

KARTA OPISU MODUŁU KSZTAŁCENIA		
Nazwa modułu/przedmiotu Zaawansowane technologie baz danych		Kod 1010512311010519882
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: 15 Ćwiczenia: - Laboratoria: 30 Projekty/seminaria: -	Liczba punktów 3	
Status przedmiotu w programie studiów (podstawowy, kierunkowy, inny) kierunkowy		(ogólnouczelniany, z innego kierunku) z danego kierunku
Obszar(y) kształcenia i dziedzina(y) nauki i sztuki nauki techniczne nauki techniczne		Podział ECTS (liczba i %) 3 100% 3 100%
Odpowiedzialny za przedmiot / wykładowca: Odpowiedzialny za przedmiot / wykładowca: Prof. dr hab. inż. Zbyszko Królikowski dr inż. Paweł Boiński email: Zbyszko.Krolikowski@cs.put.poznan.pl email: pawel.boinski@cs.put.poznan.pl tel. 61 665 2907 tel. 6652965 Instytut Informatyki Wydział Informatyki ul. Piotrowo 2, 60-965 Poznań ul. Piotrowo 3 60-965 Poznań		
Wymagania wstępne w zakresie wiedzy, umiejętności, kompetencji społecznych:		
1	Wiedza:	Efekty kształcenia ze studiów I stopnia zdefiniowane w Uchwale Senatu PP ? efekty te prezentowane są w serwisie internetowym wydziału www.fc.put.poznan.pl Student rozpoczynający ten przedmiot powinien posiadać podstawową wiedzę z systemów baz danych i języków programowania.
2	Umiejętności:	Powinien posiadać umiejętność rozwiązywania podstawowych problemów z projektowania systemów informatycznych, administrowania systemami baz danych, formułowania poleceń w języku SQL oraz umiejętność pozyskiwania informacji ze wskazanych źródeł.
3	Kompetencje społeczne	Powinien również rozumieć konieczność poszerzania swoich kompetencji / mieć gotowość do podjęcia współpracy w zespole. Ponadto w zakresie kompetencji społecznych student musi prezentować takie postawy jak uczciwość, odpowiedzialność, wytrwałość, ciekawość poznawcza, kreatywność, kultura osobista, szacunek dla innych ludzi.
Cel przedmiotu:		
Przekazanie studentom wiedzy dotyczącej projektowania i implementacji hurtowni danych i aplikacji klasy BI, w zakresie: architektury, modelowania danych, projektowania warstwy integrującej i zasilającej - ETL, rozszerzeń SQL dla aplikacji klasy BI, struktur fizycznych i indeksowych hurtowni danych, trendów rozwojowych hurtowni danych i systemów BI. Przekazanie studentom podstawowej wiedzy dotyczącej baz NoSQL.		
Rozwijanie u studentów umiejętności rozwiązywania problemów, w zakresie: projektowania i implementowania systemu hurtowni danych, oceny przydatności technologii HD i BI oraz eksploracji danych do konkretnego zastosowania.		
Kształtowanie u studentów umiejętności pracy zespołowej w ramach budowy systemów hurtowni danych. Kształtowanie u studentów umiejętności realizowania projektów HD i BI oraz eksploracji danych. Kształtowanie u studentów umiejętności korzystania ze źródeł wiedzy (np. publikacje anglojęzyczne) i samokształcenia się oraz zarządzania czasem.		
