

<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>Umiejętności:</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>Kompetencje społeczne:</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>Sposoby sprawdzenia efektów kształcenia</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"> - na podstawie odpowiedzi na pytania dotyczące materiału omówionego na wykładach, <p>b) w zakresie laboratoriów:</p> <ul style="list-style-type: none"> - na podstawie oceny bieżącego postępu realizacji zadań, <p>Ocena podsumowująca:</p> <p>a) w zakresie wykładów weryfikowanie założonych efektów kształcenia realizowane jest przez:</p> <ul style="list-style-type: none"> - ocenę wiedzy i umiejętności wykazanych w formie kolokwium zaliczeniowego z pytaniami otwartymi. <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><0;50%>: ndst., (50%;60%>: dst, (60%;70%>: dst+, (70%;80%>: db, (80%;90%>: db+, (90%;100%>: bdb</p> <ul style="list-style-type: none"> - omówienie wyników egzaminu; <p>b) w zakresie laboratoriów / ćwiczeń weryfikowanie założonych efektów kształcenia realizowane jest przez:</p> <ul style="list-style-type: none"> - ocenę realizacji zadań zleczanych na każdym zajęciach, - ocenę wiedzy i umiejętności związanych z realizacją zadań laboratoryjnych poprzez rozwiązanie jednego 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>Treści programowe</p>
<p>Program wykładów obejmuje następujące zagadnienia:</p> <ul style="list-style-type: none"> - repetytorium z baz danych, w tym: podstawy relacyjnych baz danych, projektowanie baz danych, - 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, płątka śniegu, konstelacji faktów wraz z ich oceną, implementacja wielowymiarowa),

- 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

Literatura uzupełniająca:		
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		
Bilans nakładu pracy przeciętnego studenta		
Czynność	Czas (godz.)	
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		
Obciążenie pracą studenta		
forma aktywności	godzin	ECTS
Łączny nakład pracy	77	4
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	37	2
Zajęcia o charakterze praktycznym	28	1