

KARTA PRZEDMIOTU

Nazwa przedmiotu: Hurtownie danych i systemy eksploracji danych (InfAu>SI6-HDSED-19)

Nazwa w języku polskim:

Nazwa w jęz. angielskim: Data Warehouses and Data Mining Systems

Dane dotyczące przedmiotu:

Jednostka oferująca przedmiot: Wydział Automatyki, Elektroniki i Informatyki

Przedmiot dla jednostki: Politechnika Śląska

Domyślny typ protokołu dla przedmiotu:

ZAL

Język wykładowy:

polski

Strona WWW:

<https://platforma2.polsl.pl/rau2/course/view.php?id=502>

Skrócony opis:

Celem wykładu HDiSED jest przekazanie studentom podstawowych wiadomości w zakresie koncepcji, modeli i projektowania klasycznych systemów hurtowni danych. Celem projektu jest nabycie przez studentów umiejętności projektowania hurtowni danych w różnych środowiskach ze szczególnym uwzględnieniem metodyki kursu specjalistycznego ITA-102 - hurtownie danych. Dzięki uczestnictwu w wykładzie i projekcie studenci praktycznie weryfikują umiejętność projektowania klasycznych hurtowni danych i nabywają umiejętności użycia metod eksploracji danych.

Opis:

Treści programowe:

Wykład:

1. Geneza hurtowni danych (HD) (ang. Data Warehouse) i systemów eksploracji danych (SED) (ang. Data Mining Systems).
2. Procesy ekstrakcji danych (ETL) (procesu ETL/M/R, projektowanie i graficzne modelowanie ekstrakcji danych, specjalizowane i uniwersalne systemy ETL).
3. Odtwarzanie i uaktualnianie danych (upadek procesów ETL, tradycyjne algorytmy odtwarzania, algorytmy Design-Resume (DR/JavaBeans) i ich złożoność i opłacalność).
4. Ekstrakcja i integracja danych w czasie rzeczywistym (systemy rozwojowe ETL, technika stagingu, odtwarzanie hybrydowe).
5. Modelowanie danych i przetwarzanie analityczne (model relacyjny a wielowymiarowy, modele przetwarzania analitycznego w trybie on-line (OLAP), wielowymiarowe operacje i schematy danych, klasy i architektury OLAP – analiza porównawcza).
6. Projektowanie relacyjnej architektury OLAP (ROLAP) (model koncepcyjny, logiczny, diagramy ERD, schematy gwiazdy i płatka śniegu, złożone hierarchie, cechy, zmienność wymiarów, agregacja danych).
7. Metodyka projektowania DSS klasy ROLAP (instalacje pilotażowe, wielowarstwowa infrastruktura techniczna, wybrane aspekty DSS w zakresie analizy: ex1) rynku papierosów, ex2) rynku producenta AEG, ex3) zasobów HR/DSS w elektrowni).
8. Model fizyczny architektury ROLAP (indeksy bitmapowe i złączeniowe, metody indeksowania DW Oracle 10g/11g, UB drzewo, aR-drzewo, partycjonowanie danych).
9. Perspektywy materializowane w architekturze ROLAP (SQL'owa klauzula GROUP BY, zarządzanie perspektywami, optymalizacja modelu fizycznego).
10. Przetwarzanie równoległe architektur ROLAP (typu: SE, SN, SD; klaster Beowulfa, paskowanie architektury ROLAP/SN, analizy wydajnościowe).
11. Analiza porównawcza rynkowych systemów rozwojowych OLAP.
12. Bezpieczeństwo danych w architekturze ROLAP (autentykacja użytkowników i źródeł danych, bezpieczeństwo transferu danych, kontrola dostępu do danych, ocena wpływu mechanizmów kontroli dostępu i bezpieczeństwa danych na wydajność HD).
13. Serwery ROLAP (wersjonowanie bazy danych, Red Brick Decision Server, DW Oracle 11g, silnik ETL, silnik OLAP, zarządzanie indeksami, optymalizacja zapytań, środowisko rozwojowe, silnik eksploracji danych, silnik wizualizacji)
14. Kierunki rozwoju hurtowni danych.

