

<b>KARTA OPISU MODUŁU KSZTAŁCENIA</b>		
Nazwa modułu/przedmiotu <b>Zaawansowane technologie baz danych</b>		Kod <b>1010515311010519882</b>
Kierunek studiów <b>Informatyka</b>	Profil kształcenia (ogólnoakademicki, praktyczny) <b>ogólnoakademicki</b>	Rok / Semestr <b>1 / 1</b>
Ścieżka obieralności/specjalność <b>Zaawansowane technologie internetowe</b>	Przedmiot oferowany w języku: <b>polski</b>	Kurs (obligatoryjny/obieralny) <b>obligatoryjny</b>
Stopień studiów: <b>II stopień</b>	Forma studiów (stacjonarna/niestacjonarna) <b>niestacjonarna</b>	
Godziny Wykłady: <b>12</b> Ćwiczenia: - Laboratoria: <b>16</b> Projekty/seminaria: -		Liczba punktów <b>3</b>
Status przedmiotu w programie studiów (podstawowy, kierunkowy, inny) <b>kierunkowy</b>		(ogólnouczelniany, z innego kierunku) <b>z danego kierunku</b>
Obszar(y) kształcenia i dziedzina(y) nauki i sztuki <b>nauki techniczne</b> <b>nauki techniczne</b>		Podział ECTS (liczba i %) <b>3 100%</b> <b>3 100%</b>
<b>Odpowiedzialny za przedmiot / wykładowca:</b>		
Prof. dr hab. inż. Zbyszko Królikowski email: Zbyszko.Krolikowski@cs.put.poznan.pl tel. 61 6652907 Instytut Informatyki ul. Piotrowo 2, 60-965 Poznań		dr inż. Paweł Boiński email: Pawel.Boinski@cs.put.poznan.pl tel. 61 6652965 Wydział Informatyki ul. Piotrowo 3 60-965 Poznań
<b>Wymagania wstępne w zakresie wiedzy, umiejętności, kompetencji społecznych:</b>		
1	<b>Wiedza:</b>	Efekty kształcenia ze studiów I stopnia zdefiniowane w Uchwale Senatu PP, a szczególnie efekty K_W1-2, K_W4, K_W6-15, K_U1-2, K_U4, K_U7-8, K_U14-20, K_U22-23, K_U26, K_K1-9, weryfikowane w procesie rekrutacji na studia 2 stopnia ? 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	<b>Umiejętności:</b>	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	<b>Kompetencje społeczne</b>	Powinien również rozumieć konieczność poszerzania swoich kompetencji / mieć gotowość do podjęcia współpracy zespołowej. Ponadto w zakresie kompetencji społecznych student musi prezentować takie postawy jak uczciwość, odpowiedzialność, wytrwałość, ciekawość poznawcza, kreatywność, kultura osobista, szacunek dla innych ludzi.
<b>Cel przedmiotu:</b>		
<p>Repetitorium z najważniejszych problemów projektowania i implementacji baz danych. 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, trendów rozwojowych hurtowni danych i systemów BI.</p> <p>Przekazanie studentom wiedzy dotyczącej podstaw eksploracji danych.</p> <p>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.</p> <p>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.</p>		
<b>Efekty kształcenia i odniesienie do kierunkowych efektów kształcenia</b>		
<b>Wiedza:</b>		

