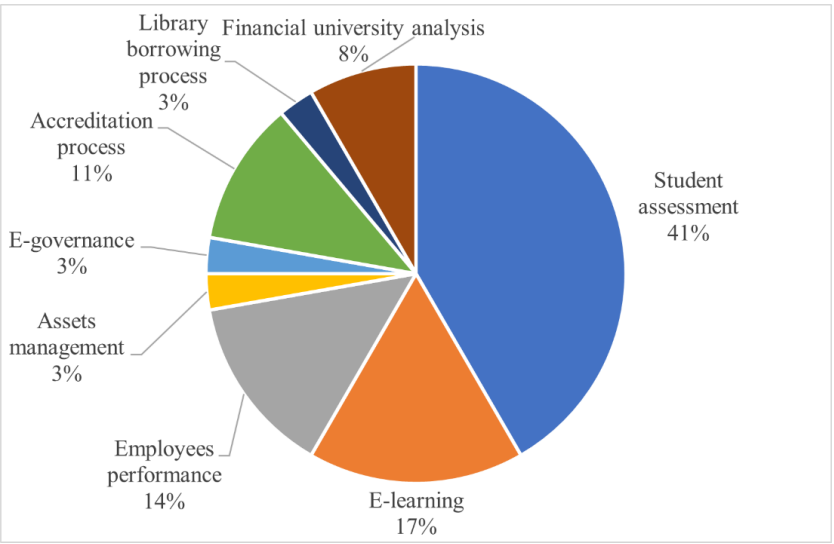

Figure 4. Percentage distribution of university fields based on thirty-one selected sources.

Measuring student performance in overall, e-learning process, evaluation staff performance, university accreditation, and university financial analysis are the most designed and analysis field on the universities. Less studies address processes as library borrowing, e-governance, and assets management.

The design of the data warehouse often employs star schemas and Kimball's methodology. A schema is not always carefully developed or accurately described; nonetheless, when it is specified, it is documented. In certain cases, the big data infrastructure was utilized, such as

Apache in conjunction with relational database systems.

Concerning the software utilized, there was a lack of information pertaining to it associated with the data source systems and the data warehouse software employed. Pentaho Data Integration was predominantly utilized for the ETL process. As the years progressed, the technology of the utilized program was enhanced. Given that the majority of data warehouses are

relational, Microsoft SQL Server is frequently cited as a prominent tool associated with various

types of databases. The integration of big data infrastructure with relational databases has received considerable attention.

In conclusion, each presented data warehouse prototype offers solutions that enhance decision-making for various stakeholders in certain domains. The benefits of each data warehouse system to various university processes are enumerated for each literature source. The contribution in each applied subject is a crucial step for enhanced analysis of higher education data and for delivering improved solutions.

This review can give some answers on the following research questions:

RQ1. How much are used the data warehouses in storing and analyzing data in universities?

Existing research indicates that the evolution of data warehouses affects their utilization in universities; however,

there is a limitation of studies in this area. Typically, multiple educational institutions do not openly reveal their data warehouse infrastructure, leading to the assumption that they have developed a proprietary system. Our effort focusses on identifying and analyzing research papers that have publicly implemented data warehouses, highlighting the methodologies employed in their research. Potential gaps that remain largely unexamined in higher education include the gathering, identification, and analysis of alumni data, as well as university project management.

RQ2. Which areas or services of the universities are more employed by data warehouse designers?

The chosen literature and Figure 4 indicate that issues involving students, such as graduation, course registration, management, and assessment of their academic accomplishments, are utilized more frequently, comprising 41% of the selected research. The significance of delivering an efficient educational process is centered on two primary stakeholders: students and academic staff. Accurate evaluation of the e-learning process and academic staff performance is crucial for improving the education process, and this evaluation is therefore highly regarded.

RQ3. How effective have been their usage?

The consulted literature suggests that implementing data warehouses only improves the overall university management, and offer solutions for governance, and other actors included. Many institutions in education haven't adopted data warehouses or data mining fully, partially because universities are nonprofit and resource-limited, cost and complexity remain barriers.

RQ4. Which are some future recommendations that can be given?

A few recommendations are:

- ✓ The universities have to improve the university information systems to design the data warehouse quickly and more correctly.
- ✓ Open-source software can be considered more often for reducing the costs of the institutions; even security issues should be considerably evaluated.
- ✓ The overall data warehouse architecture and implementation can be explained in more detail since this is an applicative process that is helpful for other researchers.

