

KARTA OPISU MODUŁU KSZTAŁCENIA		
Nazwa modułu/przedmiotu Integracja systemów transakcyjnych		Kod 1010512331010513980
Kierunek studiów Informatyka	Profil kształcenia (ogólnoakademicki, praktyczny) ogólnoakademicki	Rok / Semestr 2 / 3
Ścieżka obieralności/specjalność Technologie przetwarzania danych	Przedmiot oferowany w języku: polski	Kurs (obligatoryjny/obieralny) obligatoryjny
Stopień studiów: II stopień	Forma studiów (stacjonarna/niestacjonarna) niestacjonarna	
Godziny Wykłady: 16 Ćwiczenia: - Laboratoria: 16 Projekty/seminaria: -		Liczba punktów 4
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		Podział ECTS (liczba i %) 4 100%
Odpowiedzialny za przedmiot / wykładowca:		
<p>dr hab. inż. Robert Wrembel, prof. nadzw. email: Robert.Wrembel@put.poznan.pl tel. 61 6652991 Instytut Informatyki ul. Piotrowo 2, 60-965 Poznań</p>		
Wymagania wstępne w zakresie wiedzy, umiejętności, kompetencji społecznych:		
1	Wiedza:	Student rozpoczynający ten przedmiot powinien posiadać podstawową wiedzę z systemów baz danych i architektur systemów informatycznych.
2	Umiejętności:	Niezbędna jest umiejętność formułowania poleceń w języku SQL i umiejętność programowania w dowolnym języku proceduralnym. Student powinien ponadto posiadać umiejętność rozwiązywania podstawowych problemów z zakresu administracji i obsługi systemów baz danych oraz umiejętność pozyskiwania informacji ze wskazanych źródeł.
3	Kompetencje społeczne	Powinien również mieć gotowość do 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:		
<ol style="list-style-type: none"> 1. Wskazanie praktycznych problemów jakie rozwiązuje się integrując systemy informatyczne w dużej instytucji lub przedsiębiorstwie. 2. Przekazanie studentom wiedzy dotyczącej integracji heterogenicznych źródeł danych. 3. Przedstawienie problematyki budowy systemu rozproszonych baz danych. 4. Rozwijanie u studentów umiejętności rozwiązywania problemów. 5. Kształtowanie u studentów umiejętności pracy zespołowej w ramach tworzenia heterogenicznej, rozproszonej bazy danych i symulowania awarii w takim systemie. 6. Kształtowanie u studentów umiejętności tworzenia oprogramowania zgodnego ze standardami w ramach implementacji warstwy dostępu do bazy danych zgodnego ze standardem X/Open DTP. 		
Efekty kształcenia i odniesienie do kierunkowych efektów kształcenia		
Wiedza:		
<ol style="list-style-type: none"> 1. ma uporządkowaną, podbudowaną teoretycznie wiedzę ogólną w zakresie problematyki integracji heterogenicznych źródeł danych, architektur i technologii integracji danych, architektur systemu rozproszonych baz danych, algorytmów zatwierdzania transakcji rozproszonych, architektur replikacji danych, optymalizacji zapytań rozproszonych - [K_W4] 2. ma podbudowaną teoretycznie szczegółową wiedzę związaną z wybranymi zagadnieniami z zakresu informatyki, takimi jak: replikacja danych, algorytmy 2PC i 3PC, optymalizacja zapytań rozproszonych - [K_W5] 3. ma wiedzę o trendach rozwojowych i najistotniejszych nowych osiągnięciach w zakresie architektur i technologii integracji danych i systemów rozproszonych bazach danych - [K_W6] 4. ma podstawową wiedzę o cyklu życia systemu rozproszonych baz danych - [K_W7] 5. zna podstawowe techniki i narzędzia stosowane przy rozwiązywaniu złożonych zadań inżynierskich z zakresu integracji heterogenicznych źródeł danych, optymalizacji zapytań rozproszonych, replikacji danych - [K_W8] 		
Umiejętności:		

1. pozyskiwać informacje z literatury, dokumentacji technicznej oraz innych źródeł (w języku ojczystym i angielskim), integrować je, dokonywać ich interpretacji i krytycznej oceny, wyciągać wnioski oraz formułować i wyczerpująco uzasadniać opinie - [K_U1]
2. określić kierunki dalszego uczenia się i zrealizować proces samokształcenia z zakresu systemu rozproszonych baz danych i integracji heterogenicznych źródeł danych - [K_U5]
3. wykorzystać metody analityczne i eksperymentalne do formułowania i rozwiązywania zadań z zakresu replikacji baz danych, optymalizacji zapytań rozproszonych i usuwania awarii transakcji rozproszonych - [K_U9]
4. przy formułowaniu i rozwiązywaniu zadań z zakresu integracji rozproszonych i heterogenicznych źródeł danych - integrować wiedzę z różnych obszarów informatyki (takich jak systemy baz danych, sieci komputerowe i programowanie obiektowe) oraz zastosować podejście systemowe, uwzględniające także aspekty pozatechniczne - [K_U10]
5. formułować i testować hipotezy związane z problemami z zakresu integracji heterogenicznych źródeł danych - [K_U12]
6. ocenić przydatność i możliwość wykorzystania osiągnięć (metod i narzędzi) z zakresu replikacji i integracji wielu źródeł danych oraz produktów informatycznych do projektowania i zarządzania rozproszonymi bazami danych - [K_U13]

