

KARTA OPISU MODUŁU KSZTAŁCENIA		
Nazwa modułu/przedmiotu Analiza danych i sieci semantyczne dla Internetu Przedmiotów		Kod 1010512311010510003
Kierunek studiów Informatyka	Profil kształcenia (ogólnoakademicki, praktyczny) ogólnoakademicki	Rok / Semestr 1 / 1
Ścieżka obieralności/specjalność Internet Przedmiotów	Przedmiot oferowany w języku: polski	Kurs (obligatoryjny/obieralny) obligatoryjny
Stopień studiów: II stopień	Forma studiów (stacjonarna/niestacjonarna) stacjonarna	
Godziny Wykłady: 30 Ćwiczenia: - Laboratoria: 30 Projekty/seminaria: -		Liczba punktów 5
Status przedmiotu w programie studiów (podstawowy, kierunkowy, inny) inny		(ogólnouczelniany, z innego kierunku) ogólnouczelniany
Obszar(y) kształcenia i dziedzina(y) nauki i sztuki nauki techniczne nauki techniczne		Podział ECTS (liczba i %) 5 100% 5 100%
Odpowiedzialny za przedmiot / wykładowca:		
dr inż. Agnieszka Ławrynowicz email: agnieszka.lawrynowicz@cs.put.poznan.pl tel. (+48) 61 6653026 Wydział Informatyki ul. Piotrowo 3, 60-965 Poznań		dr hab. inż. Jerzy Stefanowski, prof. PP email: jerzy.stefanowski@cs.put.poznan.pl tel. (+48) 61 6652933 Wydział Informatyki ul. Piotrowo 3, 60-965 Poznań
Wymagania wstępne w zakresie wiedzy, umiejętności, kompetencji społecznych:		
1	Wiedza:	Student rozpoczynający ten przedmiot powinien posiadać podstawową wiedzę w zakresie technologii internetowych (w tym języka XML), podstaw logiki i baz danych.
2	Umiejętności:	Student powinien posiadać umiejętność rozwiązywania podstawowych problemów w zakresie projektowania systemów informatycznych i ich realizacji oraz umiejętność pozyskiwania informacji ze wskazanych źródeł. Powinien również rozumieć konieczność poszerzania swoich kompetencji.
3	Kompetencje społeczne	W zakresie kompetencji społecznych student musi prezentować takie postawy jak uczciwość, odpowiedzialność, wytrwałość, ciekawość poznawcza, kreatywność, kultura osobista, szacunek dla innych ludzi.
Cel przedmiotu:		
1. Wprowadzenie do analizy danych w zakresie uczenia nadzorowanego z przykładów, w tym strumieni danych i wielkich wolumenów danych (ang. BigData)		
2. Przekazanie studentom podstawowej wiedzy w zakresie technologii semantycznych, w tym koncepcji Web 3.0 (semantyczna sieć WWW).		
3. Rozwijanie u studentów umiejętności rozwiązywania problemów w zakresie sposobów użytkowania i projektowania systemów wykorzystujących technologie semantyczne.		
Efekty kształcenia i odniesienie do kierunkowych efektów kształcenia		
Wiedza:		
Umiejętności:		
Kompetencje społeczne:		
Sposoby sprawdzenia efektów kształcenia		

