

**НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ВОДНОГО ГОСПОДАРСТВА ТА ПРИРОДОКОРИСТУВАННЯ**

Навчально-науковий інститут кібернетики, інформаційних технологій та інженерії

04-05-251 S

<b>СИЛАБУС</b> навчальної дисципліни		<b>SYLLABUS</b>	
<b>Технології оброблення великих даних</b>		<b>Big data processing technologies</b>	
Шифр за ОП	<b>ПП 15.1</b>	Code in Degree Programme	
Освітній рівень: <b>магістерський (другий)</b>		Level of Education: <b>master's (second)</b>	
Галузь знань <b>Інформаційні технології</b>	<b>12</b>	Field of Knowledge <b>Information technologies</b>	
Спеціальність <b>Інформаційні системи та технології</b>	<b>126</b>	Field of Study <b>Information systems and technologies</b>	
Освітня програма: <b>Інформаційні технології в бізнесі</b>		Degree Programme: <b>Information technologies in business</b>	

РІВНЕ – 2024

Силабус навчальної дисципліни «Технології оброблення великих даних» для здобувачів вищої освіти другого ступеня «магістр», які навчаються за освітньо-професійною програмою «Інформаційні технології в бізнесі» спеціальності 126 «Інформаційні системи та технології». Рівне. НУВГП. 2024. 11 с.

ОП на сайті університету: <http://ep3.nuwm.edu.ua/id/eprint/30341>

Розробник силабусу: Джоші Олена Іванівна, к.т.н., доцент кафедри комп'ютерних технологій та економічної кібернетики

Силабус схвалений на засіданні кафедри комп'ютерних технологій та економічної кібернетики

Протокол № 1 від «27» серпня 2024 року

Завідувач кафедри: Грицюк Петро Михайлович, д.е.н., професор.

Керівник (гарант) ОП: Барановський Сергій Віталійович, к.т.н., доцент, доцент кафедри комп'ютерних технологій та економічної кібернетики.

Схвалено науково-методичною радою з якості ННІКІТІ

Протокол № 1 від «28» жовтня 2024 року

Голова науково-методичної ради з якості ННІКІТІ: Мартинюк Петро Миколайович, д.т.н., професор.

Попередня версія силабусу (вказати шифр) \_\_\_\_

© О.І.Джоші, 2024

© НУВГП, 2024

<b>ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ</b>
<b>«Технології оброблення великих даних»</b>
<b>ЗАГАЛЬНА ІНФОРМАЦІЯ</b>

Ступінь вищої освіти	Магістр
Освітня програма	Інформаційні технології в бізнесі
Спеціальність	126 Інформаційні системи та технології
Рік навчання, семестр	1-й рік, 2 семестр
Кількість кредитів	4 кредити ЄККТС
Лекції:	18 / 2 годин
Лабораторні заняття:	22 / 10 годин
Самостійна робота:	80 / 108 годин
Курсова робота:	Немає
Форма навчання	Денна / заочна
Форма підсумкового контролю	Залік
Мова викладання	державна або англійська

### ІНФОРМАЦІЯ ПРО РОЗРОБНИКА

Лектор



*Джоші Олена Іванівна, канд.техн.наук,  
доцент кафедри комп'ютерних технологій та економічної кібернетики*

Вікіситет

<https://goo.su/2LEV>

ORCID

<https://orcid.org/0000-0002-1205-0318>

Як комунікувати

[o.i.joshi@nuwm.edu.ua](mailto:o.i.joshi@nuwm.edu.ua)

Актуальні оголошення на сторінці дисципліни в системі MOODLE

### ІНФОРМАЦІЯ ПРО НАВЧАЛЬНУ ДИСЦИПЛІНУ

#### Мета та завдання

**Метою** дисципліни є формування фундаментальних теоретичних знань і практичних навичок щодо ефективного збору, зберігання, аналізу та оброблення великих обсягів даних з використанням сучасних технологій.