Kompetencje społeczne:

1. rozumie, że w systemach rozproszonych baz danych wiedza i umiejętności bardzo szybko stają się przestarzałe i potrafi systematycznie zdobywać nową wiedzę i umiejętności z tej dziedziny - [K_K1]
2. zna przykłady i rozumie przyczyny wadliwie działających rozproszonych baz danych, mechanizmów zatwierdzania rozproszonych transakcji i integracji heterogenicznych źródeł danych - [K_K4]
3. potrafi odpowiednio określić priorytety służące realizacji zadania określonego przez siebie lub innych - [K_K6]

Sposoby sprawdzenia efektów kształcenia

Efekty kształcenia przedstawione wyżej weryfikowane są w następujący sposób:

Ocena formująca:

- a) w zakresie wykładów:
 - na podstawie odpowiedzi na pytania dotyczące materiału omówionego na poprzednich wykładach,
- b) w zakresie laboratoriów:
 - na podstawie oceny bieżącego postępu realizacji zadań,

Ocena podsumowująca:

- a) w zakresie wykładów weryfikowanie założonych efektów kształcenia realizowane jest przez:
 - ocenę wiedzy i umiejętności wykazanych na egzaminie pisemnym o charakterze problemowym (student może korzystać z dowolnych materiałów dydaktycznych) i w formie testu jedno- lub wielokrotnego wyboru;

Egzamin składa się z około 6 zadań problemowych i 22 pytań testowych jedno lub wielokrotnego wyboru (proporcje między zadaniami problemowymi i testowymi mogą się zmieniać, w zależności od roku akademickiego). Maksymalnie można uzyskać 50 punktów, z czego nie więcej niż 22 za pytania testowe. Nie przyznaje się punktów ułamkowych. Kolokwium uznaje się zaliczone od 26 punktów. Przyjmuje się następującą skalę ocen i punktów:

0-25: ndst
26-30: dst
31-35: dst+
36-40: db
41-45: db+
46-50: bdb

- omówienie wyników egzaminu;
- b) w zakresie laboratoriów weryfikowanie założonych efektów kształcenia realizowane jest przez:
 - ocenę wiedzy i umiejętności związanych z realizacją zadań projektowych,
 - ocenę i obronę przez studenta sprawozdania z realizacji projektu,
 - ocenę wiedzy i umiejętności w formie testu wielokrotnego wyboru.

Uzyskiwanie punktów dodatkowych za aktywność podczas zajęć (laboratoryjnych), a szczególnie za:

- omówienia dodatkowych aspektów zagadnienia,
- umiejętność współpracy w ramach zespołu praktycznie realizującego zadanie szczegółowe w laboratorium,
- uwagi związane z udoskonaleniem materiałów dydaktycznych,
- wskazywanie trudności percepcyjnych studentów umożliwiające bieżące doskonalenia procesu dydaktycznego.

Treści programowe

Program wykładu obejmuje następujące zagadnienia:

- problematyka integracji heterogenicznych źródeł danych (charakterystyki systemów, struktur danych, samych danych),
- architektury integracji danych (system mediacyjny, system sfederowany, system hurtowni danych, web service),
- technologie dostępu do heterogenicznych źródeł danych (gateway, ODBC/JDBC, OLEDB, specjalizowane sterowniki txt i xml),
- studium przypadku integracji źródeł danych,
- problematyka budowy systemu rozproszonych baz danych,
- architektury replikacji danych,
- funkcjonalność replikacji w Oracle, DB2, SQL Server,
- zarządzanie transakcjami rozproszonymi (protokół 2PC i 3PC),
- zarządzanie blokadami w systemie rozproszonych baz danych,
- problematyka alokacji danych w systemie rozproszonych baz danych (algorytmy Best fit, Optymalizacja zysku, Statyczny, Dynamiczny),
- optymalizacja zapytań rozproszonych (proces optymalizacji, rodzaje optymalizatorów, algorytmy przeszukiwania przestrzeni planów, techniki redukcji planów zapytań rozproszonych).

