ANNOTATION

dissertation work of Oksana Leonidovna Kopnova

on the topic " Methods and algorithms for data integration based on ontologies to ensure analytical activities in a university ", submitted for the degree of Doctor of Philosophy (PhD) in the specialty 8D06101 - "Informatics, computer engineering and management"

Modern universities are complex organizational structures that include many processes and departments, the management of which requires accurate and timely information. In the context of constantly growing volumes of data and increasingly complex requirements for their processing, universities face the problem of disparate data stored in various information systems. This creates obstacles to the formation of a comprehensive view of the work of an educational institution and complicates the monitoring and management of key processes, such as resource planning, educational activities, scientific research and administrative management.

Amendments to the legislation of the Republic of Kazakhstan regarding the expansion of academic and managerial independence of universities, which were approved on April 25, 2018 by the Mazhilis of the Parliament of the Republic of Kazakhstan within the framework of the draft law "On Amendments and Additions to Certain Legislative Acts of the Republic of Kazakhstan on the Issues of Expanding the Academic and Managerial Independence of Higher Education Institutions", which amends three main laws regulating the activities of educational institutions. This bill amends three main laws regulating the activities of educational institutions: the Law "On Education"; the Law "On Science"; the Law "On the Commercialization of the Results of Scientific and (or) Scientific and Technical Activities". According to these amendments, universities are granted greater academic freedom from 65% to 85%, which allows them to independently develop educational programs focused on the needs of the labor market in a particular region, as well as form a student body by specialty. In addition, universities are granted financial independence and the opportunity to conduct educational activities in other countries. These changes create additional requirements for the organization of analytical activities of the university and its management issues. Information systems cannot fully satisfy the needs of decision makers, since information is presented either by means of outdated information systems, disparate databases, poor-quality data, or in conditions of lack of standards and consistency, etc.

The solution to these issues can be an information and analytical system (IAS) that allows for the unification and integration of data from various systems, ensuring transparency and integrity of information for analytical activities and informed decision-making at the university.

The dissertation proposes an approach based on ontologies that can integrate both structured and unstructured data and provide a holistic view of corporate information systems data, improving the decision-making process. The current state shows that in the context of rapid data growth and the complexity of their analysis,

the use of ontologies is becoming not just an advantage, but a necessity for organizations seeking to maintain competitiveness and make strategic decisions based on comprehensive and high-quality analytics. However, there are a number of scientific issues that do not sufficiently meet the needs of management in modern realities, these are:

– the need to integrate disparate data into a single information space to ensure the analytical activities of the university ;

– the lack of a unified approach to building an information space in universities, in terms of data integration and analysis;

– the need to solve various applied management problems using multi-aspect descriptions of the university system when making decisions;

– the need to use modern interactive tools that provide visual representation of large volumes of data, their analysis and convenient management .

Thus, the topic of the dissertation, devoted to the development of methods and algorithms for data integration based on ontologies to ensure analytical activities at the university, is relevant.

The aim of this dissertation is to develop methods and algorithms for data integration based on ontologies to ensure analytical activities at the university and informed management decision-making in the context of heterogeneity of data sources and the complexity of their analysis.

To achieve this goal, the following tasks have been set:

1. Conduct a systemic analysis of existing approaches to data integration based on ontologies, highlighting their advantages and disadvantages in the context of management decision-making.
2. Develop an ontology-based data integration model that takes into account both structured and unstructured data sources.
3. Develop methods and algorithms for constructing and updating ontologies that improve data consolidation and processing to support analytical activities and decision making.
4. Develop a modern interactive tool that provides visual representation of large volumes of data, their analysis and convenient management.

The object of the study is the processes of data integration in higher education institutions aimed at ensuring information and analytical activities and supporting management decisions.

Subject of the research : methods and algorithms for data integration based on ontologies, designed to unify and consolidate heterogeneous information resources of the university in order to improve analytical activities.

The following is submitted for defense:

– An ontological model for describing data integration for systematization and analysis of data in the corporate information system of a university.

– A method for presenting expert opinions for data-driven decision making

– An algorithm for forming a tuple of values and storing it in a database for integrating data from a corporate information system based on ontology.

– Method of forming the information space of the university to ensure the full functioning of the information space

– An algorithm for constructing an information and analytical system within the framework of a university's corporate information system for making management decisions.

The scientific novelty of the research lies in the following provisions:

– An ontological model for describing data integration has been developed for systematization and analysis of data in the corporate information system of the university, which is confirmed by publication in a peer-reviewed journal and author's certificate No. 0151 dated October 2, 2024

– Method of presenting expert opinions for decision-making based on data, which is confirmed by the adaptation and application of the developed methods in other universities, indicating their universality and novelty. Reflected in articles from the list of KOKNVO.

– The algorithm for forming a tuple of values and storing it in a database for integrating data of a corporate information system based on ontology, which is confirmed by the created interactive tool integrating the developed methods and demonstrating the practical applicability of the proposed solutions, is reflected in the implementation reports. In peer-reviewed journals and conferences,

– of the information space, which has not previously been implemented in this form and is confirmed by publication in peer-reviewed journals and conferences

– An algorithm for constructing an information and analytical system within the framework of a university's corporate information system for making management decisions, as confirmed by a publication in the WoS journal. Copyright certificate No. 24818 dated 06.04.2022.