**Завданнями** дисципліни є:

- ознайомлення з основними концепціями великих даних: вивчення принципів, архітектури та характеристик великих даних, а також викликів, пов'язаних з їх обробленням;
- засвоєння технологій та інструментів оброблення великих даних: опанування платформ і систем, таких як Hadoop, Spark, NoSQL бази даних та інші інструменти для зберігання, аналізу та оброблення великих даних;
- практичне застосування алгоритмів та моделей: вивчення і реалізація алгоритмів для аналізу та оброблення великих даних, таких як MapReduce, кластеризація, машинне навчання;
- набуття навичок розв'язування задач масштабованості та продуктивності.

**Посилання на розміщення освітнього компонента на навчальній платформі Moodle, на платформі освітніх програм та їхніх освітніх компонентів**

<https://exam.nuwm.edu.ua/course/index.php?categoryid=20>

#### Передумови вивчення\*

(місце навчальної дисципліни в структурно-логічній схемі)

Відповідно до структурно-логічної схеми підготовки магістрів за спеціальністю 126«Інформаційні системи та технології» для вивчення навчальної дисципліни ПП15.1«Технології оброблення великих даних» здобувачам освіти потрібно мати достатній рівень знань і вмінь з навчальних дисциплін ПП2«Нереляційні бази даних».

Вона пов'язана з навчальною дисципліною ПП6«Методи машинного навчання», що одночасно вивчається.

#### **Компетентності**

*ЗК1. Здатність до абстрактного мислення, аналізу та синтезу.*

*СК1. Здатність розробляти та застосувати ІСТ, необхідні для розв'язання стратегічних і поточних задач.*

*СК5. Здатність використовувати сучасні технології аналізу даних для оптимізації процесів в інформаційних системах.*

#### **Програмні результати навчання (ПРН). Результати навчання (РН)\***

*РН1. Відшукувати необхідну інформацію в науковій і технічній літературі, базах даних, інших джерелах, аналізувати та оцінювати цю інформацію.*

*РН8. Розробляти моделі інформаційних процесів та систем різного класу, використовувати методи моделювання, формалізації, алгоритмізації та реалізації моделей з використанням сучасних комп'ютерних засобів.*

*РН9. Розробляти і використовувати сховища даних, здійснювати аналіз даних для підтримки прийняття рішень.*

*РН11. Розв'язувати задачі цифрової трансформації у нових або невідомих середовищах на основі спеціалізованих концептуальних знань, що включають сучасні наукові здобутки у сфері інформаційних технологій, досліджень та інтеграції знань з різних галузей.*

#### **Структура та зміст навчальної дисципліни**

## **ЗМІСТОВИЙ МОДУЛЬ 1. Інструменти оброблення великих даних**

### **Тема 1. Джерела великих даних. Визначення Big Data**

Визначення великих даних. Властивості великих даних. Приклади великих даних у реальному світі. Відкриті дані. Приватність даних. Структуровані та неструктуровані дані. Статичні та динамічні дані. Розподілені дані та їх обробка.

### **Тема 2. Управління даними та життєвий цикл великих даних: фази і ключові аспекти**

Що таке управління даними? Життєвий цикл великих даних: загальне визначення. Основні фази життєвого циклу великих даних. Ключові виклики управління життєвим циклом даних (управління якістю даних, забезпечення безпеки та конфіденційності даних). Перспективи розвитку управління життєвим циклом великих даних.

### **Тема 3. Класифікація та аналіз інструментів оброблення великих даних**

Інструменти для зберігання та управління даними. Інструменти для пакетної обробки великих даних. Інструменти для потокової обробки даних. Інструменти для керування кластерами і контейнерами. Критерії вибору інструментів залежно від типу даних, обсягу та завдань.

### **Тема 4. Парадигма MapReduce**

MapReduce у розподілених обчисленнях. Аспекти роботи MapReduce. Основні етапи MapReduce. Приклади застосування MapReduce.

## **ЗМІСТОВИЙ МОДУЛЬ 2. Фреймворки оброблення великих даних**

### **Тема 5. Фреймворк Apache Hadoop**

Призначення фреймворку Apache Hadoop для обробки великих даних. Архітектура Apache Hadoop. Принципи роботи фреймворку Apache Hadoop. Обробка і управління даними в Hadoop. Масштабованість і продуктивність Hadoop. Інструменти та екосистема Apache Hadoop. Переваги та недоліки Hadoop у порівнянні з іншими фреймворками.

