

<b>KARTA OPISU MODUŁU KSZTAŁCENIA</b>		
Nazwa modułu/przedmiotu <b>Zaawansowane technologie przetwarzania danych</b>		Kod <b>1010512321010510163</b>
Kierunek studiów <b>Informatyka</b>	Profil kształcenia (ogólnoakademicki, praktyczny) <b>ogólnoakademicki</b>	Rok / Semestr <b>1 / 2</b>
Ścieżka obieralności/specjalność <b>Technologie przetwarzania danych</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>stacjonarna</b>	
Godziny Wykłady: <b>30</b> Ćwiczenia: - Laboratoria: <b>30</b> Projekty/seminaria: -		Liczba punktów <b>5</b>
Status przedmiotu w programie studiów (podstawowy, kierunkowy, inny) (ogólnouczelniany, z innego kierunku) <b>kierunkowy 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>5 100%</b> <b>5 100%</b>
<b>Odpowiedzialny za przedmiot / wykładowca:</b>		
dr inż. Tomasz Koszłajda email: Tomasz.Koszłajda@cs.put.poznan.pl tel. 61 6652960 Instytut Informatyki ul. Piotrowo 2, 60-965 Poznań		dr hab. inż. Marek Wojciechowski email: Marek.Wojciechowski@cs.put.poznan.pl tel. 61 6652962 Instytut Informatyki ul. Piotrowo 2, 60-965 Poznań
<b>Wymagania wstępne w zakresie wiedzy, umiejętności, kompetencji społecznych:</b>		
1	<b>Wiedza:</b>	Student rozpoczynający ten przedmiot powinien posiadać podstawową wiedzę z dziedziny systemów baz danych oraz obiektowego paradygmatu programowania.
2	<b>Umiejętności:</b>	Powinien posiadać umiejętność rozwiązywania podstawowych problemów występujących w projektowaniu i tworzeniu programów komputerowych, instalacji, konfigurowania i strojenia oprogramowania systemowego oraz umiejętność pozyskiwania informacji ze wskazanych źródeł. Powinien również rozumieć konieczność poszerzania swoich kompetencji i mieć gotowość do podjęcia współpracy w ramach zespołu.
3	<b>Kompetencje społeczne</b>	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>		
1. Przekazanie studentom podstawowej wiedzy dotyczącej składowania i przetwarzania zbiorów danych o specyficznych typach danych: obrazy, dokumenty tekstowe, dane przestrzenne (GIS) i strumienie danych. 2. Rozwijanie u studentów umiejętności rozwiązywania problemów związanych z modelowaniem, projektowaniem i konstrukcją programów przetwarzających duże i współdzielone repozytoria powyższych typów danych.		
<b>Efekty kształcenia i odniesienie do kierunkowych efektów kształcenia</b>		
<b>Wiedza:</b>		
1. ma zaawansowaną i pogłębioną wiedzę z zakresu systemów informatycznych zarządzających danymi przestrzennymi, obrazami, dokumentami tekstowymi i strumieniami danych; i - [K2st_W1] 2. ma zaawansowaną wiedzę szczegółową dotyczącą obiektowo-relacyjnego modelu danych; metod wyszukiwania obrazów na podstawie ich zawartości (CBIR), globalnych i lokalnych deskryptorów obrazów; baz dokumentów tekstowych, systemów GIS i systemów strumieni danych; - [K2st_W3] 3. ma wiedzę o trendach rozwojowych i najistotniejszych nowych osiągnięciach w zakresie technik przetwarzania danych tekstowych, przestrzennych i multimedialnych oraz strumieni danych - [K2st_W4] 4. ma wiedzę o procesach zachodzących w cyklu życia systemów informatycznych służących do przetwarzania strumieni danych; - [K2st_W5] 5. zna zaawansowane metody, techniki i narzędzia stosowane w wypadku obsługi danych medialnych, przestrzennych i strumieni danych, takie jak: histogramy kolorów, tekstu i kształtów dla reprezentacji zawartości obrazów, szkice i streszczenia strumieni danych. - [K2st_W6]		
<b>Umiejętności:</b>		

