ЧЕРКАСЬКИЙ ДЕРЖАВНИЙ ТЕХНОЛОГІЧНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ І СИСТЕМ

«ЗАТВЕРДЖУЮ»
Голова вченої ради
ФІТІС
_____ / Чепинога А.В. /
Протокол № 1
«30» серпня 2021 р.р.

РОБОЧА ПРОГРАМА

навчальної дисципліни

«Великі дані: системи, програмування і управління»

Шифр за ОНП – ОД$_8$

Рівень вищої освіти — Освітньо-науковий (доктор філософії)

Спеціальність - 121 - Інженерія програмного забезпечення

Освітня програма - «Інженерія програмного забезпечення»

2021-2022 навчальний рік

Робоча програма навчальної дисципліни «Великі дані: системи, програмування і управління» підготовки здобувачів освітньо-наукового рівня «Доктор філософії» за спеціальністю 121 – Інженерія програмного забезпечення, освітня програма «Інженерія програмного забезпечення» - *11 сторінок.*

Розробник:

Рідкокаша А.А., доцент кафедри програмного забезпечення автоматизованих систем, к.т.н., доцент

Робоча програма навчальної дисципліни затверджена на засіданні кафедри програмне забезпечення автоматизованих систем

Протокол № 2 від « 26 » серпня 2021 р.

Завідувач кафедри _____ /Первунінський С.М./

Обговорено та рекомендовано до затвердження методичною комісією факультету інформаційних технологій і систем

«30» серпня 2021 р., протокол № 1

Голова методичної комісії факультету інформаційних технологій і систем

_____ /Карапетян А.Р./
підпис *ПІБ*

1 ІНФОРМАЦІЯ ПРО ВИКЛАДАЧА

Прізвище, ім’я, по батькові	Рідкокаша Анатолій Андрійович
Науковий ступінь	кандидат технічних наук
Вчене звання	доцент
Посада	доцент кафедри програмного забезпечення автоматизованих систем
Місце роботи	Черкаський державний технологічний університет
Адреса кафедри	бульв. Шевченка, 460, м. Черкаси, ЧДТУ, 1корпус, каб. 605
Контактний телефон	+38(098)-524-53-43
Профайл викладача	https://pzas.chdtu.edu.ua/staff/ridkokasha-anatolij-andrijovych/
e-mail:	redanan1945@gmail.com
Профайл дисципліни	
Розклад консультацій	https://pzas.chdtu.edu.ua/debts/

2. ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

<table>
<tr><td rowspan="2">Галузь знань, спеціальність, освітня програма, освітній рівень</td><td colspan="2" rowspan="2">Загальні характеристики</td><td colspan="2">Навчальне навантаження з дисципліни</td></tr>
<tr><td>денна форма</td><td>заочна форма</td></tr>
<tr><td rowspan="2"><u>Галузь знань</u>
12 «Інформаційні технології»</td><td colspan="2" rowspan="2">Обов’язкова</td><td colspan="2">Курс підготовки:</td></tr>
<tr><td>1</td><td></td></tr>
<tr><td rowspan="2"><u>Спеціальність</u>
121 – Інженерія програмного забезпечення</td><td>Загальна кількість кредитів ЄКТС</td><td>4</td><td colspan="2">Семестр підготовки:</td></tr>
<tr><td>Загальна кількість годин</td><td>120</td><td>2</td><td></td></tr>
<tr><td rowspan="4"><u>Освітня програма</u>
«Інженерія програмного забезпечення»</td><td rowspan="2">Кількість аудиторних годин</td><td rowspan="2">48</td><td colspan="2">Лекції</td></tr>
<tr><td>32</td><td></td></tr>
<tr><td rowspan="2">Кількість годин самостійної роботи</td><td rowspan="2">72</td><td colspan="2">Практичні, семінарські</td></tr>
<tr><td>16</td><td></td></tr>
<tr><td rowspan="6"><u>Освітньо-науковий рівень</u>

Доктор філософії</td><td colspan="2" rowspan="6">Мова навчання - українська</td><td colspan="2">Лабораторні</td></tr>
<tr><td>–</td><td></td></tr>
<tr><td colspan="2">Самостійна робота</td></tr>
<tr><td>72</td><td></td></tr>
<tr><td colspan="2">Форма підсумкового контролю</td></tr>
<tr><td>іспит</td><td></td></tr>
</table>

3 МЕТА ТА ЗАВДАННЯ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Мета викладання дисципліни	Метою є представити системи, програмування і управління в опрацюванні великих даних в структурованому і систематизованому вигляді для теоретичного й практичного навчання процесам аналізу великих даних.
Завдання вивчення дисципліни	Основними завданнями є сформувати у аспірантів уміння науково-дослідної роботи проектувати та реалізовувати базу даних при виконанні типових задач діяльності «Збирання та аналіз інформації, визначення гіпотез і проблем», «Управління інформацією», «Моделювання», «Знаходження закономірностей».

