Syllabus
Course Program

Data warehouses

Specialty
121 – Software Engineering
122 – Computer Science

Institute
Institute of Computer Science and Information Technology

Educational program
Software Engineering
Computer Science and Intelligent Systems

Department
Software Engineering and Management Intelligent Technologies (321)

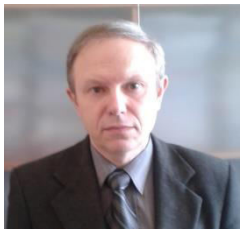
Level of education
Bachelor's level

Course type
Special (professional), Elective

Semester
4

Language of instruction
English, Ukrainian

Lecturers and course developers



Dmytro Orlovskyi

Dmytro.Orlovskyi@khpi.edu.ua
Ph.D., Associate Professor, Associate Professor of Software Engineering and Management Intelligent Technologies Department

Number of publications - more than 100.
Google Scholar:https://scholar.google.com/citations?user=bvEPOtYAAAAJ&hl
ORCID: https://orcid.org/0000-0002-8261-2988
Scopus: https://www2.scopus.com/authid/detail.uri?authorId=57202894400

More about the lecturer on the department's website

General information

Summary

The task of the discipline is the acquisition by students of the knowledge and skills necessary for solving problems related to the analytical processing of data in information systems, the development of data models that reflect the structure of data warehouses for various subject areas, the study of the theory and practice of effective organization of data warehouses

Course objectives and goals

Developing students' theoretical and practical knowledge necessary for the design and development of data warehouses in solving problems related to the development, maintenance, and quality assurance of software.

Format of classes

Lectures, laboratory classes, consultations, self-study. Final control in the form of a test.

Competencies

GC1. Ability to think abstractly, analyze and synthesize.
GC2. Ability to apply knowledge in practical situations.

GC3. Knowledge and understanding of the subject area and understanding of professional activities.
GC6. Ability to learn and master modern knowledge.
PC8. Ability to design and develop software using various programming paradigms: generalized, object-oriented, functional, logical, with appropriate models, methods and algorithms of computation, data structures and control mechanisms.
PC9. Ability to implement a multi-level computing model based on client-server architecture, including databases, knowledge and data warehouses, to perform distributed processing of large data sets on clusters of standard servers to meet the computing needs of users, including cloud services.

Learning outcomes

PLO10. To use tools for developing client-server applications, design conceptual, logical and physical models of databases, develop and optimize queries to them, create distributed databases, data warehouses and showcases, knowledge bases, including cloud services, using web programming languages.
PLO14. To apply knowledge of methodology and CASE tools for designing complex systems, methods of structural analysis of systems, object-oriented design methodology in the development and study of functional models of organizational, economic, production and technical systems.

Student workload

The total volume of the course is 150 hours (5 ECTS credits): lectures - 32 hours, laboratory classes - 32 hours, self-study - 86 hours.

Course prerequisites

Algorithmization and programming
Fundamentals of Computer Science and Artificial Intelligence
DataBases

Features of the course, teaching and learning methods, and technologies

Teaching and learning methods:
interactive lectures with presentations, discussions, laboratory classes, teamwork, case method, student feedback, problem-based learning.
Forms of assessment:
written individual assignments for laboratory work (CAS), assessment of knowledge in laboratory classes (CAS), express surveys (CAS), online tests (CAS), final/semester control in the form of a semester exam, according to the schedule of the educational process (FAS).

Program of the course

Topics of the lectures

Topic 1. Operational and analytical data processing
OLTP-systems, their purpose and features. Problems arising from the application of OLTP-systems and traditional databases for analytical data processing and decision support.
Topic 2. Basic concepts of data warehouses
Definition of data warehouse. Properties of data warehouses. Tools for working with data warehouses.
Topic 3. Data warehouses design
Categories of data in data warehouses, their classification. data warehouse schema types, their features.
Topic 4. The concept of data lakes
Purpose and main features of data lakes. Differences between data lakes and data warehouses.
Topic 5. Tools of integration and interoperability
Methods of corporate data integration. Uploading data to data lakes and data warehouses. ETL procedures and ELT procedures. Purpose and main features. Differences between ETL and ELT.
Topic 6. OLAP technologies
Definition of OLAP. FASMI test. Codd's rules for OLAP systems. The concept of a multidimensional data model. Concept of multidimensional space (hypercube). The main characteristics of a hypercube. Basic

operations on the hypercube. Data storage formats in OLAP cubes (ROLAP, MOLAP, HOLAP). General characteristics, advantages and disadvantages.

