Government of Russian Federation

Federal State Autonomous Educational Institution of High Professional Education

"National Research University Higher School of Economics"

Syllabus for the course
"Introduction to the Semantic Web Technologies"

(Семантическая паутина)

for postgraduate program in 02.06.01 Computer and Information science / 05.13.11 "Mathematical Theory and Software for Computing Machinery, Systems, and Networks", 05.13.17 "Theoretical Foundations of Computer Science"

Authors:
M.V. Zakharyaschev, visiting professor, mzakharyaschev@hse.ru

Approved by the Academic Council of the School for Postgraduate Studies in Computer Science by October 19, 2015

Moscow - 2015

1. Scope of Use

This program establishes the minimal requirements to postgraduate students' knowledge and skills for 02.06.01 Computer and Information science / 05.13.11 "Mathematical Theory and Software for Computing Machinery, Systems, and Networks", 05.13.17 "Theoretical Foundations of Computer Science", and determines the content of the course and educational techniques used in teaching the course.

The present syllabus is aimed at faculty teaching the course and postgraduate students studying 02.06.01 Computer and Information science / 05.13.11 "Mathematical Theory and Software for Computing Machinery, Systems, and Networks", 05.13.17 "Theoretical Foundations of Computer Science".

This syllabus meets the standards required by:

- Educational standards of National Research University Higher School of Economics;
- Postgraduate educational program for 02.06.01 Computer and Information science.
- University curriculum of the postgraduate program for 02.06.01 Computer and Information science / 05.13.11 "Mathematical Theory and Software for Computing Machinery, Systems, and Networks", 05.13.17 "Theoretical Foundations of Computer Science", approved in 2015.

2. Summary

This course is a gentle introduction to the theory and practice of the Semantic Web, an extension of the current Web that provides an easier way to find, share, reuse and combine information. It is based on machine-readable information and builds on XML technology's capability to define customized tagging schemes, RDF's (Resource Description Framework) flexible approach to representing data, the OWL (Web Ontology Language) schema language and SPARQL query language. The Semantic Web provides common formats for the interchange of data (where on the Web there is only an interchange of documents). It also provides a common language for recording how data relates to real world objects, allowing a person or a machine to start off in one database, and then move through an unending set of databases which are connected not by wires but by being about the same thing. Important applications of the Semantic Web technologies include Healthcare (SNOMED CT), Supply Chain Management (Biogen Idec), Media Management (BBC), Data Integration in the Oil & Gas industry (Chevron, Statoil), Web Search and E-commerce.

3. Learning Objectives

The learning objective of the course "The Semantic Web" is to

(i) introduce the theoretical foundations of the Semantic Web, including the standard W3C data, query and ontology languages such as Resource Description Framework (RDF), SPARQL query language, the Web Ontology Language OWL, and the corresponding knowledge representation technologies;

(ii) provide the students with practical skills of building ontologies and querying the Web;

(iii) overview the current applications of the Semantic Web technologies in health care, media management, and industry.

4. Learning Outcomes

After completing the study of the discipline the PhD student should:

(i) understand fundamental concepts, advantages and limits of the Semantic Web;

(ii) understand fundamental concepts, advantages and limits of the Semantic Web;
(iii) understand and use the RDF framework and associated technologies such as RDFa and SPARQL;
(iv) understand and use the ontology language OWL 2 and its profiles;
(v) understand the principles of ontology-based data access;
(vi) understand the basics of the underlying knowledge representation and reasoning formalisms such as description logic.

After completing the study of the discipline the PhD student should have developed the following competencies:

Competence	Code	Descriptors (indicators of achievement of the result)	Educative forms and methods aimed at generation and development of the competence
The ability to carry out research in the field of professional activity using current research methods and information and communication technologies.	ОПК-1	Students obtain necessary knowledge to understand and formulate the theoretical difficulty of problems they are solving.	Tutorials and classes
The ability to choose and effectively use educational technologies, methods, and tools for ensuring a desired level of personal and professional development of the learner.	ОПК-3	Students learn current applications of the Semantic Web technologies, apply knowledge of the standard W3C data, query and ontology languages such as Resource Description Framework (RDF) in practical tasks	Lectures, tutorials
The ability to develop and analyze models of information processes and structures	ПК-5	The PhD student is able to develop and analyze ontology languages and semantic web models, implement them in a programming language and select a proper model for particular problems.	Lectures, tutorials, and assignments.

5. Place of the Discipline in the Postgraduate Program Structure

This is an elective course for 05.13.11 "Mathematical Theory and Software for Computing Machinery, Systems, and Networks", 05.13.17 "Theoretical Foundations of Computer Science".

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. The Semantic Web	32	4	4	24
2. RDF and SPARQL	48	6	6	36
3. Ontology engineering	64	8	8	48
4. Description Logic and formal semantics	64	8	8	48
5. Ontology-based data access.	58	8	6	44
Total	266	34	32	200

7. Requirements and Grading

Homework		Weekly
Test	3	In-class tests. Preparation time – 80 min.
Exam	1	Written exam with discussion. Preparation time – 120 min.

8. Assessment

The assessments consist of written in-class tests.

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.

9. The grade formula

The final exam is worth 37% of the final mark.

Final course mark *is obtained from the following formula*:

Final=0.63*(Test1+Test2+Test3) + 0.37*(Exam).

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	

10. Course description.

1. The Semantic Web.

The history of the Semantic Web. Syntactic vs semantic web. Ontologies in (Computer) Science. The layered approach to the Semantic Web. XML, the tree model of XML documents, XML Schema. Querying XML documents, XPath.

2. RDF and SPARQL.

Resource Description Framework (RDF). RDF Schema. RDF/S semantics. Linked Data. Terse RDF Triple Language Turtle. SPARQL query language. SPARQL endpoints.

3. Ontology engineering.

Requirements for ontology languages. From RDFS to OWL. Three species of OWL. OWL ontologies. OWL ontologies in life sciences and industry. Protégé ontology editor and framework for building intelligent systems. Reasoning with OWL. Open vs. closed worlds.

4. Description Logic and formal semantics.

Description logic EL (the OWL 2 EL profile of OWL 2). Description logic ALC. Tableau algorithms. Description logics extending ALC. First-order logic.

5. Ontology-based data access.

Instance data as ABoxes. Entailment regimes for SPARQL queries. Ontology-based data access with Ontop.

11. Educational technologies

The following educational technologies are used in the study process:

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

Methods of Instruction:

Practical classes using PC. Individual course projects. Ontology engineering tasks.

12. Final exam guidelines

http://www.dcs.bbk.ac.uk/~michael/sw15/sw15.html

13. Reading and Materials

Literature:

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

Additional literature:

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

Literature for self-study:

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

(ii) 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.

(iii) Lecture slides and online resources provided by the lecturer.

14. Equipment.

Sufficient PC quantity for students. Distant course support.