COURSE OFFERED IN THE DOCTORAL SCHOOL

Code of the course	4606-ES-000000C-0022	Name of the course	Polish	Reprezentacja wiedzy w sztucznej inteligencji
			English	**Knowledge Representation in Artificial Intelligence (KRAI)**
Type of the course	General courses			
Course coordinator	**Prof. Mieczysław Muraszkiewicz**			
Implementing unit	WEiTI	Scientific discipline / disciplines*	information and communication technology	

Level of education	Doctoral studies	Semester	Winter and Summer		
Language of the course	English				
Type of assessment:	Graded credit	Number of hours in a semester	60	ECTS credits	4
Minimum number of participants	10	Maximum number of participants	24	Available for students (BSc, MSc)	~~Yes~~/No

Type of classes		Lecture	Auditory classes	Project classes	Laboratory	Seminar
Number of hours	in a week	2	0	2	0	0
	in a semester	30	0	30	0	0

* does not apply to the Researcher's Workshop

1. Prerequisites

- Knowledge of basic concepts of artificial intelligence.
- Knowledge of logic (first order predicate calculus).
- Knowledge of basic notions and theorems of graph theory.
- Programming skills in at least one of the following languages: C, C ++, Java, Python, R.
- General interest in issues of a philosophical nature, particularly in epistemology.

2. Course objectives

The aim of the course is to familiarize Doctoral School students with the main methods and techniques of knowledge representation that are used in the field of artificial intelligence, especially in relation to machine learning systems, decision support systems, intelligent information systems, and data analysis systems. The lecture prepares students to design assorted components of such systems and to undertake research in this field.

The lecture is organised as a two steps endeavour. First, basic concepts, including knowledge representation methods and artificial intelligence are introduced, and then an overview of selected knowledge representation methods is given. The motto of the lecture is R. Hamming's saying: "The purpose of computing is insight, not numbers." The lecture is accompanied by students' projects that are practical applications of the lecture's topics.

3. Course content (separate for each type of classes)

Lecture

1. **Introduction. Basic concepts: data, information, knowledge, intelligent system (2 hours)**
   A linguistic approach is used to define the concepts of data, information, knowledge, and inteligent system. Data is defined as a pair of an arbitrary alphabet and a set of rules that allow one to build up strings composed of the alphabet elements. The data has no meaning. Information is meant as data plus an interpretation and/or as assertions, thus its includes semantics. Knowledge is represented by assertions in which is embedded a reasoning feature (a syllogism). Intelligent systems are such systems that demonstrate at least the ability to reason and learn. Examples of the above notions are discussed.

2. **Knowledge representation (5 hours)**
A formal definition of knowledge representation as a pair of a chosen language and a set of knowledge manipulation operators, with an emphasis on inference mechanism(s) is given. The inference mechanism is a crucial element of this definition as it determines the level of "intelligence" of the system. The choice of knowledge representation determines the effectiveness of an intelligent system; it also influences how it is designed, the costs of its implementation, and the way how it is used.

3. **Intelligence. Artificial intelligence (2 hours)**
Descriptive definitions of the concepts of intelligence and artificial intelligence are provided. In relation to human intelligence, the approach proposed by H. Gardner, who perceives intelligence as a multidimensional property of the human brain/mind is examined. Artificial intelligence is defined using its definition given by J. McCarthy. A brief survey of artificial intelligence and its evolutionary milestones are presented. The role machine learning plays in artificial intelligence is discussed.

4. **Classical logic as a method of knowledge representation. Information systems in logic (5 hours)**
In the lecture, classical logic plays a special role as a method of knowledge representation, constituting a specific point of reference for other methods (benchmark). Noteworthy, information systems in logic are theories. An example of an information system in logic is given, showing that answering questions is about proving theorems.

5. **Information systems in non-classical logic (2 hours)**
R. Reiter's default logic and L. Zadeh's fuzzy logic allow one to describe typical/casual situations, though burdened with uncertainty. An example of an information system based on Reiter's logic is provided.