### **Тема 6. Фреймворк Apache Spark**

Призначення фреймворку Apache Spark для обробки та аналізу великих даних. Принципи роботи та складові Apache Spark. Операції Spark. Структури даних Spark. Читання та запис даних в Spark. Робота зі Spark.

### **Тема 7. Потокова обробка даних у реальному часі**

Архітектура та технології потокової обробки даних (Apache Kafka, Apache Flink). Обробка потоків подій в реальному часі. Приклади використання потокової обробки даних у бізнесі.

#### **Теми лабораторних робіт:**

1. Джерела великих даних.
2. Життєвий цикл великих даних.
3. Інструментів оброблення великих даних.
4. Парадигма MapReduce.
5. Фреймворк Apache Hadoop.
6. Фреймворк Apache Spark.
7. Потокова обробка даних у реальному часі.

#### **Теми самостійних робіт:**

Отримання сертифікату за результатами неформальної освіти на освітніх платформах онлайн-курсів (англ. MOOC). Зміст онлайн-курсу має відповідати змісту навчальної дисципліни, він повинен мати приблизно обсяг 10±2 годин.

## **БЮДЖЕТ ЧАСУ ТА ШКАЛА ОЦІНЮВАННЯ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ ЗА ФОРМАМИ НАВЧАННЯ**

Теми

Денна форма навчання

Заочна форма навчання

	Обсяг, год.				Балів	Обсяг, год.				Балів
	Всього					Всього				
		лекції	лабораторні	сам. роб.			лекції	лабораторні	сам. роб.	
<b>Модуль 1</b>										
Тема1	14	2	2	10	10	13	1	-	12	10
Тема 2	14	2	2	10	5	17	-	2	15	5
Тема 3	25	4	6	15	10	17	-	2	15	10
Тема 4	14	2	2	10	5	14		2	12	5
<b>Модульний контроль 1</b>	-	-	-	-	<b>20</b>	-	-	-	-	<b>20</b>
<b>Всього</b>	<b>67</b>	<b>10</b>	<b>12</b>	<b>45</b>	<b>50</b>	<b>61</b>	<b>1</b>	<b>6</b>	<b>54</b>	<b>50</b>
<b>Модуль 2</b>										
Тема5	22	3	4	15	10	21	1	2	18	10
Тема6	17	3	4	10	10	20	-	2	18	10
Тема7	14	2	2	10	10	18	-	-	18	10
<b>Модульний контроль 2</b>	-	-	-	-	<b>20</b>	-	-	-	-	<b>20</b>
<b>Всього</b>	<b>53</b>	<b>8</b>	<b>10</b>	<b>35</b>	<b>50</b>	<b>59</b>	<b>1</b>	<b>4</b>	<b>54</b>	<b>50</b>
<b>Разом</b>	<b>120</b>	<b>18</b>	<b>22</b>	<b>80</b>	<b>100</b>	<b>120</b>	<b>2</b>	<b>10</b>	<b>108</b>	<b>100</b>

### СТРУКТУРА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ У РОЗРІЗІ РЕЗУЛЬТАТІВ НАВЧАННЯ

Теми	Програмні результати навчання (ПРН, РН*)				Методи, технології, засоби навчання
	1	8	9	11	
<b>МОДУЛЬ 1</b>					
<b>ЗМІСТОВИЙ МОДУЛЬ 1</b>					
Тема1	+				Лекції, лабораторні заняття за індивідуальними варіантами, самостійна робота; Платформа Moodle, Google сервіси, навчальні посібники, наукові публікації, тощо.
Тема 2		+		+	
Тема 3		+	+	+	
Тема 4		+		+	
Модульний контроль 1	+	+	+	+	Тестові завдання
<b>ЗМІСТОВИЙ МОДУЛЬ 2</b>					
Тема5	+		+	+	Лекції, лабораторні заняття за індивідуальними варіантами, самостійна робота; Платформа Moodle, Google сервіси, навчальні посібники, наукові публікації, тощо.
Тема6	+	+		+	
Тема7	+	+		+	
Модульний контроль 2	+	+	+	+	Тестові завдання

### Форми та методи навчання

**Методи навчання:** лекції, лабораторні заняття за індивідуальними варіантами, самостійна робота; тестові завдання.

**Технології навчання:** платформа Moodle, Google сервіси, навчальні посібники, наукові публікації, тощо.

### Інструменти, обладнання, програмне забезпечення

Для виконання лабораторних та самостійних робіт передбачено використання програмного забезпечення з відкритим вихідним кодом.