Efekty kształcenia i odniesienie do kierunkowych efektów kształcenia		
Wiedza:		
1. ma zaawansowaną i pogłębioną wiedzę z zakresu hurtowni danych oraz baz NoSQL, podstaw teoretycznych ich budowania oraz metod, narzędzi i środowisk programistycznych wykorzystywanych do ich implementacji - [K2st_W1]		
2. ma zaawansowaną wiedzę szczegółową dotyczącą hurtowni danych i baz NoSQL - [K2st_W3]		
3. ma wiedzę o trendach rozwojowych baz NoSQL - [K2st_W4]		
4. ma zaawansowaną i szczegółową wiedzę o procesach zachodzących w cyklu życia hurtowni danych - [K2st_W5]		
5. zna zaawansowane metody, techniki i narzędzia stosowane przy rozwiązywaniu złożonych zadań inżynierskich i prowadzeniu prac badawczych w zakresie hurtowni danych i analizy danych - [K2st_W6]		
Umiejętności:		

<p>1. potrafi pozyskiwać informacje na temat zaawansowanych systemów baz danych z literatury oraz innych źródeł (w języku ojczystym i angielskim), integrować je, dokonywać ich interpretacji i krytycznej oceny - [K2st_U1]</p> <p>2. potrafi wykorzystać do formułowania i rozwiązywania zadań inżynierskich i problemów badawczych metody eksperymentalne - [K2st_U4]</p> <p>3. potrafi - przy formułowaniu i rozwiązywaniu zadań inżynierskich z zakresu zaawansowanych systemów baz danych - integrować wiedzę z różnych obszarów informatyki (np. inżynieria oprogramowania, administrowanie systemami informatycznymi, bazy danych) oraz zastosować podejście systemowe - [K2st_U5]</p> <p>4. potrafi ocenić przydatność i możliwość wykorzystania nowych osiągnięć (metod i narzędzi) oraz nowych produktów informatycznych (w szczególności systemów zarządzania bazami danych, środowisk programistycznych ETL) - [K2st_U6]</p> <p>5. potrafi określić kierunki dalszego uczenia się i zrealizować proces samokształcenia (m.in. wykorzystując dostępne przewodniki po narzędziach ETL) - [K2st_U16]</p>
Kompetencje społeczne:
<p>1. rozumie, że w informatyce wiedza i umiejętności z zakresu ZSBD bardzo szybko stają się przestarzałe - [K2st_K1]</p> <p>2. rozumie znaczenie wykorzystywania najnowszej wiedzy z zakresu ZSBD w rozwiązywaniu problemów badawczych i praktycznych (m.in. dobór odpowiednich narzędzi i metod dla zasilania i odświeżania hurtowni danych) - [K2st_K2]</p>

<p style="text-align: center;">Sposoby sprawdzenia efektów kształcenia</p> <p>W zakresie wykładów weryfikowanie założonych efektów kształcenia realizowane jest przez ocenę wiedzy i umiejętności wykazanych na kolokwium zaliczeniowym z pytaniami (5-7) problemowymi i otwartymi. Kolokwium uznaje się za zaliczony po uzyskaniu ponad 50% możliwych do zdobycia punktów. Przyjmuje się następującą skalę ocen i punktów:</p> <p><0;50%>: ndst., (50%;60%>: dst, (60%;70%>: dst+, (70%;80%>: db, (80%;90%>: db+, (90%;100%>: bdb</p> <p>W zakresie laboratoriów / ćwiczeń weryfikowanie założonych efektów kształcenia realizowane jest przez:</p> <ul style="list-style-type: none">- ocenę realizacji zadań zleczanych na każdych zajęciach,- ocenę wiedzy i umiejętności związanych z realizacją zadań laboratoryjnych poprzez rozwiązanie sprawdzianu otwartego (w trakcie semestru) oraz kolokwium (w formie testu) na koniec semestru.- uzyskiwanie punktów dodatkowych za aktywność podczas zajęć, a szczególnie za:- omówienia dodatkowych aspektów zagadnienia,- uwagi związane z udoskonaleniem materiałów dydaktycznych. <p>W zakresie laboratorium przyjmuje się następującą skalę ocen w zależności od liczby uzyskanych punktów: <0;50%>: ndst., (50%;60%>: dst, (60%;70%>: dst+, (70%;80%>: db, (80%;90%>: db+, (90%;100%>: bdb.</p>
<p style="text-align: center;">Treści programowe</p> <p>Program wykładów obejmuje następujące zagadnienia:</p> <ul style="list-style-type: none">- problematyka przetwarzania BigData, w tym: BI;- architektury systemu hurtowni danych (podstawowa ETL, z warstwą ODS, z warstwą data mart, architektura ELT, BigData),- modelowanie danych (konceptualny model wielowymiarowy, implementacja relacyjna - schematy gwiazdy, płotka śniegu, konstelacji faktów wraz z ich oceną, implementacja wielowymiarowa),- struktury fizyczne plików hurtowni danych oraz struktury indeksowe,- bazy NoSQL do zastosowań web-owych i chmurowych (Dlaczego NoSQL ? Bazy klucz ? wartość, bazy dokumentów, bazy grafowe, ?) <p>Program zajęć laboratoryjnych podzielono