System integration - course description

General information

Course name	System integration
Course ID	11.3-WE-BizEIP-SystInteg-Er
Faculty	Faculty of Computer Science, Electrical Engineering and Automatics
Field of study	E-business
Education profile	practical
Level of studies	First-cycle Erasmus programme
Beginning semester	winter term 2021/2022

Course information

Semester	6
ECTS credits to win	5
Course type	obligatory
Teaching language	english
Author of syllabus	• dr hab. inż. Marek Kowal, prof. UZ

Classes forms

The class form	Hours per semester (full-time)	Hours per week (full-time)	Hours per semester (part-time)	Hours per week (part-time)	Form of assignment
Lecture	30	2	-	-	Exam
Laboratory	30	2	-	-	Credit with grade

Aim of the course

Presentation of the software used for data integration. Overview of typical data sources. Presentation of data cleaning and transformation methods. Familiarize students with the architecture of ETL systems (Extract, Transform, Load). Developing practical skills in operating selected systems supporting ETL processes. Presentation of the issue of semantic data integration. Familiarization with integration platform architectures.

Prerequisites

Databases, Data warehouses and reporting services

Scope

Data integration using ETL (Extract, Transform, Load) systems. The architecture of ETL systems. Data sources: relational databases, flat files, XML, JSON. Data transformation and integration methods. Profiling and data cleansing. Software tools supporting the design and deployment of ETL processes. Structure of the ETL package. Learning to program ETL packages using SQL Server Integration Services (SISS). Configuration of SISS packages.

The architecture of the integration platforms. Enterprise ESB (Enterprise Service Bus) architecture. SOA (Service-Oriented Architecture). Microservices. REST and SOAP technologies.

Teaching methods

Lecture - conventional lecture using a video projector.
Laboratory - practical exercises in the computer laboratory.

Learning outcomes and methods of theirs verification

Outcome description	Outcome symbols	Methods of verification	The class form
The student knows the key concepts, definitions, and application areas of data integration methods for the purposes of business data analytic		• an examination test with score scale	• Lecture
The student knows the functioning of the microservices architecture and the corporate architecture of the ESB data bus.		• an examination test with score scale	• Lecture
The student knows the organization and architecture of ETL systems in data warehouses.		• an examination test with score scale	• Lecture
Development of an exemplary concept and an IT systems integration plan for a given problem		• a discussion • a written assignment	• Laboratory
The student can design and program an ETL package with a data transformation operation, and then secure it and run it on the server.		• a quiz • carrying out laboratory reports	• Laboratory

Outcome description	Outcomesymbols	Methods of verification	The class form
The student can prepare an ETL package for data extraction from heterogeneous sources.		• a quiz • carrying out laboratory reports	• Laboratory
The student knows methods of modeling and data storage in ETL subsystems for the needs of a data warehouse.		• an examination test with score scale	• Lecture
The student can create various data flow models using control flow components.		• a quiz • carrying out laboratory reports	• Laboratory
The student knows REST and SOAP technologies.		• an examination test with score scale	• Lecture
The student can design and program an ETL package to load data in a data warehouse.		• a quiz • carrying out laboratory reports	• Laboratory
The student can prepare a report on a selected system integration technology and legal solutions regulating data processing in data integration systems.		• a discussion • a written assignment	• Laboratory

Assignment conditions

Lecture - the passing criteria is to obtain positive grades from tests carried out at least once in a semester.

Laboratory - the passing criterion is to obtain positive marks for laboratory exercises and tests.

Final mark components = lecture: 50% + teaching laboratory: 50%

Recommended reading

1. Kimball, R., Caserta J., The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data, Wiley, 2004.
2. AnHai, D., Halevy A., Ives Z., Principles of Data Integration, Morgan Kaufman, 2012.
3. Aspin, A., SQL Server 2012 Data Integration Recipes: Solutions for Integration Services and Other ETL Tools, APress, 2012.
4. Morris, J., Practical Data Migration, The British Computer Society, 2012.
5. Ehrig, M., Ontology Alignment: Bridging the Semantic Gap, Springer, 2006.
6. Ghosh P.: Semantic Integration of Applications: Application Integration By Linking Semantically Related Objects Shared Across Applications, CreateSpace Independent Publishing Platform, 2016
7. Chappell D.: Enterprise Service Bus: Theory in Practice, O'Reilly, 2004

Further reading

1. Cody, R.: Cody's Data Cleaning Techniques Using SAS, SAS Press, 2008.
2. Egger, N., Fiechter, J-M. R., Salzmann R., Sawicki R, Thielen T., SAP BW Data Retrieval: Mastering the ETL Process, SAP Press, 2006.
3. Bobak, A.: Connecting the Data: Data Integration Techniques for Building an Operational Data Store (ODS), Technics Publications LLC, 2012.
4. McGilvray, D., Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information, Morgan Kaufmann, 2008.
5. Maydanchik, A., Data Quality Assessment, Technics Publications, 2007.
6. Ozsu, T., Valduriez P., Principles of Distributed Database Systems, Springer, 2011.

Notes

Modified by dr hab. inż. Marek Kowal, prof. UZ (last modification: 21-07-2021 09:36)