6. **Semantic Networks (2 hours)**
A definition of semantic networks is introduced. From a formal point of view, they are graphs whose nodes represent objects or classes of objects and edges (arches) show the relations between the objects. The fact that an inference is a "concatenation" of graph's arches is indicated as a "convenient" feature of the network, which remarkably facilitates reasoning, though being a possible trap. It is emphasised that semantic networks as a prefiguration of ontologies and to a certain extent the semantic web.

7. **Semantic atoms (2 hours)**
A definition of semantic atoms (semantic atoms, semantic primitives, semantic universals) is given. A genesis and destiny of semantic atoms i.e., the creation of meta-language(s) in which it is possible to write sentences of a given natural language in a syntactic abbreviated form and thus create intelligent knowledge based systems is presented.

8. **Frames and scripts (2 hours)**
M. Minsky's frames mainly describe the structural aspect of the world modelled in the information or a decision system, showing the objects and relations occurring in it. The scripts proposed by R. Abelson and R. Schank as an extension of the framework concept to represent the behavioural aspect of the world are discussed.

9. **Ontologies (2 hours)**
A definition of ontology is given in this module. Ontologies are presented as a tool for modelling both structural and behavioural aspects of the real world and creating information systems, including intelligent systems, in which inference can be made in relation to the objects and situations they represent. While representing domain knowledge ontologies are also platforms for

reconciliation, authorization and sharing of knowledge within the communities that make use of them.

10. **Neural networks (2 hours)**
This lecture is not conceived as a ground lecture on neural networks, its purpose is to provide general information about how knowledge is represented and manipulated in neural networks. An example of an information system using a deep neural network is provided.

11. **Natural Language Processing - NLP (2 hours)**
Core linguistics concepts and tasks of NLP (text corpora, regular expressions, tokenization, stemming, tagging, parsing; machine translation, summarisation, sentiment analysis, question-answering) are presented, and it is explained why NLP is challenging from a machine learning perspective. The language modelling and its application potential are presented and discussed.

12. **Naive Bayes Classifier (2 hours)**
Naive Bayes classifier is presented as an example of a main element of an intelligent information system that is also a decision-making system. Noteworthy, its knowledge representation method as afar as the inference mechanism is concerned uses probability. This is because the data set is an information model of the real world; and it answers questions such as: to which of the existing classes the indicated object should be classified?

Project

The lecture is accompanied by a project that includes either planning and execution of a simple experiment using a selected technique(s) discussed in the lecture or another technique(s) in the field of artificial intelligence and related fields, or creating a simple "intelligent" application based on a specific knowledge representation model. Examples of projects are conducting an experiment of factual/text data mining, development of a parser for a certain language of semantic atoms, implementation of a search mechanism in a certain class of semantic networks, development of a program for identifying user preferences based on the analysis of his queries. There are no prerequisites for the development tools to be used for carrying out the design/implementation work.

4. Learning outcomes

	Learning outcomes description	Reference to the learning outcomes of the WUT DS	Learning outcomes verification methods*
Knowledge			
K01	knows the key concepts in the field of knowledge representation (data, information, knowledge, answering questions) and artificial intelligence (reasoning, machine learning, deep learning, neural networks, natural language processing, recommender systems)	SD_W2, SD_W3, SD_W5	written exam, oral exam
K02	knows the main methods of knowledge representation used for modelling the real world and designing and implementing components of artificial intelligence applications	SD_W2, SD_W3, SD_W5	written exam, oral exam
K03	knows general principles and rules of designing knowledge representation	SD_W2, SD_W3, SD_W5	written exam, oral exam