na następujące części:</p> <ol style="list-style-type: none">1. Wprowadzenie do środowiska ćwiczeniowego<ul style="list-style-type: none">- studium przypadku,- źródła danych, schemat hurtowni danych,- podstawy metodyki Agile BI.2. Wprowadzenie do obsługi narzędzia Pentaho Data Integration<ul style="list-style-type: none">- podstawowe pojęcia,- repozytorium,- transformacja oparta na jednym źródle danych,- transformacja podrzędna.3. Obsługa wielu źródeł danych<ul style="list-style-type: none">- rozbudowa istniejących transformacji i transformacji podrzędnych o dodatkowe źródło danych,- sterowanie ścieżką przepływu danych,- metody łączenia danych.4. Dodatkowe transformacje<ul style="list-style-type: none">- metody eliminowania duplikatów,- automatyczne generowanie danych dla wymiarów,

<ul style="list-style-type: none"> - zasilanie tabeli faktów. - podstawy metodyki Agile BI. <p>5. Zaawansowane transformacje</p> <ul style="list-style-type: none"> - źródła danych oparte na plikach CSV, wykrywanie zmian w źródłach danych, - operacyjna składnica danych, odświeżanie hurtowni danych. <p>6. Nowoczesne źródła danych</p> <ul style="list-style-type: none"> - dokumenty XML, usługi sieciowe. <p>7. Profilowanie i czyszczenia danych, dane historyczne</p> <ul style="list-style-type: none"> - wykrywanie błędów w danych (dane referencyjne, wzorce danych), - automatyczne poprawianie błędów, naprawianie błędów w źródłach danych, - modyfikacja transformacji w celu przechowywania danych historycznych dla zmieniających się wymiarów. <p>8. Poprawa wydajności procesu ETL, tematyczne hurtownie danych</p> <ul style="list-style-type: none"> - masowe ładowanie danych (Oracle, PostgreSQL, MySQL) - wyliczanie agregatów z danych, przykład tematycznej hurtowni danych. <p>9. Przetwarzanie danych w hurtowniach danych za pomocą języka SQL i jego rozszerzeń.</p> <p>Zajęcia są prowadzone w formie zajęć ćwiczeniowych przy komputerach, przy czym każdy student pracuje samodzielnie. Każde zadanie jest poprzedzone krótką prezentacją a następnie omówione zagadnienia są ćwiczone w praktyce.</p> <p>Metody dydaktyczne:</p> <ol style="list-style-type: none"> 1. wykład: prezentacja multimedialna, prezentacja jest uzupełniana krótkimi przykładami prezentowanymi w sposób tradycyjny z wykorzystaniem tablicy. 2. laboratoria: prezentacja multimedialna, prezentacja jest uzupełniana krótkimi przykładami prezentowanymi w sposób tradycyjny z wykorzystaniem tablicy, wykonywanie ćwiczeń w hurtowni danych, omawianie trudniejszych ćwiczeń przy tablicy, odpowiedzi na pytania na bieżąco, rozwiązywanie problemów na bieżąco 		
<p>Literatura podstawowa:</p> <ol style="list-style-type: none"> 1. Z.Królikowski, Hurtownie danych ? Struktury logiczne i fizyczne, Wydawnictwo Naukowe PP, 2008 2. A.Chodkowska-Gyurics, Hurtownie danych, Teoria i praktyka, PWN 2014 3. P.J.Sadalage, M.Flower, ?NoSQL - Kompendium wiedzy?, Helion, 2015 		
<p>Literatura uzupełniająca:</p> <ol style="list-style-type: none"> 1. Jiang B.: Constructing Data Warehouses with Metadata-driven Generic Operators, and more: Architecture, Methodology, and Paradigm; Concepts, Algorithms, and Operators; Principles, Recommendations, and Exercises. DBJ Publishing, 2011, ISBN-13: 978-3033029200 2. Dokumentacja Pentaho Data Integration https://wiki.pentaho.com/ 3. Matt Casters, Roland Bouman, Jos Van Dongen: Pentaho Kettle Solutions, John Wiley & Sons 2010 4. A.Pelikant, Hurtownie danych ? od przetwarzania analitycznego do raportowania, Wyd. Helion, 2011 		
Bilans nakładu pracy przeciętnego studenta		
Czynność	Czas (godz.)	
1. udział w zajęciach laboratoryjnych	30	
2. realizacja zadań (poza czasem laboratoryjnym):	10	
3. udział w konsultacjach (mogą być realizowane drogą elektroniczną) związanych z realizacją procesu kształcenia, w szczególności projektu	2	
4. przygotowanie do sprawdzianów	8	
5. udział w wykładach	15	
6. zapoznanie się ze wskazaną literaturą / materiałami dydaktycznymi w czasie realizacji projektu (10 stron tekstu naukowego = 1 godz.), 100 stron	10	
7. przygotowanie do kolokwium zaliczającego wykład	5	
Obciążenie pracą studenta		
forma aktywności	godzin	ECTS
Łączny nakład pracy	80	3
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	47	2
Zajęcia o charakterze praktycznym	40	2