<p>Efekty kształcenia przedstawione wyżej weryfikowane są w następujący sposób:</p> <p>Ocena formująca:</p> <p>a) w zakresie wykładów: ? na podstawie odpowiedzi na pytania dotyczące materiału omówionego na poprzednich wykładach.</p> <p>b) w zakresie laboratoriów: ? na podstawie oceny bieżącego postępu realizacji zadań cząstkowych.</p> <p>Ocena podsumowująca:</p> <p>a) w zakresie wykładów weryfikowanie założonych efektów kształcenia realizowane jest przez: ? ocenę wiedzy i umiejętności podczas realizacji w grupie projektu informatycznego zawierającego elementy technologii semantycznych. ? ocenę wiedzy z zakresu analizy danych na podstawie kolokwium zaliczeniowego</p> <p>b) w zakresie laboratoriów weryfikowanie założonych efektów kształcenia realizowane jest przez: ? ocenę prezentowanego przez studenta sprawozdania z realizacji projektu w ramach publicznej prezentacji projektu.</p> <p>Uzyskiwanie punktów dodatkowych za aktywność podczas zajęć, a szczególnie za: ? omówienia dodatkowych aspektów zagadnienia, ? efektywność zastosowania zdobytej wiedzy podczas rozwiązywania zadanego problemu, ? uwagi prowadzące do udoskonalenia materiałów dydaktycznych lub procesu dydaktycznego.</p>	
<p>Treści programowe</p>	
<p>Program wykładu obejmuje następujące zagadnienia:</p> <p>Analiza danych w zakresie uczenia nadzorowanego z przykładów (cel: zrozumienie danych i generalizacji danych na nowe przypadki). Podejścia: drzewa decyzyjne, podejścia regułowe, naiwny klasyfikator Bayesowski. Strumienie danych w Internecie Przedmiotów. Analiza wielkich wolumenów danych (ang. Big Data) z Internetu Przedmiotów. Pojęcie Internetu Semantycznego. Warstwowa architektura języków Internetu Semantycznego. Klasyczne sieci semantyczne i trójkowy model danych. Język opisu zasobów RDF Reprezentacja wiedzy za pomocą ontologii. Język zapytań SPARQL. Pojęcie końcówki SPARQL. Silniki zapytań SPARQL. Modelowanie metadanych. Semantyczne systemy Wiki. Przykłady aktualnych inicjatyw związanych z semantyczną integracją danych i wiedzy (http://schema.org, Graf Wiedzy Google), ontologie i schematy metadanych dot. sieci sensorów (W3C SSN) oraz wykorzystanie technologii semantycznych w Internecie Przedmiotów.</p> <p>Program laboratorium obejmuje następujące zagadnienia:</p> <p>Analiza danych w zakresie statystyki opisowej (miary opisu, rozproszenia) oraz badanie zależności między zmiennymi liczbowymi.</p> <p>Analiza danych (uczenie nadzorowane) z wykorzystaniem narzędzia WEKA (drzewa decyzyjne, systemy regułowe, naiwny klasyfikator Bayesowski). Reprezentacja danych w modelu RDF. Modelowanie ontologii z wykorzystaniem edytora ontologii (Protégé). Modelowanie metadanych (schema.org, JSON-LD). Przetwarzanie danych semantycznych (repozytoria trójek np. Virtuoso, nierelacyjna baza dokumentów MongoDB). Odpytywanie heterogenicznych źródeł wiedzy za pomocą języka SPARQL. Tworzenie i zarządzanie treścią oraz modelowanie i wyszukiwanie wiedzy w środowisku Semantic Media Wiki.</p>	
<p>Literatura podstawowa:</p> <ol style="list-style-type: none"> 1. Systemy uczące się, Michał Skorzybut, Mirosław Krzyśko, Tomasz Górecki, Waldemar Wołyński, WNT, 2008. 2. Data Mining: Practical Machine Learning Tools and Techniques, Third Edition, Morgan Kaufmann, 2011 3. Ontologie w systemach informatycznych, Krzysztof Goczyła, EXIT, 2011 4. Eksploracja danych. Metody i algorytmy. Tadeusz Morzy, PWN, 2013. 	
<p>Literatura uzupełniająca:</p> <ol style="list-style-type: none"> 1. Semantic Web for the Working Ontologist, Dean Allemang and Jim Hendler, Morgan Kaufmann, 2008 2. Working with MediaWiki, Yaron Koren, WikiWorks Press, 2012 (aktualizacja 2014) 3. Agnieszka Ławrynowicz. 2017. Semantic data mining. An ontology-based approach. Studies on the Semantic Web, Vol. 29. IOS Pres/AKA Verlag. 	
<p>Bilans nakładu pracy przeciętnego studenta</p>	
<p>Czynność</p>	<p>Czas (godz.)</p>

1. udział w wykładach	30
2. udział w zajęciach laboratoryjnych	30
3. przygotowanie do zajęć laboratoryjnych	10
4. dokończenie (w ramach pracy własnej) ćwiczeń laboratoryjnych	10
5. udział w konsultacjach związanych z realizacją ćwiczeń laboratoryjnych i projektu (również drogą elektroniczną)	5 10
6. napisanie projektu zaliczeniowego	10
7. zapoznanie się ze wskazaną literaturą i materiałami dydaktycznymi oraz przygotowanie się do zajęć	8
8. przygotowanie prezentacji multimedialnej projektu oraz udział w końcowej prezentacji projektów	10
9. przygotowanie do kolokwium zaliczeniowego i udział w kolokwium (8+2 godz.)	2
10. omówienie wyników kolokwium	
Obciążenie pracą studenta	
forma aktywności	godzin
ECTS	
Łączny nakład pracy	125
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	69
Zajęcia o charakterze praktycznym	68