6. Conclusions

The aim of this systematic review was to summarize and analyze how the data warehouse systems are utilized in higher education. The analysis was based on several factors, such as university field, data warehouse modelling, data sources, specific software used, and the benefits provided from their implementation. The analysis is historical, grouping the selected literature according to time periods. Creating data warehouse solutions for universities enables clearer predictions of students' performance on exams, their learning processes, course management, and the lecturers' contributions to teaching and research. It also provides solutions for the management, such as project management and financial analysis. This review also identifies that most of the analyzed cases relate to students' management, and the preferred modelling case is the star schema and bottom-up approach. The selected data sources for the higher education institutions came from internal management systems, but also spreadsheets, and manual data were considered. The choice and explanation of data warehouse software were not always illustrated. This literature review makes a contribution to the higher education system to understand how effective and helpful decision-making is at universities using data warehouse technology.

References

Abdullah, Z. A., & Obaid, T. A. (2016). Design and Implementation of Educational Data Warehouse Using OLAP. International Journal of Computer Science and Network, 5(5), 824-827.

Ahmed, M. K., Muhammad, A. B., Adamu, A., & Umar, A. Y. (2023). Building a Secured Data Warehouse for a University Staff Management System: A Case Study of Gombe State University, Gombe. American Journal of Mathematical and Computer Modelling, 7(4), 55-60. doi:10.11648/j.ajmcm.20220704.12

Bassil, Y. (2012). A Data Warehouse Design for A Typical University Information System. Journal of Computer Science & Research (JCSCR), 1(6), 12-17.

Bhatia, P. (2019). Data Mining and Data Warehousing - Principles and Practical Techniques. Cambridge University Press.

Bondarev, A., Zakirov, D., & Zakirov, D. (2015). Data warehouse on Hadoop platform for decision support systems in education. 2015 Twelve International Conference on Electronics Computer and Computation (ICECCO) (pp. 1-4). Almaty, Kazakhstan: IEEE. doi:10.1109/ICECCO.2015.7416884

Budiarta, K., Wijaya, P. A., & Indrapartha, C. G. (2017). Analysis and Design of Data Warehouse on Academic STMIK STIKOM Bali. International Journal of Engineering and Emerging Technology, 2(1).

Čeliković, M., Aleksić, S., & Luković, I. (2009). A data warehouse system for monitoring university education process. Proceedings of the 4th International Conference on Information Technology (ICIT 2009), (pp. 1-7). Al-Zaytoonah University of Jordan.

Dahlan, A., & Wibowo, F. W. (2017). Design of Library Data Warehouse Using SnowFlake Scheme Method, Case Study : Library Database of Campus XYZ. 2016 7th International Conference on Intelligent Systems, Modelling and Simulation (pp. 318-322). Bangkok, Thailand: IEEE. doi:10.1109/ISMS.2016.71

Dell'Aquila, C., Di Tria, F., Lefons, E., & Tangorra, F. (2007). An Academic Data Warehouse. Proceedings of the 7th WSEAS International Conference on Applied Informatics and Communications, (pp. 229-235). Athens, Greece.

Di Tria, F., Lefons, E., & Tangorra, F. (2015). Academic Data Warehouse Design Using a Hybrid Methodology. Computer Science and Information Systems, 12(1), 135-160. doi:10.2298/CSIS140325087D

Dimokas, N., Mittas, N., Nanopoulos, A., & Angelis, L. (2008). A Prototype System for Educational Data Warehousing and Mining. 2008 Panhellenic Conference on Informatics,, (pp. 199-203). Samos, Greece.

Dyk, L. v. (2008). A Data Warehouse Model for Micro-Level Decision Making in Higher Education. The Electronic Journal of e-Learning, 6(3), 235-244. Retrieved from https://academic-publishing.org/index.php/ejel/article/view/1549

Edastama, P., Dudhat, A., & Maulani, G. (2021). Use of Data Warehouse and Data Mining for Academic data. A case study at a national university. International Journal of Cyber and IT Service Management(IJCITSM), 1(2), 206-215. Retrieved from https://doi.org/10.34306/ijcitsm.v1i2.55

Gladić, D., & Petrovački, J. (2021). Using a Data Warehouse System to Monitor and Analyze Student Achievement in Teaching Process : Student paper. 2021 20th International Symposium INFOTEH-JAHORINA (INFOTEH) (pp. 1-6). East Sarajevo, Bosnia and Herzegovina: IEEE. doi:10.1109/INFOTEH51037.2021.9400685