### Порядок оцінювання програмних результатів навчання/результатів навчання

Для досягнення мети та завдань курсу здобувачам освіти потрібно засвоїти теоретичний матеріал та здати поточні модульні контролю знань, а також вчасно виконати лабораторні і самостійні роботи. В результаті вони зможуть отримати такі обов'язкові бали:

**60** балів – поточна складова:

50 балів – виконання лабораторних і самостійних робіт, опитування та дискусії під час лекційних занять;

10 балів – захист лабораторних і самостійних робіт англійською мовою;

**20** балів – модульний контроль 1;

**20** балів – модульний контроль 2.

Вивчення навчальної дисципліни закінчується **заліком**.

Усього **100** балів.

Модульні контролю відбуваються у формі тестування. Загальна кількість питань по кожному з модулів – мінімум 150, з них у тесті 30 питань з п'ятьма варіантами відповідей з наступним розподілом балів:

1-й рівень: 20 питань по 0,4 бали кожне, всього 8 балів;

2-й рівень: 9 питань по 1 балу кожне, всього 9 балів;

3-й рівень: 1 питання по 3 бали, всього 3 бали;

Разом 20 балів.

Оцінювання результатів навчання проводиться у % від кількості балів, виділених на завдання:

0 % – завдання не виконано;

40 % – завдання виконано частково та містить суттєві помилки методичного або розрахункового характеру;

60 % – завдання виконано повністю, але містить суттєві помилки у розрахунках або в методиці;

80 % – завдання виконано повністю і вчасно, проте містить окремі несуттєві недоліки (висновки, оформлення тощо);

100 % – завдання виконано правильно, вчасно і без зауважень.

### Рекомендована література (основна, допоміжна)

#### ОСНОВНА

1. Талах М.В. Технології обробки Big Data. Навчальний посібник. Чернівці : Чернівецький нац. ун-т, 2024. 454 с.
2. Олещенко Л.М. Технології оброблення великих даних: конспект лекцій. Київ : КПІ ім. Ігоря Сікорського, 2021. 227 с.
3. Ланде Д.В., Субачі Ю., Гладуна Я. Оброблення надвеликих масивів даних (Big Data): навчальний посібник. Київ : ІСЗЗІ КПІ ім. Ігоря Сікорського, 2021. 168 с.

#### ДОПОМІЖНА

1. Таран В.І., Гордієнко Ю.Г., Стіренко С.Г. Технології Big Data: практикум. Навчальний посібник. Київ : КПІ ім. Ігоря Сікорського, 2022. 56 с.
2. Олещенко Л.М. Технології оброблення великих даних: комп'ютерний практикум. Київ : КПІ ім. Ігоря Сікорського, 2021. 85 с.
3. Marz N., Warren J. Big Data: Principles and Best Practices of Scalable Real-time Data Systems. MANNING: Shelter Island. 2015. 330 p.
4. Bhalla R. Introduction to Big Data. LPU. 2023. 297 p.
5. Buaya R., Calheiros R., Dastjerdi A. Big Data Principles and Paradigms. Morgan Kaufmann. 2016. 496 p.
6. Prajapati V. Big Data Analytics with R and Hadoop. [Packt Publishing] open source. Birmingham – Mumbai. 2013. 238 p.

#### Інформаційні ресурси в Інтернет

1. Цифрові сервіси для освіти України. Платформи <https://mooc4ua.online/platforms>

#### Поєднання навчання та досліджень

Здобувачі другого (магістерського) рівня вищої освіти мають можливість додатково отримати бали в межах поточної складової за виконання індивідуальних завдань дослідницького характеру, можуть бути долучені до опублікування наукових статей і тез з тематики навчальної дисципліни, участі в науково-практичних конференціях ІТ спрямування.

