

ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ

(повне найменування вищого навчального закладу)

Кафедра Програмних засобів

“ЗАТВЕРДЖУЮ”

Ректор (перший проректор)

_____ В.Г. Прушківський
“ _____ ” _____ 2016 року

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

МЕТОДИ АНАЛІЗУ ДАНИХ ТА ЇХ ОРГАНІЗАЦІЯ

(код і назва навчальної дисципліни)

спеціальність 122«Комп’ютерні науки»

(код і назва спеціальності)

освітня програма «Інформаційні технології проектування»

(назва освітньої програми)

інститут, факультет Інститут інформатики та радіоелектроніки, факультет
комп’ютерних наук і технологій

(назва інституту, факультету)

Робоча програма «Методи аналізу даних та їх організація» для студентів
(назва навчальної дисципліни)
 магістрів за спеціальністю 122«Комп'ютерні науки», освітня програма
 «Інформаційні технології проектування», 2016 р. – 10 с.

Розробники: (вказати авторів, їхні посади, наукові ступені та вчені звання)
 Федорончак Тетяна Василівна, доц. каф. ПЗ, к.т.н., доц.

Робоча програма затверджена на засіданні кафедри
програмних засобів

Протокол від. “___” _____ 2016 року № ___

Завідувач кафедри програмних засобів

_____ (Субботін С.О.)
(підпис) (прізвище та ініціали)
 “___” _____ 2016 року

Схвалено методичною комісією вищого навчального закладу за спеціальністю
122«Комп'ютерні науки та інформаційні технології»
(код, назва)

Протокол від. “___” _____ 2016 року № ___

“___” _____ 20__ року Голова _____ (М.М. Касьян)
(підпис) (прізвище та ініціали)

1. Опис навчальної дисципліни

Найменування показників	Галузь знань, напрям підготовки, освітньо-кваліфікаційний рівень	Характеристика навчальної дисципліни	
		денна форма навчання	заочна форма навчання
Кількість кредитів – 4	Галузь знань <u>12 Інформаційні технології</u> (шифр і назва)	Нормативна	
	Спеціальність: <u>122«Комп'ютерні науки»</u> (шифр і назва)		
Модулів – 2	Освітня програма: <u>«Інформаційні технології проектування»</u>	Рік підготовки:	
Змістових модулів – 2		5-й	
Індивідуальне науково-дослідне завдання (назва)		Семестр	
Загальна кількість годин – 120		10-й	
Тижневих годин для денної форми навчання: аудиторних – 3 самостійної роботи студента – 5, 6	Освітньо-кваліфікаційний рівень: магістр	Лекції	
		28 год.	
		Практичні, семінарські	
		--	
		Лабораторні	
		14 год.	
		Самостійна робота	
		78 год.	
		Індивідуальні завдання: --	
		Вид контролю: іспит	

Примітка.

Співвідношення кількості годин аудиторних занять до самостійної і індивідуальної роботи становить:
для денної форми навчання – 0,5

2. Мета та завдання навчальної дисципліни

Мета – придбання фундаментальних знань, вмінь та навичок, необхідних для збору та представлення інформації у вигляді, придатному для машинної обробки, а також наступного статистичного та інтелектуального аналізу структурованих даних.

Завдання дисципліни полягає в ознайомленні студентів з новітніми концепціями розвитку Всесвітньої мережі World Wide Web, а саме з семантичною павутиною (Semantic web) - технологією, що дозволяє трансформувати інформаційне наповнення Інтернету на глобальну базу знань, методами семантичного та інформаційного пошуку, методами інтелектуального аналізу даних, що використовуються для виявлення знань в мережі Інтернет, методами статистичної обробки даних.

У результаті вивчення навчальної дисципліни студент повинен

з н а т и :

- методи збору, інтеграції та попередньої обробки даних про предметну область досліджень;
- методи забезпечення інтероперабельності програмних систем;
- методи представлення даних у вигляді семантичних моделей;
- інтелектуальні методи аналізу даних;

в м і т и :

- виконувати семантичний опис даних і знань про предметні області комп'ютерних інформаційних технологій;
- виконувати витяг знань з баз і сховищ даних для розробки та використання математичних моделей і програмного забезпечення шляхом обробки даних з використанням методів математичної статистики, машинного навчання та штучного інтелекту;
- використовувати моделювання для вирішення задач прогнозування та прийняття рішень в різних предметних областях
- інтегрувати програмні системи і програми та забезпечити їх інтероперабельність (сумісність);
- застосовувати набуті знання для здійснення професійної діяльності при розробці, налагодженні та експлуатації ІС та технологій.