<p>1. ma uporządkowaną, podbudowaną teoretycznie wiedzę ogólną w zakresie architektury systemów informatycznych, w szczególności baz danych, hurtowni danych i eksploracji danych. - [K2st_W2]</p> <p>2. ma podbudowaną teoretycznie szczegółową wiedzę związaną z wybranymi zagadnieniami z zakresu informatyki, takimi jak: architektury i technologie integracji danych, metodyka projektowania systemu informatycznego (w szczególności hurtowni danych), technologie analizy i eksploracji danych, modele składowania danych. - [K2st_W3]</p> <p>3. ma wiedzę o trendach rozwojowych i najistotniejszych nowych osiągnięciach w dziedzinie przetwarzania danych i hurtowni danych - [K2st_W4]</p> <p>4. ma podstawową wiedzę o cyklu życia systemów informatycznych, w szczególności systemów hurtowni danych - [K2st_W5]</p> <p>5. zna podstawowe metody, techniki i narzędzia stosowane przy rozwiązywaniu złożonych zadań inżynierskich z wybranego obszaru informatyki (projektowania i implementowania systemu hurtowni danych, projektowania i implementowania schematów hurtowni danych) - [K2st_W6]</p>
<p><b>Umiejętności:</b></p> <p>1. potrafi pozyskiwać informacje z literatury, baz danych oraz innych źródeł (w języku ojczystym i angielskim), integrować je, dokonywać ich interpretacji i krytycznej oceny, wyciągać wnioski oraz formułować i wyczerpująco uzasadniać opinie (także przygotowywać prezentacje i raporty techniczne) - [K2st_U1]</p> <p>2. potrafi wykorzystywać do formułowania i rozwiązywania zadań inżynierskich i prostych problemów badawczych metody eksperymentalne - [K2st_U4]</p> <p>3. 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>4. potrafi - przy formułowaniu i rozwiązywaniu zadań inżynierskich - integrować wiedzę z różnych obszarów informatyki (np. inżynieria oprogramowania, administrowanie systemami informatycznymi, bazy danych, technologie internetowe) oraz zastosować podejście systemowe, uwzględniające także aspekty pozatechniczne - [K2st_U5]</p> <p>5. potrafi określić kierunki dalszego uczenia się i zrealizować proces samokształcenia (m.in. wykorzystując dostępne przewodniki po narzędziach ETL) - [K2st_U16]</p>
<p><b>Kompetencje społeczne:</b></p> <p>1. rozumie, że w informatyce wiedza i umiejętności bardzo szybko stają się przestarzałe - [K2st_K1]</p> <p>2. rozumie znaczenie wykorzystywania najnowszej wiedzy z zakresu informatyki w rozwiązywaniu problemów badawczych i praktycznych (m.in. dobór odpowiednich narzędzi i metod dla zasilania i odświeżania hurtowni danych) - [K2st_K2]</p>

<p><b>Sposoby sprawdzenia efektów kształcenia</b></p>
<p>Efekty kształcenia przedstawione wyżej weryfikowane są w następujący sposób:</p> <p>Ocena formująca:</p> <p>a) w zakresie wykładów:</p> <ul style="list-style-type: none"> <li>- na podstawie odpowiedzi na pytania dotyczące materiału omówionego na wykładach,</li> </ul> <p>b) w zakresie laboratoriów:</p> <ul style="list-style-type: none"> <li>- na podstawie oceny bieżącego postępu realizacji zadań,</li> </ul> <p>Ocena podsumowująca:</p> <p>a) w zakresie wykładów weryfikowanie założonych efektów kształcenia realizowane jest przez:</p> <ul style="list-style-type: none"> <li>- ocenę wiedzy i umiejętności wykazanych w formie kolokwium zaliczeniowego z pytaniami otwartymi.</li> </ul> <p>Kolokwium składa się z szeregu pytań (4-6) otwartych. Kolokwium uznaje się za zaliczone po uzyskaniu ponad 50% możliwych do zdobycia punktów. Przyjmuje się następującą skalę ocen i punktów:</p> <p>&lt;0%;50%&gt;: ndst., (50%;60%&gt;: dst, (60%;70%&gt;: dst+, (70%;80%&gt;: db, (80%;90%&gt;: db+, (90%;100%&gt;: bdb</p> <ul style="list-style-type: none"> <li>- omówienie wyników egzaminu;</li> </ul> <p>b) w zakresie laboratoriów / ćwiczeń weryfikowanie założonych efektów kształcenia realizowane jest przez:</p> <ul style="list-style-type: none"> <li>- ocenę realizacji zadań zleczanych na każdych zajęciach,</li> <li>- ocenę wiedzy i umiejętności związanych z realizacją zadań laboratoryjnych poprzez rozwiązanie jednego kolokwium (w formie testu) na koniec semestru.</li> <li>- uzyskiwanie punktów dodatkowych za aktywność podczas zajęć, a szczególnie za:</li> <li>- omówienia dodatkowych aspektów zagadnienia,</li> <li>- uwagi związane z udoskonaleniem materiałów dydaktycznych.</li> </ul> <p>W zakresie laboratorium przyjmuje się następującą skalę ocen w zależności od liczby uzyskanych punktów: &lt;0%;50%&gt;: ndst., (50%;60%&gt;: dst, (60%;70%&gt;: dst+, (70%;80%&gt;: db, (80%;90%&gt;: db+, (90%;100%&gt;: bdb.</p>
<p><b>Treści programowe</b></p>
<p>Program wykładów obejmuje następujące zagadnienia:</p> <ul style="list-style-type: none"> <li>- repetytorium z baz danych, w tym: podstawy relacyjnych baz danych, projektowanie baz danych,</li> <li>- architektury systemu hurtowni danych (podstawowa ETL, z warstwą ODS, z warstwą data mart, architektura ELT, BigData),</li> <li>- modelowanie danych (konceptualny model wielowymiarowy, implementacja relacyjna - schematy gwiazdy, płątka śniegu, konstelacji faktów wraz z ich oceną, implementacja wielowymiarowa),</li> </ul>