Güneş, İ., & Birgin, M. K. (2023). Implementing data warehouse infrastructure for an e-learning system. Gümüşhane University Journal of Science, 13(3), 750-766. Retrieved from https://doi.org/10.17714/gumusfenbil.1239468

Guorui, L. (2023). Practical analysis of data warehouse technology in college education management. Academic Journal of Computing & Information Science, 6(12), 132-137. doi:10.25236/AJCIS.2023.061215

Ingham, J. (2000). Data Warehousing: A Tool for the Outcomes Assessment Process. IEEE Transactions on Education, 43(2), 132-136.

Inmon, W. H. (2002). Building the data warehouse, 3rd edition. Wiley Computer Publishing.

Ismaili, B., & Besimi, A. (2024). A Data Warehousing Framework for Predictive Analytics in Higher Education: A Focus on Student at-Risk Identification. SEEU Review, 19(2), 43-57. doi:https://doi.org/10.2478/seeur-2024-0020.

Ivančević, V., Čeliković, M., Aleksić, S., & Luković, I. (2011). An Application of Educational Data Mining Techniques at Faculty of Technical Sciences in Novi Sad. ICIT 2011 The 5th International Conference on Information Technology.

Jonker, A., Holdsworth, J., & Kosinski, M. (2025, November 11). IBM Think. Retrieved from What is a data warehouse?: https://www.ibm.com/think/topics/data-warehouse

Kimball, R., & Ross, M. (2013). The data warehouse Toolkit,: the complete guide to diemnsional modeling. Wiley.

Kumaladewi, N., & Sugiarti, Y. (2016). Design analysis of data warehouse for lecturer performance evaluation (Case study: Faculty of science and technology UIN Jakarta). 2016 4th International Conference on Cyber and IT Service Management, (pp. 1-6). Bandung, Indonesia. doi:10.1109/CITSM.2016.7577531

Lapura, E. F., Fernandez, J. K., Pagatpat, M. J., & Dinawanao, D. D. (2018). Development of a University Financial Data Warehouse and its Visualization Tool. Procedia Computer Science, 135, 587-595. doi:10.1016/j.procs.2018.08.229

Lv, Y. (2025). Research on the Application of Data Warehouse in Education Evaluation System. Research on the Application of Data Warehouse in Education Evaluation System, 19(1). Retrieved from https://doi.org/10.4018/IJCINI.377804

Manjunath, T. N., Ravindra, H. S., Umesh, I. M., & Ravikumar, G. K. (2012). Realistic Analysis of Data Warehousing and Data Mining Application in Education Domain. International Journal of Machine Learning and Computing, 2(4), 419-422. doi:10.7763/IJMLC.2012.V2.157

Mindoro-Mesana, J. M., Mesana, J. P., & Mariño, L. M. (2022). Design and Development of Romblon State University Romblon Campus Accreditation Data Warehouse. Romblon State University Research Journal, 4(2), 17-24.

Mohammed, M. A., Hasson, A. R., Shawkat, A. R., & Al-khafaji, N. J. (2012). E-government architecture uses data warehouse techniques to increase information sharing in Iraqi universities. 2012 IEEE Symposium on E-Learning, E-Management and E-Services (pp. 1-5). Kuala Lumpur, Malaysia: IEEE. doi:10.1109/IS3e.2012.6414947

Mohammed, M., Ibrahim, H. B., Shawkat, A. R., & Hasson, A. R. (2013). Implementation of Data warehouse Architecture for E-government of Malaysian Public Universities to Increase Information sharing between them. Rural ICT Development (RICTD) International Conference 2013, (pp. 1-9). Malacca, Malaysia.

Moscoso-Zea, O., Paredes-Gualtor, J., & LujáN-Mora, S. (2018). A Holistic View of Data Warehousing in Education. IEEE Access, 6, 64659-64673. doi:10.1109/ACCESS.2018.2876753

Moturi, C., & Emurugat, A. (2015). Prototyping an Academic Data Warehouse: Case for a Public University in Kenya. British Journal of Applied Science & Technology, 8(6), 550-557. doi:10.9734/BJAST/2015/14774

Neamah, A. F. (2021). Adoption of Data Warehouse in University Management: Wasit University Case Study. nternational Conference on Applied and Practical Sciences ICAPS (2021), (pp. 1-12). doi:10.1088/1742-6596/1860/1/012027

Nebic, Z., & Mahnic, V. (2010). Data warehouse for an e-learning platform. 14th WSEAS International Conference on Computers, Part of the 14th WSEAS CSCC Multiconference;, (pp. 415-420). Corfu Island; Greece.