3. Програма навчальної дисципліни

Змістовий модуль 1. Семантичний опис даних, системне інтегрування різних за природою даних, інтероперабельність.

Тема 1. Інтеграція та інтероперабельність (сумісність) даних.

Проблема інтеграції даних. Методи синтаксичної та семантичної інтеграції даних. Інтероперабельність даних.

Тема 2. Інтелектуальна обробка та інтелектуальна інтеграція даних.

Методи інформаційного пошуку. Задача інтелектуального пошуку. Роль технологій Semantic Web у інтелектуалізації пошуку. Задача інтелектуальної інтеграції інформації (ініціатива І3 – Intelligent Information Integration).

Проблеми інтелектуальної інтеграції: різноманітність і автономність інформаційних ресурсів на інфраструктурному, мовному, синтаксичному, семантичному рівнях.

Змістовий модуль 2 Семантичний веб та його технології

Тема 1. Семантичний веб

Поняття Semantic Web, Структура базової моделі Semantic Web. Поняття онтології. Агенти та сервіси. Представлення знань для Semantic Web. Linked Data в середовищі Semantic Web. Концепція Open Data. Технологія інтелектуального аналізу у мережі Інтернет.

Тема 2. Семантичні технології веб

Розширювана мова розмітки (XML, eXtensible Markup Language). Загальна схема опису ресурсів RDF (Resource Description Framework), RDF Schema. Метадані, Dublin Core. Онтології, мова опису веб-онтологій OWL (Web Ontology Language). Мови запитів до RDF сховищ (SPARQL, Protocol and RDF Query Language). Принцип «логічного висновку», дескриптивна логіка, мова правил веб-семантики (SWRL, Semantic Web Rule Language).

4. Структура навчальної дисципліни

Назви змістових модулів і тем	Кількість годин											
	денна форма						заочна форма					
	усь го	у тому числі					усь ого	у тому числі				
		л	п	лаб	інд	с.р.		л	п	лаб	інд	с.р.
1	2	3	4	5	6	7	8	9	10	11	12	13
Змістовий модуль 1. Семантичний опис даних, системне інтегрування різних за природою даних, інтероперабельність.												
Тема 1. Інтеграція та інтероперабельність (сумісність) даних.	24	6				18						
Тема 2. Інтелектуальна обробка та інтелектуальна інтеграція даних.	24	6				18						
Разом за змістовим модулем 1	48	12		-		6						
Змістовий модуль 2. Семантичний веб та його технології												
Тема 1. Семантичний веб	30	8		2		20						

Тема 2. Семантичні технології веб	42	8		12		22						
Разом за змістовим модулем 2	72	16		14		40						
Разом	120	28	-	14	-	78						

5. Теми семінарських занять

Не передбачено.

6. Теми практичних занять

Не передбачено.

7. Теми лабораторних занять

№ з/п	Назва теми	К-ть годин
1	Представлення даних в форматі XML	2
2	Представлення даних в форматі RDF	2
3	Мова представлення онтологій OWL	4
4	Інженерія знань на основі логіки	4
5	Інтеграція RDF та HTML	2
	Усього	14

8. Самостійна робота

№ з/п	Назва теми	К-ть годин
1	Методи очищення зібраних сирих даних. Витяг знань з баз і сховищ даних і їхня інтеграція	4
2	Особливості інтелектуальної інтеграції інформації у додатках e-Commerce, e-Science, e-Government, Grid.	10
3	Технологія інтелектуального аналізу у мережі Інтернет (Web Mining): Web Content Mining; Web Structure Mining; Web Usage Mining.	10
	Усього	44

Перелічені теми повністю виносяться на самостійний розгляд студентів.

Самостійна робота студента включає:

1. повторення матеріалу, засвоєного на лекціях, самостійне опанування частини теоретичного матеріалу, підготовку до лабораторних робіт, підготовку до

іспиту;

2. ознайомлення з програмним забезпеченням для виконання лабораторних робіт, вивчення навчальних матеріалів по роботі з ними : EditiX; MR3; Protégé; SW Workbench; Jena; Fact++.

Загалом самостійна робота студента складає 78 год.

9. Індивідуальні завдання

Не передбачено.