The practical significance of the study lies in the possibility of direct application of the developed methods and algorithms in the activities of higher education institutions:

- The proposed methods allow for the integration of heterogeneous data, which provides more complete and reliable information for university managers and analysts.
- Systematization and analysis of data helps to identify bottlenecks in the educational process and allows taking measures to eliminate them, thereby improving the educational process.
- The use of modern technologies for data integration and analysis strengthens the university’s position in the educational space, increasing its attractiveness for applicants and partners.
- The developed solutions can be adapted and scaled for other educational institutions and organizations facing similar problems of integrating and analyzing large volumes of data.

Personal contribution of the dissertation candidate. As part of the dissertation research, the author independently completed the following volume of scientific and practical work:

– a deep systems analysis of modern methods of data integration based on

ontologies was carried out, especially in the context of higher education institutions;

– the shortcomings and limitations of existing information and analytical support systems in the field of education were identified, which made it possible to identify areas for improvement;

– an original model of data integration based on ontologies has been developed, taking into account the specifics of university information systems and allowing for the integration of both structured and unstructured data sources;

– new methods and algorithms for constructing and updating ontologies are proposed, which improve data consolidation processes and optimize their processing for analytical purposes;

– An experimental implementation of the developed system was carried out in a real university, and the system was configured and adapted to the specific requirements of the university.

– Practical advice on implementing the system in other educational institutions is offered, taking into account possible features and requirements. The results of the study were published in a number of scientific articles and reported at specialized conferences, the main idea and work in them belongs to the author.

Work approbation. The main results of the dissertation were reported at: International scientific and practical conference "Socio-economic and legal systems: modern vision" (Omsk, Siberian Institute of Business and Information Technology, March 2, 2016); International Kazakh-Russian seminar "Digital University" (Almaty, Kazakh National Pedagogical University named after Abay, February 21-23, 2018); at scientific seminars of the department "Information Systems" and " BigData " of Al- Farabi Kazakh National University . The results of the dissertation have been implemented at Al- Farabi Kazakh National University (implementation certificates in Appendix B). Training seminars were held among deans and heads of departments, directors of departments on the capabilities of the information and analytical system. Also on July 9, 2018. The results of the scientific work were heard at a meeting of the interdepartmental seminar of the information technology faculty of NSUEM (Novosibirsk). She was a member of the research group of the project AP08053145 Applied research of data management in the education system for decision-making (on the example of the university). Also, within the framework of the project 0207-17-ГК "Comprehensive automation of the activities of higher educational institutions, research institutes and the implementation of digital services", the results of the dissertation research were commercialized .

Publications. 27 scientific papers have been published on the topic of the dissertation, including 9 in journals recommended by the Committee for Control in Education and Science of the Republic of Kazakhstan, 10 in the proceedings of international conferences, including 1 indexed in the RSCI. Also, 5 articles have been published in journals indexed in the Scopus database , including 1 in conference materials, 1 in a periodical (20th percentile), 3 in a periodical (37th percentile). Two articles in journals indexed in WoS A monograph on the topic “Data management strategy for organizing the education system (using the university as an example)” was published ISBN 978-601-223-612-5. Received 2 copyright

certificates No. 24818 (computer program) and No. 27664 (work of science), 3 implementation certificates.

Structure and volume of the work. The dissertation consists of an introduction, four chapters, a conclusion, a bibliography (177 titles) and appendices. The main content of the dissertation is presented on 164 pages of typewritten text, illustrated with tables and figures.

The **first chapter** examines the theoretical foundations of ontology-based data integration in educational systems. An analysis of existing approaches to data integration is conducted, including traditional methods and the use of ontologies that ensure semantic compatibility and unification of heterogeneous data. Particular attention is paid to the advantages of ontologies for educational institutions, such as support for management decision-making, automation of analytical activities, and integration of external and internal data sources. The main challenges associated with the implementation of ontologies are highlighted, including the complexity of development and high resource requirements, which, despite these limitations, makes them an important tool for the effective management of educational and administrative activities in universities.

The second chapter of the dissertation is devoted to the development of methods and algorithms for data integration based on ontologies to improve the efficiency of analytical activities in universities. This chapter analyzes the needs of universities for analytical data, proposes a mathematical model for data integration, and describes algorithms for constructing and updating ontologies. Particular attention is paid to methods for processing both structured and unstructured data, which allows for holistic and automated data integration for making informed management decisions.

The third chapter is devoted to the development of the architecture of the information and analytical system (IAS) for managing educational and scientific activities at the university. The chapter examines in detail the basic principles of building a system based on the corporate information system (CIS) of the university, and also proposes a concept for developing analytical reports using cloud technologies. The integration of the system with various data sources, including questionnaires and administrative systems, is described. Much attention is paid to the creation of scenarios of analytical reports for different levels of university management, which allows optimizing decision-making processes. The implementation of the system makes it possible to automate data collection, increase their accuracy and reduce the time for preparing analytical materials.

The fourth chapter is devoted to the experimental verification and implementation of the proposed IAS in the corporate system of the university. The chapter presents the results of the system efficiency assessment based on the analysis of data obtained during its operation. The system demonstrated a high degree of automation of scientific and educational activity management processes, improved the monitoring of educational and scientific activity, and also increased the quality of management decisions. The analysis of the results of the system use was conducted, confirming its efficiency and the feasibility of further development and adaptation for other universities .