Projekt:

Celem projektów jest przygotowanie studentów do realizacji projektów inżynierskich związanych z klasycznymi hurtowniami danych i systemami eksploracji danych oraz zaprezentowanie szczegółowej wiedzy podstawowej z zakresu ich inżynierii. Projekt obejmuje 2 części:

1. Kurs specjalistyczny „Hurtownie danych” (w ramach programu Microsoft IT Academy dla nauczycieli akademickich wg książki pt. ITA-102 - hurtownie danych).
2. Wskazane projekty zespołowe - lista projektów do wyboru.

Liczba punktów ECTS: 3

Całkowita liczba godzin: 90 (kontaktowo 65h / praca własna 25h)

Wykład: 30h

Projekt: 30h

Inne (zaliczenie wykładu i projektu): 5h

Praca własna studenta: przygotowanie do zaliczenia wykładu (5h), przygotowanie projektu (15h), przygotowanie prezentacji projektu (5h)

w tym liczba punktów ECTS uzyskanych w ramach zajęć prowadzonych z bezpośrednim udziałem nauczycieli akademickich lub innych osób prowadzących zajęcia i studentów: 2,17.

Literatura:

1. Publikacje Zespołu Teorii Przestrzeni Danych i Algorytmów (dostępne w Bibliotece Głównej Politechniki Śląskiej: <https://www.bg.polsl.pl/expertus/new/bib/>).
2. Kurs specjalistyczny „Hurtownie danych” w ramach programu Microsoft IT Academy dla nauczycieli akademickich wg książki pt. ITA-102 - hurtownie danych autorstwa: Gorawski M., Gorawski M.J., Bańkowski. S. Wydawnictwo: Microsoft IT Academy Publication, 2008, 198 s.

Efekty uczenia się:

Wiedza:

K1A_W09 Posiada elementarną wiedzę z zakresu metod projektowania urządzeń cyfrowych w podstawowych technologiach (w tym

USOS: Szczegóły przedmiotu: InfAu>SI6-HDSED-19, w cyklu: <brak>, jednostka dawcy: <brak>, grupa przedm.: <brak>

programowalnych) oraz ich oddziaływania na otoczenie.

K1A_W10 Posiada elementarną wiedzę na temat systemów informatycznych czasu rzeczywistego, zasad ich projektowania i programowania.

K1A_W12 Ma teoretyczną wiedzę ogólną w zakresie: architektury systemów komputerowych, systemów operacyjnych, sieci komputerowych i technologii sieciowych, systemów wbudowanych oraz projektowania i implementacji prostych systemów komputerowych.

K1A_W22 Ma podstawową wiedzę dotyczącą zarządzania, w tym zarządzania jakością, i prowadzenia działalności gospodarczej.

Umiejętności:

K1A_U27 Jest przygotowany do efektywnego uczestniczenia w inspekcji oprogramowania.

Metody i kryteria oceniania:

Wykład:

Praca zaliczeniowa w formie prezentacji na temat wybranego zagadnienia poruszanego w toku wykładów.

Kryterium zaliczenia: terminowe dostarczenie pracy zaliczeniowej oraz jej pozytywna ocena.

Projekt:

Zaliczenie pisemne w formie pracy projektowej: sprawozdanie z projektu wstępnego oraz sprawozdanie z projektu zasadniczego.

Kryterium zaliczenia: dostarczenie i zaprezentowanie sprawozdania z pracy projektowej zgodnie z określonymi wymaganiami zadania projektowego.

Sylabus obowiązuje od roku akademickiego 2025/2026, a jego zawartość nie podlega zmianom w trakcie trwania semestru.

Punkty przedmiotu w cyklach:

<bez przypisanego programu>			
Typ punktów	Liczba	Cykl pocz.	Cykl kon.
Europejski System Transferu Punktów (ECTS)	3	2021/2022-L	