10. Методи навчання

Організаційні форми навчання: лекційні заняття, лабораторні заняття, самостійна робота студентів.

Основні методи активного навчання: розповідь, бесіда, дискусія, діалог, індивідуальний/груповий проект, презентація.

11. Методи контролю

Усний контроль у вигляді індивідуального та фронтального опитування.

Письмовий контроль у вигляді самостійних письмових робіт, поточного тестування.

12. Розподіл балів, які отримують студенти

Поточне тестування та самостійна робота				Сума
ЗМ 1		ЗМ 2		
T1	T2	T1	T2	
20	20	30	30	100

T1, T2 ... T12 – теми змістових модулів.

Шкала оцінювання: національна та ECTS

Сума балів за всі види навчальної діяльності	Оцінка ECTS	Оцінка за національною шкалою	
		для екзамену, курсового проекту (роботи), практики	для заліку
90 – 100	A	відмінно	зараховано
85-89	B	добре	
75-84	C		
70-74	D	задовільно	
60-69	E		
35-59	FX	незадовільно з можливістю повторного складання	не зараховано з можливістю повторного

			складання
0-34	F	незадовільно з обов'язковим повторним вивченням дисципліни	не зараховано з обов'язковим повторним вивченням дисципліни

13. Методичне забезпечення

1. Конспект лекцій.

2. Методичні вказівки до лабораторних робіт з дисципліни “Методи аналізу даних та їх організація” для студентів напряму підготовки 6.050101 «Комп’ютерні науки» спеціальності 8.05010102 «Інформаційні технології проектування» / Д.М. Піза, Т.В. Юр. – Запоріжжя: ЗНТУ, 2013. – 50 с.

14. Рекомендована література

Базова

1. Ситник В.Ф. Інтелектуальний аналіз даних (дейтамайнінг): навчальний посібник / В.Ф. Ситник, М.Т. Краснюк. – К.: КНЕУ, 2007. – 376 с.

2. Барсегян А. А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP: учебное пособие для студ. Вузов / А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И.И. Холод. – 2-е изд. – СПб.: БХВ-Петербург, 2007. – 384 с.

3. Бородакий Ю.В. Вероятностно-статистические методы обработки данных в информационных системах / Ю.В. Бородакий, Н.А. Крицына, Ю.П. Кулябичев, Ю.Ю. Шумилов. – М.: Радио и связь, 2003. – 264 с.

4. Савицкий Н.И. Технологии организации, хранения и обработки данных: учеб. пособие / Н.И. Савицкий. – М: ИНФРА-М, 2001. – 232 с.

5. Hebel John. Semantic Web Programming / John Hebel, Matthew Fisher, Ryan Blace, Andrew Perez-Lopez. – Wiley Publishing, 2009. – 648 p.

6. Allemang Dean. Semantic Web for the working ontologist : effective modeling in RDFS and OWL / Dean Allemang, Jim Hendler. – 2nd ed. - Morgan Kaufmann Publishers, 2011. – 354 p.

7. Онтологии и тезаурусы: модели, инструменты, приложения : учебное пособие / Б.В. Добров, В.В. Иванов, Н.В. Лукашевич, В.Д. Соловьев. – М.: БИНОМ. Лаборатория знаний, 2009. – 173 с.

8. Маннинг Кристофер Д. Введение в информационный поиск / Маннинг Кристофер Д., Рагхаван Прабхакар, Шютце Хайнрих. – М.: ООО"И.Д.Вильямс", 2011. – 528 с.

Допоміжна

9. Олійник А.О. Інтелектуальний аналіз даних: навчальний посібник для студ. вищ. навч. закл.: рек. МОНУ / А.О. Олійник, С.О. Субботін, О.О. Олійник. – Запоріжжя: ЗНТУ, 2012. – 278 с.

10. Segaran Toby. Programming the Semantic Web / Toby Segaran, Colin Evans, Jamie Taylor.. - O'Reilly Media, 2009. - 302 p.

11. Pollock Jeffrey T. Semantic Web For Dummies / Jeffrey T. Pollock. – Wiley, 2009. – 432 p.

12. Amerland David. Google Semantic Search / David Amerland. – Que Publishing, 2013. – 240 p.

13. Інформаційні ресурси

1. W3C Semantic Web Activity [Електронний ресурс]. Режим доступу: <http://www.w3.org/2001/sw/>

2. Semantic Web Community Wiki [Електронний ресурс]. Режим доступу: <http://semanticweb.org>