1. **Course title:** DEVELOPMENT TRENDS IN DATA ENGINEERING

2. **Course code**

3. **Validity of course description:** 2018/2019

4. **Level of studies:** MSc programme

5. **Mode of studies:** intramural studies

6. **Field of study:** COMPUTER SCIENCE (FACULTY SYMBOL)

7. **Profile of studies:**

8. **Programme:** DATABASES AND SYSTEMS ENGINEERING

9. **Semester:** 1, 2

10. **Faculty teaching the course:** Institute of Informatics, RAu2

11. **Course instructor:** Henryk Josiński, PhD

12. **Course classification:**

13. **Course status:** compulsory

14. **Language of instruction:** English


Student understands relational database model and architecture of a database management system. He/she has skills to prepare the database schema with normalized relations and is capable of formulating SQL queries as well as creating database applications.

16. **Course objectives:** The course includes lecture and lab exercises. Purpose of the subject is to teach students how to develop and use modern database systems. The goal of the lecture is to teach / present students topics / issues from the area of data engineering related to processing of diverse, large and variable data sets: data models, data structures, areas of application, selected mechanisms of database management system, query languages. The goal of the lab exercises is to practice the processing of diverse data sets based on different data models: data storing, modifying, retrieving.

17. **Description of learning outcomes:**

<table>
<thead>
<tr>
<th>Nr</th>
<th>Learning outcomes description</th>
<th>Method of assessment</th>
<th>Teaching methods</th>
<th>Learning outcomes reference code</th>
</tr>
</thead>
<tbody>
<tr>
<td>W1</td>
<td>Student understands the concept of a non-relational database model and is capable of indicating areas of application for presented database solutions.</td>
<td>exam, reports, tests</td>
<td>lecture, laboratory</td>
<td>K2A_W07 K2A_W08 K2A_W09</td>
</tr>
<tr>
<td>W2</td>
<td>Student understands the differences between row- and column-oriented data structures and understands the consequences arising therefrom.</td>
<td>exam, reports, tests</td>
<td>lecture, laboratory</td>
<td>K2A_W08 K2A_W09 K2A_U11</td>
</tr>
<tr>
<td>W3</td>
<td>Student understands the necessity of elimination of certain ACID properties in NoSQL databases in favor of CAP compromises.</td>
<td>exam, reports, tests</td>
<td>lecture, laboratory</td>
<td>K2A_W05 K2A_W06 K2A_W07</td>
</tr>
<tr>
<td>U1</td>
<td>Student is capable of creating databases of different types as well as processing diverse data sets based on different data models: data storing, modifying, retrieving.</td>
<td>exam, reports, tests</td>
<td>laboratory</td>
<td>K2A_U01 K2A_U11 K2A_U13</td>
</tr>
<tr>
<td>U2</td>
<td>Student is capable of properly using XML.</td>
<td>exam, reports, tests</td>
<td>lecture, laboratory</td>
<td>K2A_U01 K2A_U10 K2A_U13</td>
</tr>
<tr>
<td>U3</td>
<td>Student is aware of influence of appropriate selection of database solutions on processing</td>
<td>exam, reports, tests</td>
<td>lecture, laboratory</td>
<td>K2A_U11 K2A_U12</td>
</tr>
</tbody>
</table>
18. Teaching modes and hours
Lecture / BA / MA Seminar / Class / Project / Laboratory
Lecture 30 h., Laboratory 45h

19. Syllabus description:
Lectures:
Big Data. The 3/5V model, data storing, programming and processing.
NoSQL databases (2 lectures). Data models used by NoSQL databases: „key-value“ (Amazon Dynamo), wide column (Google BigTable, Cassandra), graph (Neo4j and Cypher query language as well as Apache TinkerPop tools), and document (MongoDB, CouchDB).
Full-Text Search (FTS) mechanism. Indexing binary data. Data search modes. Predicates and operators used in queries. FTS in selected database systems.
Query languages for XML – XPath, XQuery.
Architecture, administration and security of a database management system based on the case of Oracle DBMS.
The Flashback Queries (Flashback Database, Flashback Table, Flashback Drop) mechanism in Oracle DBMS. Retrieving metadata and historical data for a specific time interval.
Laboratory:
12 lab exercises (3h/week):
Big Data – a model, data storing, programming and processing.
Document, graph and column-oriented data models in NoSQL databases (3 lab exercises).
IBM solidDB in-memory database.
Fuzzy databases.
Distributed database and data replication (2 lab exercises).
Architecture, administration and security in Oracle DBMS.
Flashback queries in Oracle DBMS.
XML – mapping between relational databases and XML, XML in database management systems (2 lab exercises).
The database management systems used in lab exercises: Oracle, solidDB, Cassandra, MongoDB, CouchDB, Neo4j, MS SQL Server, MySQL, PostgreSQL.

20. Examination: after 2nd semester – written exam

21. Primary sources:
G. Harrison: Next Generation Databases: NoSQLand Big Data (1st Edition)
22. Secondary sources:
A. Ploetz, D. Kandhare, S. Kadambi, X. Wu: Seven NoSQL Databases in a Week: Get up and running with the fundamentals and functionalities of seven of the most popular NoSQL databases, 2018.
Internet sources presented during the lectures and laboratories

23. Total workload required to achieve learning outcomes

<table>
<thead>
<tr>
<th>Lp.</th>
<th>Teaching mode</th>
<th>Contact hours / Student workload hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lecture</td>
<td>30/30</td>
</tr>
<tr>
<td>2</td>
<td>Classes</td>
<td>/</td>
</tr>
<tr>
<td>3</td>
<td>Laboratory</td>
<td>45/45</td>
</tr>
<tr>
<td>4</td>
<td>Project</td>
<td>/</td>
</tr>
<tr>
<td>5</td>
<td>BA/ MA Seminar</td>
<td>/</td>
</tr>
<tr>
<td>6</td>
<td>Other</td>
<td>15/0</td>
</tr>
<tr>
<td></td>
<td>Total number of hours</td>
<td>90/75</td>
</tr>
</tbody>
</table>

24. Total hours: 165

25. Number of ECTS credits: 3

26. Number of ECTS credits allocated for contact hours: 3

27. Number of ECTS credits allocated for in-practice hours (laboratory classes, projects): 2

26. Comments:

Approved:

………………………………………………………………………………………..
(date, Instructor’s signature)........................................................................

………………………………………………………………………………………..
(date, the Director of the Faculty Unit signature)........................................