Ningning, G. (2010). Proposing Data Warehouse and Data Mining in Teaching Management Research. 2010 International Forum on Information Technology and Applications (pp. 436-439). Kunming, China: IEEE. doi:10.1109/IFITA.2010.286

Pesonen, J., Fomkin, A., & Jokipii, L. (2018). Building Data Science Capabilities into University Data Warehouse to Predict Graduation. European Journal of Higher Education IT(1), 1-11. Retrieved from https://www.eunis.org/erai/2018-1/

Philippe, B. K., Nsabimana, T., Ndayisaba, L., Kala, J. R., Ndikumagenge, J., & Mushengezi, E. Z. (2023). Decision-Making Information System for Academic Careers in Congolese Universities: From Analysis to Design of a Data Warehouse. Open Journal of Applied Sciences, 13, 2395-2407. doi:10.4236/ojapps.2023.1312187

Rudniy, A. (2022). Data Warehouse Design for Big Data in Academia. Computers, Materials & Continua, 71(1), 979-922. Retrieved from https://doi.org/10.32604/cmc.2022.016676

Santoso, L. W., & Yulia. (2017). Data Warehouse with Big Data Technology for Higher Education. 4th Information Systems International Conference 2017. Bali, Indonesia. Retrieved from https://doi.org/10.1016/j.procs.2017.12.134

Serasinghe, C., Jayakody, D., Dayananda, K., & Asanka, D. (2021). Design and Implementation of Data Warehouse for a Higher Educational Institute in Sri Lanka. 2021 6th International Conference for Convergence in Technology (I2CT) (pp. 1-5). Maharashtra, India: IEEE. doi:10.1109/I2CT51068.2021.9417820

Sihombing, D. (2022). Academic Data Warehouse Modeling in Higher Education Using Nine-Step Design Methodology. Journal of Information Systems and Informatics, 4(4), 1126-1134.

Sinaga, A. S., & Girsang, A. S. (2016). University Accreditation using Data Warehouse. Journal of Physics: Conference Series, 801, 1-8. doi:10.1088/1742-6596/801/1/012030

Solodovņikova, D., & Niedrīte, L. (2005). Using Data Warehouse Resources for Assessment of E-Learning Influence on University Processes. ADBIS 2005, Advances in Databases and Information Systems, Communications of the Ninth East-European Conference on Advances in Databases and Information Systems. Tallinn, Estonia.

Sutedja, I., Yudha, P., Khotimah, N., & Vasthi, C. (2018). Building A Data Warehouse to support Active Management: Analysis and Design. 2018 International Conference on Information Management and Technology (ICIMTech), (pp. 460-465). Bina Nusantara University, Jakarta, Indonesia.

Ta'a, A., Bakar, M. S., Chit, S. C., & Soid, M. H. (2017). Data Warehouse system for blended learning in institutions of higher education. e-Academia Journal, 6(2), 144-155.

Vardarlıer, P., & Silahtaroğlu, G. (2016). Journal of Research in Business & Social Science, 5(1), 1-14.

Wai, T. T., & Aung, S. S. (2009). Metadata Based Student Data Extraction from Universities Data Warehouse. 2009 International Conference on Signal Processing Systems (pp. 670-673). Singapore: IEEE. doi:10.1109/ICSPS.2009.101

Williamson, B. (2018). The hidden architecture of higher education: building a big data infrastructure for the "smarter university". International Journal of Educational Technology in Higher Education, 15, 1-26. Retrieved from https://doi.org/10.1186/s41239-018-0094-1

Yu, X. (2021). The Application of Data Warehouse in Teaching Management in Colleges and Universities. Journal of Physics: Conference Series, 1738, 1-8. doi:10.1088/1742-6596/1738/1/012090

Yulianto, A. A., & Kasahara, Y. (2020). Data Warehouse System for Multidimensional Analysis of Tuition Fee Level in Higher Education Institutions in Indonesia. International Journal of Advanced Computer Science and Applications, 11(6), 541-550.

Zhao, Y., Llorente, A. P., & Gómez, M. S. (2021). Digital competence in higher education research: A systematic literature review. Computers & Education, 168. Retrieved from https://doi.org/10.1016/j.compedu.2021.104212