4 РЕЗУЛЬТАТИ НАВЧАННЯ

№ з/п	Результати навчання
	РН1. Мати передові концептуальні та методологічні знання з інженерії програмного забезпечення та дотичних до неї міждисциплінарних напрямів
	$РН_3$. Пропонувати нові ефективні методи і моделі розроблення, впровадження, супроводу програмного забезпечення та управління відповідними процесами на всіх етапах життєвого циклу
1	$РН_9$. Формулює та вирішує задачі оптимізації, адаптації, прогнозування, керування та прийняття рішень щодо процесів, засобів та ресурсів розробки, впровадження, супроводу та експлуатації програмного забезпечення.
2	$СК_1$**.** Здатен виконувати оригінальні дослідження, досягати наукових результатів, які створюють нові знання в інженерії програмного забезпечення та дотичних до неї міждисциплінарних напрямах
3	$СК_3$. Розробляє нові та вдосконалює існуючи моделі, методи, засоби, процеси у сфері програмної інженерії, які забезпечують розвиток або надають нові можливості технологіям розробки та використання програмного забезпечення.
4	$СК_4$**.** Здатен до продукування нових ідей і розв'язання комплексних проблем у сфері інженерії програмного забезпечення та дотичних до неї міждисциплінарних напрямах.
5	$СК_5$**.** Здатен до застосування сучасних методологій, методів та інструментів інженерії програмного забезпечення в педагогічній та науковій діяльності

5 ПРЕРЕКВІЗИТИ

Здобувачі ступеня «доктор філософії» повинні мати базові знання з дисциплін «Бази даних» (ОР бакалавр), «Об'єктно-орієнтоване програмування» (ОР бакалавр), «Методологія та організація досліджень у програмуванні» (ОР магістр). Дисципліна «Великі дані: системи, програмування і управління» також базується на знаннях і навичках, отриманих аспірантами

при вивченні дисциплін «Програмні агенти та мультиагентні системи», «Філософія науки та методологія наукових досліджень».

6 ПОСТРЕКВІЗИТИ

На базі дисципліни «Великі дані: системи, програмування і управління» будується «Науково-дослідна робота аспіранта». Набуті компетентності використовуються у професійній діяльності

7 ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Тема 1. Інформаційні моделі та системи
Основні визначення. Структури даних. Типи завдань. Зменшення розмірності. Види обробки. Програмні системи
Тема 2. Розподілена обробка інформації
Розподілені обчислення як сучасні концептуальні знання в інженерії програмного забезпечення. Сучасні методології обробки інформації. Попередня обробка вхідних даних. Згортання попередньо оброблених даних. Модель MapReduce. Стадія Map. Стадія Shuffle. Стадія Reduce. Інструменти Hive та Pig
Тема 3. Побудова розподілених обчислень у методах і моделях розроблення, впровадження, супроводу програмного забезпечення
Кластер Hadoop. Утіліти. Hadoop – демонстраційна віртуальна машина
Тема 4. Нереляційні розподілені бази даних
BigTable. Hbase. Методи, засоби і процеси
Тема 5. Статистичне опрацювання великих обсягів даних
Методи, програмування. Порівняння статистичних пакетів
Тема 6. Інтелектуальна обробка
Інтелектуальний аналіз даних. Генетичний алгоритм. Програмування.
Тема 7. Опрацювання невизначеності у великих даних
Рівні невизначеності. Формування агрегату. Побудова результатів. Нові знання
Тема 8. Моделювання
Модель сутність-характеристика. Прототипування. Побудова залежностей, прогнозів. Оптимізація, адаптація, прогнозування, керування та прийняття рішень щодо процесів, засобів та ресурсів розробки

8 ТЕМАТИЧНИЙ ПЛАН НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

№ теми	Назва модулів і тем	Форми організації навчання, кількість годин						Література, інформаційні ресурси
		Денна форма			Заочна форма			
		Лекції	Практичні заняття	Самостійна робота	Лекції	Практичні заняття	Самостійна робота	
1	Інформаційні моделі та системи	2	1	6				основна: 1, 2, 4, 5; додаткова: 1, 3, 4
2	Розподілена обробка	4	2	10				основна: 1,2,4,5