	systems for artificial intelligence applications		
K04	knows the methods of assessing knowedge representation systems and its applications	SD_W2, SD_W3, SD_W4, SD_W5	written exam, oral exam
Skills			
S01	can analyse information needs and adjust the knowledge representation method	SD_U1, SD_U2, SD_U4, SD_U6,	project report, presentation
S02	knows how to implement knowledge representation methods or modify their existing implementations	SD_U1, SD_U2, SD_U4, SD_U6, SD_U7, SD_U8	project report, presentation
S03	can assess the quality of an intelligent system based on the effectiveness of its knowledge representation method	SD_U1, SD_U2, SD_U4, SD_U6, SD_U7, SD_U8	project report, presentation
Social competences			
SC01	uses appropriate methods of oral and written communication in the field of formulating knowledge representation based tasks	SD_K1, SD_K2, SD_K4	project
SC02	effectively cooperates in a team on research and/or knowledge representation method implementation	SD_K1, SD_K2, SD_K4	project

*Allowed learning outcomes verification methods: exam; oral exam; written test; oral test; project evaluation; report evaluation; presentation evaluation; active participation during classes; homework; tests

5. Assessment criteria

Checking the assumed learning outcomes is carried out by:
- summarizing the knowledge and skills demonstrated in the written and oral exam,
- assessment of knowledge and skills related to the implementation of project tasks assessment of the implementation and experimental work performed and the quality of documentation,
- formative assessment related to participation in consultations and an interactive form of lecture.

6. Literature

<u>Basic bibliography:</u>

1. Barr A., Cohen, P.R., Feigenbaum E. A., *The Handbook of Artificial Intelligence*, vol. I, II, III, IV Addison-Wesley, 1989.
2. Brachman R., Levesque H. (editors): *Readings in Knowledge Representation*, Morgan Kaufmann, 1985.
3. Brachman R., Levesque H.: *Knowledge Representation and Reasoning*, Morgan Kaufmann, 2004.
4. Gelfond M., Kahl Y., *Knowledge Representation, Reasoning, and the Design of Intelligent Agents: The Answer-Set Programming Approach*, Cambridge University Press, 2014.
5. van Harmelen F., Lifschitz V., Porter B. (eds), *Handbook of Knowledge Representation*, Elsevier, 2008.
6. Russel S., Norvig P.: *Artificial Intelligence. A Modern Approach*. Pearson Education Inc., 4th edition, 2020.
7. Sharda R., Delen D., *Analytics, Data Science, & Artificial Intelligence: Systems for Decision Support,* Pearson, 11th edition, 2019.

8. Sowa J.F.: *Knowledge Representation: Logical, Philosophical, and Computational Foundations*, Brooks Cole Publishing Co., Pacific Grove, 2000.

Supplementary bibliography:

9. Cichosz P.: *Systemy uczące się*. Warszawa, WNT, 2001.
10. Cichosz P.: *Data Mining Algorithms. Explained Using R*. Wiley, 2015.
11. Jakus G., Milutinovic V., Omerovic S., *Concepts, Ontologies, and Knowledge Representation*, Springer, 2013.
12. Lemos N.: *An Introduction to the Theory of Knowledge*, Cambridge University Press, 2007.
13. Muraszkiewicz M., Rybiński H.: *Bazy danych*, Wydawnictwo Akademickie, 1993.
14. Nowak R., Muraszkiewicz M. (eds). *Sztuczna inteligencja dla inżynierów*, Wydawnictwo Politechniki Warszawskiej, 2022.
15. Ullman J. D, Garcia-Molina H., Widorn J.: *Database Systems The Complete Book*, Pearson Education; 2nd edition, 2013.

Also: Weka or (other) Library(s), github, R package(s), Python package(s).

7. PhD student's workload necessary to achieve the learning outcomes**		
No.	Description	Number of hours
1	Hours of scheduled instruction given by the academic teacher in the classroom	60
2	Hours of consultations with the academic teacher, exams, tests, etc.	5
3	Amount of time devoted to the preparation for classes, preparation of presentations, reports, projects, homework	50
4	Amount of time devoted to the preparation for exams, test, assessments	5
	Total number of hours	**120**
	ECTS credits	**4**

** 1 ECTS = 25-30 hours of the PhD students work (2 ECTS = 60 hours; 4 ECTS = 110 hours, etc.)