COURSE DESCRIPTION CARD - SYLLABUS

Course name
Knowledge Extraction [S2SI1E>EKW]

Course

Field of study
Artificial Intelligence

Year/Semester
1/2

Area of study (specialization)
–

Profile of study
general academic

Level of study
second-cycle

Course offered in
English

Form of study
full-time

Requirements
compulsory

Number of hours

Lecture
15

Laboratory classes
0

Other (e.g. online)
0

Tutorials
0

Projects/seminars
15

Number of credit points

3,00

Coordinators

dr hab. inż. Agnieszka Ławrynowicz prof. PP
agnieszka.lawrynowicz@put.poznan.pl

Lecturers

Prerequisites

A student starting this course should have basic knowledge of: web technologies (including XML, JSON data representation formats), logic, database systems, artificial intelligence, natural language processing and Python programming. The student should also understand the necessity of extending their competences and be willing to cooperate within a team. Moreover, in terms of social competence, a student should demonstrate such attitudes as honesty, responsibility, perseverance, cognitive curiosity, creativity, personal culture, respect for other people.

Course objective

1. To provide students with knowledge of knowledge graphs and knowledge extraction. 2. To develop in students the ability to solve problems in how to use and implement methods for knowledge extraction and knowledge-intensive systems. 3. To develop students' skills in applying artificial intelligence and data analysis methods to extract knowledge and construct knowledge graphs. 4. To develop in students the skills of teamwork.

Course-related learning outcomes

Knowledge The student:

1. Has advanced and in-depth knowledge in the field of widely understood information systems using knowledge bases, theoretical basis of their construction and methods, tools and programming environments used for their implementation [K2st_W1].
2. Has structured and theoretically underpinned general knowledge related to key issues in knowledge extraction [K2st_W2].
3. Has advanced detailed knowledge related to selected issues in knowledge extraction [K2st_W3].
4. Has knowledge of development trends and the most significant new developments in computer science and other selected related scientific disciplines within the field of knowledge extraction [K2st_W4].
5. Knows advanced methods, techniques and tools used in solving complex engineering tasks and conducting research work in the area of knowledge extraction [K2st_W6].
Skills The student:
1. Is able to acquire information from literature, databases and other sources, integrate it, interpret and critically evaluate it, draw conclusions and formulate and fully justify opinions [K2st_U1].
2. Is able to plan and carry out experiments, including measurements and computer simulations, interpret the obtained results and draw conclusions, and formulate and verify hypotheses related to complex engineering problems in the field of knowledge extraction and simple research problems [K2st_U3].
3. Is able to use analytical, simulation and experimental methods to formulate and solve engineering tasks in the field of knowledge extraction and simple research problems [K2st_U4].
4. Is able - when formulating and solving tasks in the field of knowledge extraction - to integrate knowledge from different areas of computer science (and, if necessary, also knowledge from other scientific disciplines) and apply a system approach, taking into account also non-technical aspects [K2st_U5].
5. Is able to assess the usefulness and applicability of new developments (methods and tools) of modeling and knowledge engineering and new IT products using ontologies and knowledge graphs [K2st_U6].
6. Is able to critically analyze existing technical solutions in the field of knowledge extraction and propose their improvements (enhancements) [K2st_U8].
7. Is able to assess the suitability of methods and tools for solving an engineering task involving the construction or evaluation of an information system or its components (in particular, ontologies, knowledge graphs or other artifacts in the area of knowledge representation), including recognizing the limitations of these methods and tools [K2st_U9].
8. Is able - using, among others, conceptually new methods - to solve complex information technology tasks, including non-standard tasks and tasks with a research component, in particular in the area of knowledge extraction [K2st_U10].
Social competences The student:
1. Understands that in computer science, knowledge and skills become obsolete very quickly [K2st_K1].
2. Understands the importance of using the latest knowledge in computer science, especially in the area of knowledge extraction, in solving research and practical problems [K2st_K2].

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:

Formative assessment:
(a) in terms of lectures: based on answers to questions on the material discussed and activity.
b) for the project: based on assessing the progress of the partial tasks concerning the realization of the project.
Summative assessment:
a) in the scope of lectures, verification of the assumed educational effects is realized by the exam.
b) in the scope of the project, the verification of the assumed learning outcomes is realized by: evaluation of skills related to the implementation of the project.
Pass mark: 50% of the points.

Programme content

Knowledge graphs (entity and property centric models)
Knowledge engineering (methodologies, tools, good practices)
Entity recognition
Relation extraction
Entity linking and semantic tagging
Knowledge graph refinement (link prediction, entity and triple classification, fact checking)
Knowledge extraction and construction from pre-trained large language models (GPT-3 etc.)

Course topics

none

Teaching methods

lecture: multimedia presentation, demonstration of exemplary solutions, quizes
project: reading and analysing literature on the topic of the project, designing the solution, implementation, evaluation, teamwork

Bibliography

Basic 1. Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia d'Amato, Gerard de Melo, Claudio Gutierrez, Sabrina Kirrane, José Emilio Labra Gayo, Roberto Navigli, Sebastian Neumaier, Axel-Cyrille Ngonga Ngomo, Axel Polleres, Sabbir M. Rashid, Anisa Rula, Lukas Schmelzeisen, Juan Sequeda, Steffen Staab, Antoine Zimmermann (2021) Knowledge Graphs, Synthesis Lectures on Data, Semantics, and Knowledge, No. 22, 1–237, DOI: 10.2200/S01125ED1V01Y202109DSK022, Morgan & Claypool, available at https://kgbook.org
2. Tom Heath and Christian Bizer, Linked Data: Evolving the Web into a Global Data Space (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool, 2011, available at http://linkeddatabook.com/book
Additional 1. Mayank Kejriwal, Craig A. Knoblock, and Pedro Szekely (Eds.). 2021. Knowledge Graphs: Fundamentals, Techniques, and Applications. The MIT Press.
2. Michael Uschold, Demystifying OWL for the Enterprise, Morgan Claypool Publishers, 2018
3. Dean Allemang, Jim Hendler, Fabien Gandon, Semantic Web for the Working Ontologist, Third Edition, ACM Books, 2020
4. Robert Arp, Barry Smith, and Andrew D. Spear. 2015. Building Ontologies with Basic Formal Ontology. The MIT Press.
5. Agnieszka Ławrynowicz, Semantic data mining. An ontology-based approach, Studies on the Semantic Web, vol. 29. IOS Pres/AKA Verlag 2017
6. Altinok, Duygu. Mastering spaCy: An end-to-end practical guide to implementing NLP applications using the Python ecosystem. Packt Publishing Ltd, 2021.
7. Sandra Kublik, Shubham Saboo: GPT-3: Building Innovative NLP Products Using Large Language Models, O'Reilly Media, 2022

Breakdown of average student's workload

	Hours	ECTS
Total workload	75	3,00
Classes requiring direct contact with the teacher	30	1,50
Student's own work (literature studies, preparation for laboratory classes/ tutorials, preparation for tests/exam, project preparation)	45	1,50