- problematyka integracji danych dla hurtowni danych oraz projektowanie warstwy ETL,
- podstawy eksploracji danych.

Zajęcia laboratoryjne prowadzone są w formie czterech 4-godzinnych zajęć, odbywających się w laboratorium. Program zajęć podzielono na osiem części:

1. Wprowadzenie do środowiska ćwiczeniowego
  - studium przypadku,
  - źródła danych,
  - schemat hurtowni danych,
  - podstawy metodyki Agile BI.
2. Wprowadzenie do obsługi narzędzia Pentaho Data Integration
  - podstawowe pojęcia,
  - repozytorium,
  - transformacja oparta na jednym źródle danych,
  - transformacja podrzędna.
3. Obsługa wielu źródeł danych
  - rozbudowa istniejących transformacji i transformacji podrzędnych o dodatkowe źródło danych,
  - sterowanie ścieżką przepływu danych,
  - metody łączenia danych.
4. Dodatkowe transformacje
  - metody eliminowania duplikatów,
  - automatyczne generowanie danych dla wymiarów,
  - zasilanie tabeli faktów.
  - podstawy metodyki Agile BI.
5. Zaawansowane transformacje
  - źródła danych oparte na plikach CSV,
  - wykrywanie zmian w źródłach danych,
  - operacyjna składnica danych,
  - odświeżanie hurtowni danych.
6. Nowoczesne źródła danych
  - dokumenty XML,
  - usługi sieciowe.
7. Profilowanie i czyszczenia danych, dane historyczne
  - 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.
8. Poprawa wydajności procesu ETL, tematyczne hurtownie danych
  - masowe ładowanie danych (Oracle, PostgreSQL, MySQL)
  - wyliczanie agregatów z danych,
  - przykład tematycznej hurtowni danych.

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.

Metody dydaktyczne:

1. wykład: prezentacja multimedialna, prezentacja jest uzupełniana krótkimi przykładami prezentowanymi w sposób tradycyjny z wykorzystaniem tablicy, rozwiązywanie problemów i omawianie rozwiązań na 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.

#### Literatura podstawowa:

1. R. Kimball, J. Caserta: The Data Warehouse ETL Toolkit. Wiley, 2004, ISBN 0-7645-6757-8-54500
2. Golfarelli M., Rizzi S.: Data Warehouse Design: Modern Principles and Methodologies. McGraw-Hill Osborne, 2009, ISBN-13: 978-0071610391

<b>Literatura uzupełniająca:</b>		
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 <a href="https://wiki.pentaho.com/">https://wiki.pentaho.com/</a>		
3. Matt Casters, Roland Bouman, Jos Van Dongen: Pentaho Kettle Solutions, John Wiley & Sons 2010		
<b>Bilans nakładu pracy przeciętnego studenta</b>		
<b>Czynność</b>	<b>Czas (godz.)</b>	
1. udział w zajęciach laboratoryjnych	16	
2. samodzielne dokończenie ćwiczeń laboratoryjnych	12	
3. udział w konsultacjach (mogą być realizowane drogą elektroniczną) związanych z realizacją procesu kształcenia, w szczególności projektu	4	
4. przygotowanie do sprawdzianów	8	
5. udział w wykładach	16	
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 zaliczeniowego wykład	1	
8. omówienie wyników kolokwium		
<b>Obciążenie pracą studenta</b>		
<b>forma aktywności</b>	<b>godzin</b>	<b>ECTS</b>
Łączny nakład pracy	77	3
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	37	1
Zajęcia o charakterze praktycznym	28	1