	інформації							додаткова: 1,2,4,6
3	Побудова розподілених обчислень	4	3	10				основна: 2, 3, 4, 5 додаткова: 1-6
4	Нереляційні розподілені бази даних	2	2	10				основна: 2, 4, 5; додаткова: 1, 3,4
5	Статистичне опрацювання великих обсягів даних	6	2	10				основна: 1, 3, 4 додаткова: 1-6
6	Інтелектуальна обробка	6	2	10				основна: 1, 3, 4 додаткова: 1,
7	Опрацювання невизначеності	4	2	10				основна: 1, 3, 4 додаткова: 1,
8	Моделювання	4	2	6				основна: 1, 3, 4 додаткова: 1,
	Разом	32	16	72				

9 ПРАКТИЧНІ ЗАНЯТТЯ

№ з/п	Тема заняття	Кількість годин	
		Денна	Заочна
1	Інформаційні моделі	1	
2	Розподілена обробка інформації	2	
3	Map-reduce	3	
4	Розподілені бази даних.	2	
5	Статистичне опрацювання	2	
6	Інтелектуальний аналіз	2	
7	Опрацювання невизначеності	2	
8	Моделювання	2	
	Усього годин	16	

МЕТОДИЧНЕ ЗАБЕЗПЕЧЕННЯ

1. Методичні рекомендації до практичних занять

10 САМОСТІЙНА РОБОТА

№ з/п	Назва теми	Кількість годин
1.	Інформаційні моделі	6
2.	Розподілена обробка інформації	10

3	Концепція розподілу обчислень MapReduce	10
4	Розподілені бази даних.	10
5	Статистичне опрацювання	10
6.	Інтелектуальний аналіз	10
7.	Опрацювання невизначеності	10
8.	Моделювання	6
	Усього годин	72

11 СИСТЕМА ОЦІНЮВАННЯ НАВЧАЛЬНИХ ДОСЯГНЕНЬ

11.1 Методи контролю

Поточний контроль знань, вмінь та навичок аспірантів передбачає застосування таких видів:

- контрольні завдання;
- практичні завдання;
- обговорення проблеми, дискусія;
- аналіз конкретних ситуацій (поданих у вигляді усного, текстового або графічного матеріалу);
- презентації результатів роботи;
- інші.

Об'єктами поточного контролю знань аспірантів є:

– систематичність, активність, своєчасність та результативність роботи над вивченням програмного матеріалу дисципліни, у т.ч. виконання домашніх завдань та розв'язання задач;

– виконання завдань для самостійного опрацювання.

Оцінювання знань аспірантів здійснюється за 100-бальною шкалою:

– максимальна кількість балів при оцінюванні знань аспірантів становить за поточну успішність 60 балів, на іспиті – 40 балів;

– при оформленні документів за залікову сесію використовується таблиця відповідності оцінювання знань аспірантів за різними системами:

СУМА БАЛІВ	ОЦІНКА ECTS	Оцінка за національною шкалою	
		екзамен	залік
90 – 100	A	відмінно	зараховано
82 – 89	B	добре	
74 – 81	C		
64 – 73	D	задовільно	
60 – 63	E		
35 – 59	FX	незадовільно	не зараховано

1 – 34	F		

Об’єктом ***підсумкового контролю знань*** аспірантів є ***іспит.*** Для відповіді аспірант повинен розв’язати контрольні завдання та відповісти на теоретичні питання.

На ***іспит*** виносяться базові питання та завдання, що потребують творчого підходу та вміння опрацьовувати лекційний матеріал.

Результати іспиту аспірантів денної форми навчання оцінюються в діапазоні від 0 до 40 балів.

11.2 Питання до іспиту

1.Великі дані. Визначення. Витягування великих даних. Аналітика

2. Методи і техніки аналізу. Mapreduce. Реалізації Mapreduce

3.Hadoop. Склад. Hadoop MapReduce

4. Зберігання великих даних. Big Table, HBase, Hive

5.Створення таблиць, запитів у HBase

6.Статистичний аналіз. Лінійна регресія

7. Множинна регресія

8. Мова R у роботі з статистикою

9. Mining Data для великих даних

10. Пошук асоціативних зв’язків

11. Невизначеність даних. Опрацювання

12. Моделювання. Сутність-зв’язок

13.Візуалізація даних

14. Ролі у роботі з великими даними

11.3 Критерії оцінювання

ДЕННА ФОРМА

Критерії оцінювання знань	Кількість балів, *максимум*
Виконання практичного завдання №1	5
Виконання практичного завдання №2	10
Виконання практичного завдання №3	10
Виконання практичного завдання №4	10
Виконання практичного завдання №5	10
Виконання практичного завдання №6	10
Виконання практичного завдання №7	5