Topic 7. Peculiarities of data warehouses implementation

Variants of data warehouses implementation. Corporate Information Factory concept. Data Warehouse Bus concept.

Topic 8. Key performance indicators in analytical systems

Purpose and main features of key performance indicators. The concept of corporate performance management.

Topics of the workshops

Workshops are not provided within the discipline.

Topics of the laboratory classes

Topic 1. Creating an operational database (on the example of using the DBMS Microsoft SQL Server)
Topic 2. Creating a data warehouse (on the example of using the Microsoft SQL Server DBMS)
Topic 3. Creating ETL procedures for load data into the data warehouse
Topic 4. Creating a multidimensional database and tools for working with it
Topic 5. Development and application of key performance indicators

Self-study

Individual assignments are not provided in the curriculum.
Students are recommended with additional materials (videos, articles) for self-study and processing.

Course materials and recommended reading

Key literature

1. J. Kunigk, I. Buss, P. Wilkinson, L. George, Architecting Modern Data Platforms, O'Reilly Media, Inc., 2019, 633 p.
2. E. Sciore, Database Design and Implementation: Second Edition, Springer Nature, 2020, 468 p.
3. P. Baltzan, Business Driven Information Systems, McGraw-Hill Education, 2021, 809 p.
4. A. Simon, Data Lakes For Dummies, John Wiley & Sons, Inc., 2021, 387 p.
5. Data Management Body of Knowledge. Second Edition, Technics Publications, 2017, 778 p.

Additional literature

1. D. Petkovic, Microsoft SQL Server 2019: A Beginner's Guide. Seventh Edition, McGraw Hill Professional, 2020, 896 p.
2. Handbook of Big Data Analytics. Volume 1: Methodologies, The Institution of Engineering and Technology, 2021, 390 p.
3. Handbook of Big Data Analytics. Volume 2: Applications in ICT, security and business analytics. – The Institution of Engineering and Technology, 2021. – 419 p.
4. A. Meier, M. Kaufmann, SQL & NoSQL Databases: Models, Languages, Consistency Options and Architectures for Big Data Management, Springer, 2019, 229 p.
5. J. O. Padallan, Distributed Database Architecture, Arcler Press, 2021, 266 p.

Assessment and grading

Criteria for assessment of student performance, and the final score structure

Description of the final score structure, course requirements, and necessary steps to earn points, especially paying attention to self-study and individual assignments.

Grading scale

Total points	National	ECTS
90–100	Excellent	A
82–89	Good	B
75–81	Good	C
64–74	Satisfactory	D
60–63	Satisfactory	E
35–59	Unsatisfactory (requires additional learning)	FX
1–34	Unsatisfactory (requires repetition of the course)	F

Norms of academic integrity and course policy

The student must adhere to the Code of Ethics of Academic Relations and Integrity of NTU "KhPI": to demonstrate discipline, good manners, kindness, honesty, and responsibility. Conflict situations should be openly discussed in academic groups with a lecturer, and if it is impossible to resolve the conflict, they should be brought to the attention of the Institute's management.
Regulatory and legal documents related to the implementation of the principles of academic integrity at NTU "KhPI" are available on the website: http://blogs.kpi.kharkov.ua/v2/nv/akademichna-dobrochesnist/

Approval

Approved by 08.06.2023 Head of the department
Ihor HAMAIUN

08.06.2023 Guarantors of the educational programs
Andrii KOPP
Uliya LITVINOVA