1. potrafi pozyskiwać informacje z literatury, baz danych integrować je, dokonywać ich interpretacji i krytycznej oceny; - [K2st_U1]
2. potrafi wykorzystać do formułowania i rozwiązywania zadań inżynierskich metody analityczne oraz eksperymentalne, np. dobrać właściwy rozmiar streszczeń strumieni danych dla zadanej wielkości błędu; - [K2st_U4]
3. potrafi integrować wiedzę z różnych obszarów informatyki: języków programowania, grafiki komputerowej i Internetu; - [K2st_U5]
4. potrafi ocenić przydatność i możliwość wykorzystania nowych osiągnięć oraz nowych produktów informatycznych - [K2st_U6]
5. potrafi dokonać krytycznej analizy istniejących rozwiązań technicznych oraz zaproponować ich ulepszenia, np. przez rozszerzenia modelu obiektowego; - [K2st_U8]
6. potrafi ocenić przydatność metod i narzędzi służących do składowania i wyszukiwania obrazów; - [K2st_U9]
7. potrafi ocenić złożoność obliczeniową algorytmów i problemów, - [K2st_U10]
<b>Kompetencje społeczne:</b>
1. rozumie, że w informatyce wiedza i umiejętności bardzo szybko stają się przestarzałe, na przykładzie ewolucji multimedialnych baz danych lub obiektowych baz danych; - [K2st_K1]
2. rozumie znaczenie wykorzystywania najnowszej wiedzy z zakresu informatyki w rozwiązywaniu problemów badawczych i praktycznych, np w dziedzinie rozpoznawania obrazów w samochodach autonomicznych. - [K2st_K2]

### Sposoby sprawdzenia efektów kształcenia

Efekty kształcenia przedstawione wyżej weryfikowane są w następujący sposób:

Ocena formująca:

a) w zakresie wykładów:

- obecności i aktywności w trakcie wykładów: szukanie odpowiedzi na pytania zadawane przez wykładowcę, krytyczne podejście do tłumaczenia wykładowców, zainteresowanie rozszerzeniem zakresu wykładów, znajdowanie błędów w materiałach wykładowych,

b) w zakresie laboratoriów:

- na podstawie oceny przygotowania do realizacji zadań,  
- na podstawie dyskusji na temat materiału prezentowanego na wykładach

Ocena podsumowująca:

a) w zakresie wykładów weryfikowanie założonych efektów kształcenia realizowane jest przez:

- ocenę wiedzy i umiejętności wykazanych na egzaminie pisemnym o charakterze problemowym, który składa się z kilku otwartych zadań, polegających na przykład na odtworzeniu działania wybranych algorytmów, liczbowej weryfikacji danej hipotezy, itp. Dla uzyskania oceny 3.0 wymagane jest uzyskanie co najmniej 50% punktów. W ocenie finalnej uwzględnia jest również aktywność w trakcie wykładów.

- omówienie wyników egzaminu,

b) w zakresie laboratoriów weryfikowanie założonych efektów kształcenia realizowane jest przez:

- rozliczenie z realizacji ćwiczeń laboratoryjnych,

- ocenę i obronę przez studenta projektu,

Uzyskiwanie punktów dodatkowych za aktywność podczas zajęć, a szczególnie za:

- efektywność zastosowania zdobytej wiedzy podczas rozwiązywania zadanych problemów,

- uwagi związane z udoskonaleniem materiałów dydaktycznych,

- wskazywanie trudności percepcyjnych studentów umożliwiające bieżące doskonalenia procesu dydaktycznego.

### Treści programowe

1. Nowe dziedziny zastosowań baz danych i ich specyfika. Nieadekwatność klasycznych systemów baz danych do rozwiązywania problemów w nowych dziedzinach zastosowań. Potrzeba znalezienia nowych generacji systemów baz danych, opracowania nowych modeli danych i nowych rozwiązań systemowych.

2. Obiektowo-relacyjne bazy danych; nowe konstrukcje: klasy, dziedziczenie klas, obiektowy typ danych, konstruktory złożonych typów danych; systemowy identyfikator obiektów; referencje między obiektami, polimorficzne kolekcje obiektów; język definiowania, zapytań i przetwarzania obiektów w bazie danych.

3. Przechowywanie i wyszukiwanie obrazów. Wyszukiwanie obrazów za pomocą ich zawartości (CBIR). Deskryptory obrazów: deskryptory kolorów, tekstury i kształtów. Lokalne deskryptory obrazów: punkty charakterystyczne (punkty zainteresowania). Algorytmy identyfikacji i deskryptory punktów charakterystycznych.

4. Bazy dokumentów tekstowych. Specyfika wyszukiwania dokumentów tekstowych, miary jakości wyszukiwania. Reprezentacje dokumentów tekstowych. Algorytmy wyszukiwania wzorca w dokumentach tekstowych: Knutha-Morrisa-Pratta i Boyera-Moora. Miary odległości między dokumentami tekstowymi. Pliki odwrócone. Numeryczne sygnatury tekstów. Reprezentacja tekstów jako punktów w przestrzeni wielowymiarowej.

