Government of Russian Federation
Federal State Autonomous Educational Institution of High Education
National Research University «Higher School of Economics»

Faculty of Computer Science
School of Data Analysis and Artificial Intelligence

Syllabus for the course
«Semantic Technologies»

for educational program «Applied Mathematics and Informatics»,
specialization 01.03.02 «Applied Mathematics and Informatics», Bachelor's program

Author:
Michael V. Zakharyaschev, professor, mzakharyaschev@hse.ru

Approved on the meeting of the School of Data Analysis and Artificial Intelligence
Head of the School Sergei O. Kuznetsov
_____

«__»_____2017 г.

APPROVED BY
Academic supervisor of the educational program «Applied Mathematics and Informatics» in specialization 01.03.02 «Applied Mathematics and Information Science» Anton .S. Konushin
_____

«__»_____2017 г.

Recommended by Academic Council of the Programme «Applied Mathematics and Information Science»

«__»_____2017 г.

Manager of the School of Data Analysis and Artificial Intelligence Larisa I. Antropova
_____

Moscow, 2017

1. Scope of Use

The present program establishes minimum requirements on students' knowledge and skills, and determines the content of the course.
The present syllabus is aimed at the department teaching the course, their teaching assistants, and students of the Bachelor's program 01.03.02 «Applied Mathematics and Informatics».
This syllabus meets the standards required by:

- Educational standards of National Research University Higher School of Economics;
- Educational program «Applied Mathematics and Informatics» of Federal Bachelor's Degree Program 01.03.02, 2017;
- University curriculum of the educational program «Applied Mathematics and Informatics» (01.03.02) for 2017.

2. Summary

This course is an introduction to Semantic Technologies that provide easier ways to find, share, reuse and combine information. Semantic Technologies define and link data on the Web or within an enterprise by developing languages to express rich, self-describing interrelations of data in a form that machines can process. They provide an abstraction layer above existing IT technologies that connects data, content and processes. Semantic Technology standards developed by W3C include

- a flexible data model RDF (Resource Description Framework) for storing data in graph databases
- schema and ontology languages for describing concepts and relationships (RDFS and OWL)
- the query language SPARQL designed to query data across various systems and databases and to retrieve and process data stored in RDF format.

Applications of Semantic Technologies range from Linked Data, Wikidata, Healthcare and Pharma Industry, Supply Chain Management, Publishing and Media Management, Web Search and E-commerce to Data Integration in the Oil & Gas industry.

3. Learning Objectives

The aims of this module are to

- introduce the theoretical foundations of Semantic Technologies, including the languages RDF/S, SPARQL, the Web Ontology Language OWL;
- provide the students with practical skills of modelling data using RDF/S, querying RDF triplestores, relational databses and XML documents, building ontologies and using datalog;
- overview the current applications of Semantic Technologies in health care, media management, and industry;
- demonstrate a few standard algorithms for classification of concepts in ontologies and answering queries.

4. Learning outcomes

By the end of the module, the student should be able to:

- understand fundamental concepts, advantages and limitations of Semantic Technologies;
- understand and use the RDF framework and associated technologies such as RDFa and SPARQL;
- understand and use the ontology language OWL 2 and its profiles;
- understand the principles of ontology-based data access and integration;
- understand the basics of knowledge representation with description logics;
- understand and use deductive database systems.

5. Place of the Discipline in the Program Structure

This is a compulsory course for specialization "Artificial Intelligence and Structural analysis", 01.03.02 "Applied Mathematics and Informatics".

The following knowledge and competences are useful for better understanding of the course:

- Basic English language, both oral and written.
- Discrete Mathematics.
- Logic.
- Programming.
- Data Bases.