Zajęcia laboratoryjne prowadzone są w formie piętnastu 2-godzinnych ćwiczeń, odbywających się w laboratorium. Program laboratorium obejmuje następujące zagadnienia:

Podstawy zarządzania rozproszonymi systemami baz danych: ogólna architektura rozproszonych systemów baz danych, zarządzanie zdalnymi połączeniami, problem uwierzytelniania użytkowników i szyfrowania danych, nazwy globalne baz danych, łączniki bazy danych, zapewnianie przezroczystości lokalizacji. Replikacja danych: architektura i mechanizm tworzenia migawek (perspektyw zmierzających do materializowanych), typy replikacji standardowa (jednokierunkowa) pełna, standardowa przyrostowa, dwukierunkowa pełna, dwukierunkowa przyrostowa, problemy zdalnego odświeżania, warunki automatycznego odświeżania zdalnych migawek. Zbieranie informacji potrzebnych do replikacji migawek opartych o skomplikowane zapytania SQL i warunki ich odświeżania, grupy odświeżania migawek. Porównanie mechanizmów tworzenia i odświeżania perspektyw zmierzających do materializowanych w różnych systemach zarządzania bazami danych, przykłady rzeczywistych replikacji i doboru metod do ich rozwiązania. Zarządzanie danymi i podstawy replikacji w nierelacyjnych bazach danych (tzw. bazach NO-SQL). Problem zatwierdzania transakcji rozproszonych: algorytm 2PC, symulowanie awarii procesu zatwierdzania, ćwiczenia praktyczne z naprawiania poszczególnych typów awarii transakcji rozproszonej. Optymalizacja zapytań rozproszonych: odpowiedzi do zapytań rozproszonych w różnych systemach zarządzania bazą danych, wskazywanie węzła przetwarzającego, warunki połączeń tabel w rozproszonych bazach danych, analiza kosztów zapytań w zależności od liczby i kierunku przesyłanych danych. Integracja wielu heterogenicznych baz danych za pomocą wbudowanych sterowników, własnego kodu korzystającego ze standardu X/Open DTP. Omówienie problemów napotkanych podczas integracji wielu systemów zarządzania bazami danych wykonanej w ramach projektu zespołowego. Przykłady rzeczywistych problemów integracji systemów transakcyjnych i doboru właściwych metod do ich rozwiązania.

Cześć wymienionych wyżej treści programowych realizowana jest w ramach pracy własnej studenta.

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. ćwiczenia laboratoryjne: rozwiązywanie zadań, ćwiczenia praktyczne, burza mózgów, praca w zespole, pokaz multimedialny

Literatura podstawowa:

1. M.T. Özsu, P. Valduriez: Principles of Distributed Database Systems. 3rd edition, Springer Verlag, 2011
2. T. Conolly, C. Begg: Database Systems - a Practical Approach to Design, Implementation, and Management. Addison-Wesley, 2002

Literatura uzupełniająca:

1. S.K. Rahimi, F.S. Haug: Distributed Database Management Systems. A Practical Approach. IEEE Computer Society, Wiley, 2010
2. K. Stocker, D. Kossmann, R. Braumandl, A. Kemper: Integrating semi-join-reducers into state-of-the-art query processors. Proc. of ICDE, 2001
3. Y-F. Huang, J-H. Chen, Fragment Allocation in Distributed Database Design, July 2000, pp. 491-506
4. A. Brunstrom, S. T. Leutenegger, R. Simha, Experimental Evaluation of Dynamic Data Allocation Strategies in a Distributed Database With Changing Workloads, no. TR-95-2, 1995, pp. 1-15
5. P.M. G. Apers, Data Allocation in Distributed Database Systems, ACM Transactions on Database Systems, vol. 13, no. 3, September 1988, pp. 263-304

Bilans nakładu pracy przeciętnego studenta

Czynność	Czas (godz.)
----------	--------------

1. udział w zajęciach laboratoryjnych:	16
2. przygotowanie do ćwiczeń laboratoryjnych:	8
3. udział w konsultacjach (mogą być realizowane drogą elektroniczną) związanych z realizacją procesu kształcenia, w szczególności ćwiczeń laboratoryjnych / projektu	4 20
4. napisanie programu / programów, uruchomienie i weryfikacja (czas poza zajęciami laboratoryjnymi)	8
5. przygotowanie do sprawdzianów / kolokwium	16
6. udział w wykładach	20
7. zapoznanie się ze wskazaną literaturą / materiałami dydaktycznymi (10 stron tekstu naukowego = 1 godz.), 200 stron	12
8. przygotowanie do egzaminu, udział w egzaminie (2godz.) oraz omówienie wyników egzaminu (2 godz.)	
Obciążenie pracą studenta	
forma aktywności	godzin
ECTS	
Łączny nakład pracy	104
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	40
Zajęcia o charakterze praktycznym	44