5. Wielowymiarowe struktury danych. Wielowymiarowe indeksy: R-drzewa, R+-drzewa, R\*-drzewa i R-drzewa Hilberta. Wyszukiwanie najbliższych sąsiadów za pomocą R-drzew. Struktury danych i algorytmy dla dużej liczby wymiarów.

6. Przechowywanie i przetwarzanie w bazach danych dużych obiektów binarnych (BLOB) i tekstowych (CLOB). Zalety i wady przechowywania dużych obiektów w bazie danych i w systemie plików.

7. Przestrzenne bazy danych: reprezentacja danych przestrzennych w bazach danych, indeksowanie przestrzennych baz

danych, zapytania do przestrzennych baz danych z wykorzystaniem relacji przestrzennych.

8. Standardy dotyczące przetwarzania danych tekstowych, przestrzennych i multimedialnych: SQL/MM, MPEG-7.

9. Dziedzina przemysłowych zastosowań baz danych: zarządzanie sieciami komputerowymi, zarządzanie sieciami telekomunikacyjnymi, sterownie ruchem ulicznym, zastosowania wojskowe, giełda. Specyfika nowych zastosowań przemysłowych. Strumienie danych. Systemy zarządzania strumieniami danych. Typowe zastosowania dla przetwarzania. Specyfika przetwarzania strumieni danych. Model danych strumieni danych. Języki zapytań na strumieniach danych. Rejestracja ciągłych zapytań. Operacje blokujące. Okna ślizgające się po strumieniach danych. Metody przetwarzania danych w strumieniach. Próbkowanie zbiornikowe. Zastosowanie falki Haara do tworzenia streszczeń danych. Techniki szkieletowania danych. Histogramy eksponentialne. Zastosowanie mikro-grup w procesie grupowania i klasyfikacji.

**Literatura podstawowa:**

1. Advanced Database Systems, Carlo Zaniolo, Morgan Kaufman, 1997, Part IV Spatial, Text and Multimedia Databases
2. Data Streams: Models and Algorithms, Charu Aggarwal, Springer, 2006
3. Principles of Multimedia Database Systems, V.S. Subrahmanian, Morgan Kaufmann, 1998
4. Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer Widom, Systemy baz danych. Kompletny podręcznik. Wydanie II, 2011, Rozdział 14.4. Indeksy wielowymiarowe
5. Dokumentacja systemu zarządzania bazą danych Oracle

**Literatura uzupełniająca:**

1. Managing and Mining Multimedia Databases, Bhavani Thuraisingham, CRC Press, 2001
2. Distributed Multimedia Database Technologies Supported by MPEG-7 and MPEG-21, Harald Kosch, CRC Press, 2003
3. Specyfikacja standardu SQL/MM
4. Specyfikacja standardu MPEG-7
5. Norbert Beckmann, Hans-Peter Kriegel, Ralf Schneider, Bernhard Seeger, The R\*-tree: An Efficient and Robust Access Method for Points and Rectangles, Proceeding SIGMOD '90
6. Anton Guttman, R-trees. A Dynamic Index Structure for Spatial Searching Proceeding SIGMOD '84
7. David B. Lomet, Betty Salzberg, The hB-tree: a multiattribute indexing method with good guaranteed performance, Readings in database systems (2nd ed.) Pages 136 ? 152
8. Krzysztof Jankiewicz, Marek Wojciechowski, Standard SQL/MM: SQL Multimedia and Application Packages, Materiały IX Seminarium PLOUG, 2004

**Bilans nakładu pracy przeciętnego studenta**

<b>Czynność</b>	<b>Czas (godz.)</b>
1. udział w zajęciach laboratoryjnych	30
2. udział w konsultacjach związanych z realizacją procesu kształcenia, w szczególności ćwiczeń laboratoryjnych (częściowo realizowane drogą elektroniczną)	20
3. realizacja projektu (czas poza zajęciami laboratoryjnymi)	20
4. przygotowanie do zajęć laboratoryjnych	30
5. udział w wykładach	1
6. omówienie wyników egzaminu	20
7. przygotowanie do egzaminu i obecność na egzaminie: 18 godz. + 2 godz.	10
8. zapoznanie się ze wskazaną literaturą / materiałami dydaktycznymi (10 stron tekstu naukowego = 1 godz.), 100 stron	

**Obciążenie pracą studenta**

<b>forma aktywności</b>	<b>godzin</b>	<b>ECTS</b>
Łączny nakład pracy	125	5
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	65	3
Zajęcia o charakterze praktycznym	50	2