Іспит	40
Разом	100

12 РЕКОМЕНДОВАНА ЛІТЕРАТУРА

Основна

1. Coronel C. Database Systems: Design, Implementation, and Management / Carlos Coronel and Steven Morris. – 12th edition. – Cengage Learning, 2017. – 791 p.
2. Гнатюк В. Вступ до R на прикладах / В. Гнатюк. Харків: ХНЕУ, 2010. – 108 с.
3. Рідкокаша А.А. Основи систем штучного інтелекту: навч. посіб. / А.А.Рідкокаша, К.К. Голдер. Черкаси. Відлуння – Плюс, 2002. – 224 с.
4 Rajaraman A. Mining of Massive Datasets / Rajaraman A., Ullman J. D. Stanford: Kosmix, Inc., Stanford Univ., 2011. – 324 p.
5. Бахрушин В.Є. Методи аналізу даних: Навч. посіб. / В.Є. Бахрушин. – Запоріжжя: КПУ, 2011. – 268 .

Додаткова

1. Lars George. HBase: The Definitive Guide. O'Reilly Media. 2011. 554 с.
2. Apache HBase Reference Guide, 2011 [Електронний ресурс] – Режим доступу до ресурсу: http://hbase.apache.org/book/book.html (дата звернення 1.08.2021 р).
3. Фаулер М., Садаладж Пр. Дж. NoSQL: методология разработки нереляционных баз данных. К: Диалектика, 2013. 192 с.
4. The MongoDB 4.4 Manual. Technical documentation – 2021. [Електронний ресурс] – Режим доступу до ресурсу: https://docs.mongodb.com/manual/ (дата звернення 1.08.2021 р).
5. Journal of Big Data [Electronic resource] http://journalofbigdata.springeropen.com/ (last access: 14.02.21). – Title from the screen
6. Big Data Research Registered in The Netherlands, Registration No. 33156677, BTW No. NL 005033019B01. [Electronic resource] https://www.journals.elsevier.com/big-data-research/ (last access: 14.02.21). – Title from the screen
7. Кравець І. М. Розподілення навантаження в базах даних великого об’єму методом горизонтальної фрагментації Автореферат дисертації Київ: НАУ, 2010. – 22 с. https://www.ikravets.com/research/2010/09/15/dissertation-distribution-of-load-balancing-in-large-volume-databases-using-horizontal-fragmentation (дата звернення 1.08.2021 р).
8. Голуб С.В., Салапатов В.І. Рідкокаша А.А. Програмні агенти і мультиагентні системи. Конспект лекцій. Черкаси. ЧДТУ. 2021. Електронне видання

9. McCreary D. Making Sense of NoSQL: A guide for managers and the rest of us./ D. McCreary, A. Kelly. // Manning Publications. – 2013. – P. 312.

10. Швороб І.Б. Методи та засоби екстракції та аналізу слабоструктурованих текстових даних на основі документо-орієнтовного графа. – Кваліфікаційна наукова праця на правах рукопису. Дисертація на здобуття наукового ступеня кандидата технічних наук (доктора філософії) за спеціальністю 10.21.02 «Структурна, прикладна та математична лінгвістика» – Національний університет «Львівська Політехніка» МОН України, Львів, 2018. [Електронний ресурс] – Режим доступу до ресурсу: https://old.lpnu.ua/research/disscoun/d-3505205/shvorob-iryna-bogdanivna (дата звернення 1.08.2021 р).

13 ІНФОРМАЦІЙНІ РЕСУРСИ

1. https://dbfiddle.uk/ - DataBase Fiddle

2. http://www.niss.gov.ua/ - сайт Національного Інституту стратегічних досліджень (Україна)

3. http://www.aries.eu.int/ - європейські інформаційно-аналітичні центри.

4. http://www.search.global.epnet.com/ - електронний формат наукових журналів

5. https://www.cyberforum.ru/ – Форум програмістів і сисадмінів

6.http://www.fe.msk.ru/win/infomarket/welcome.html - Каталог ресурсів Internet в сфері економіки і фінансів.

7. http://www.ukrstat.gov.ua/ - Державний комітет статистики

8. http://www.itaiep.doc.gov/bisnis/bisnis.html - Інформаційна служба BISNIS

14 ПЕРЕЛІК НОРМАТИВНИХ ДОКУМЕНТІВ

1. ISO/IEC 2382:2015, Information technology — Vocabulary — Part 1: Terms and definitions:
2. ДСТУ ISO/IEC 2382:2017- Інформаційні технології. Словар термінів

3. ДСТУ 3008:2015 ІНФОРМАЦІЯ ТА ДОКУМЕНТАЦІЯ. ЗВІТИ У СФЕРІ НАУКИ І ТЕХНІКИ. СТРУКТУРА ТА ПРАВИЛА ОФОРМЛЮВАННЯ