6. Schedule

Topic	Total hours	Contact hours		Self-study
		Lectures	Practice lessons	
1. Introduction	14	2	2	10
2. XML/XML Schema, XPath	14	2	2	10
3. Querying relational databases	14	2	2	10
4. RDF/RDFs, language Turtle	14	2	2	10
5. Query language SPARQL	14	2	2	10
6. Ontology-based data access	14	2	2	10
7. OWL, Ontology engineering	16	2	4	10
8. Reasoning with OWL	24	4	4	16
9. Deductive databases	28	4	4	20
Total	152	22	24	106

7. Requirements and Grading

Homework	10	Weekly exercises
Exam	1	Written exam with discussion. Preparation time – 120 min.

8. Assessment

The assessments consist of several in-class practical tasks.

Final assessment is the final exam. Students have to demonstrate knowledge of the material covered during the entire course and the ability to apply the materials.

The grade formula:

Final course mark is obtained from the following formula:

$$Final = 0.6 \cdot Grade_{Cumulative} + 0.4 \cdot Grade_{Exam}$$

where *the final exam* is worth 40% of the final mark and *the average mark for cumulative grade* 60%.

All grades having a fractional part greater than 0.5 are rounded up.

Table of Grade Accordance

Ten-point Grading Scale	**Five-point Grading Scale**	
1 – very bad 2 – bad 3 – no pass	Unsatisfactory - 2	**FAIL**
4 – pass 5 – highly pass	Satisfactory – 3	**PASS**
6 – good 7 – very good	Good – 4	
8 – almost excellent 9 – excellent 10 – perfect	Excellent – 5	

9. Course Plan

1. Introduction to the module. Ontologies in (Computer) Science. Knowledge graphs. Schema.org. Wikidata.

   Seminar: building a Don Corleone family ontology.

2. Is XML a semantic technology? The tree model of XML documents, XML Schema. Querying XML documents with XPath. JSON (JavaScript Object Notation).

   Seminar: Practical querying XML documents with XPath.

3. Querying relational databases: a quick survey.

Seminar: Practical querying relational databases with PostgreSQL.

4. Resource Description Framework (RDF). RDF Schema. RDF/S semantics. Terse RDF Triple Language Turtle.

   Seminar: producing RDF triples: by hand, from tables, from natural texts.

5. RDF query Language SPARQL.

   Seminar: Querying RDF triplestores: DBpedia, German cities. Setting up and querying an Apache Fuseki triplestore.

6. Introduction to ontology-based data access (OBDA). OBDA platform Ontop.

   Seminar: setting up ontology-based access to the IMDB database.

7. Requirements for ontology languages. From RDFS to OWL. OWL ontologies. Ontology. Ontology engineering. OWL ontologies in life sciences and industry.

   Seminar: designing an ontology using Protege.

8. Reasoning with OWL. Introduction to Description Logic and formal semantics.

   Seminar: answering queries with Protege.

9. Deductive databases. Knowledge representation and query answering with datalog.

   Seminar: query answering with DLV and RDFox.

10. Methods of Instruction

The following educational technologies are used in the study process:

- slides provided by the lecturer
- discussion and analysis in the exercise classes
- office hours

Practical classes using PC. Ontology engineering tasks.

11. Reading and Materials

Literature:

(1) P. Szeredi, G. Lukacsy and T. Benko. The Semantic Web Explained. The technology and mathematics behind Web 3.0. Cambridge University Press, 2014. ISBN 978-0-521-70036-8

(2) David Wood, Marsha Zaidman and Luke Ruth. Linked Data. Manning Publications, 2013. ISBN 9781617290398.

(3) Pascal Hitzler, Markus Kroetzsch and Sebastian Rudolph. Foundations of Semantic Web Technologies. Chapman & Hall, 2009. ISBN 978-1420090505.

(4) Grigoris Antoniou and Frank van Harmelen. A Semantic Web Primer. MIT Press, 2004. ISBN 0-262-01210-3.

Additional literature:

(1) Handbook on Ontologies, eds. S. Staab and R. Studer, Springer 2009

12. Equipment

Sufficient PC quantity for students. Distant course support.