## ПОЛІТИКИ ВИКЛАДАННЯ ТА НАВЧАННЯ

### Перелік соціальних, «м'яких» навичок (soft skills)

Вміння комунікувати з людьми, критичне мислення, загальнокультурна грамотність, здатність до навчання, самоорганізація, самоосвіта, самовдосконалення, здатність логічно обґрунтовувати позицію, формування власної думки та інше.

### Дедлайни та перескладання

Завдання навчальної дисципліни повинні бути виконані до початку екзаменаційної сесії. Завдання є своєчасно виконаним протягом тижня від дня проведення лабораторного заняття. У разі відсутності на занятті незалежно від причини здобувач освіти зобов'язаний самостійно виконати завдання відповідно до індивідуального варіанту завдання. Звіти лабораторних робіт розміщуються на навчальній платформі Moodle.

Допускається у порядку, встановленому НУВГП, вільне відвідування студентами навчальних занять відповідно до [Порядку про вільне відвідування студентами навчальних занять в НУВГП](#)

[Порядок ліквідації академічних заборгованостей у НУВГП](#)

### Неформальна та інформальна освіта

Визнання (перезарахування) результатів навчання, здобутих шляхом неформальної та інформальної освіти, відбувається відповідно до [Положення про неформальну та інформальну освіту в НУВГП](#).

Здобувачі освіти можуть пройти відкриті онлайн курси, близькі за темами, передбаченими даною навчальною дисципліною, на таких освітніх платформах як Prometheus, Coursera, edEx, edEra, FutureLearn тощо. При цьому важливо, щоб знання та навички, що формуються під час проходження певного онлайн-курсу чи його частин, мали зв'язок з результатами навчальної дисципліни та мали можливість бути перевірені в підсумковому оцінюванні. Перед початком проходження обраних курсів необхідно їх вибір узгодити з викладачем.

### Правила академічної доброчесності

- Здобувачі освіти та викладач несуть спільну відповідальність за створення сприятливого творчого навчального середовища, яке базується на взаємній повазі: здобувачі освіти повинні дотримуватися [Кодексу честі студентів](#), а викладач – [Етичного кодексу викладача НУВГП](#)
- Здобувачі освіти мають самостійно виконувати і здавати на оцінювання результати власних зусиль та оригінальної праці
- Здобувачі освіти мають дотримуватися [Положення про виявлення та запобігання академічного плагіату в НУВГП](#)
- Лабораторні та самостійні роботи виконуються здобувачами освіти за індивідуальними варіантами із дотриманням вимог [Положення про академічну доброчесність у НУВГП](#)
- У випадку виявлення плагіату лабораторних або самостійних робіт здобувач освіти не отримує бали і має виконати завдання повторно.
- Порядок проведення семестрового поточного та підсумкового контролів здійснюється відповідно до [Положення про семестровий поточний та підсумковий контроль навчальних досягнень здобувачів вищої освіти](#)

### Вимоги до відвідування

- Заняття відбуваються згідно [роzkладу](#) офлайн або онлайн за допомогою [Google Meet](#)
- Консультації проводяться за потреби в режимі онлайн за допомогою [Google Meet](#) за графіком узгодженим із навчальними групами.
- Здобувачам освіти не дозволяється пропускати заняття без поважних причин.
- У випадку пропусків занять з поважних або неповажних причин здобувачі освіти мають самостійно опрацювати теоретичний матеріал і виконати лабораторні та самостійні роботи за індивідуальними варіантами відповідно до вимог [Положення про академічну доброчесність у НУВГП](#).
- Студент має право оформити індивідуальний графік навчання згідно [Положення про індивідуальний графік навчання студентів денної форми навчання НУВГП](#)
- Здобувачі освіти можуть під час занять використовувати ноутбуки, планшети, смартфони, тощо, але виключно в навчальних цілях.

Лектор,  
к.т.н., доцент кафедри комп'ютерних технологій та  
економічної кібернетики

О.І.Джоші

Автор  
Доцент кафедри комп'ютерних технологій та  
економічної кібернетики

Олена ДЖОШІ

Затверджено

Проректор з науково-педагогічної та навчальної  
роботи

Валерій СОРОКА



документ підписаний КЕП  
Номер документа СИЛ №1373  
Підписувач Сорока Валерій Степанович  
Підписувач (дані КЕП):  
Сертифікат 3FAA9288358EC00304000009B6C3700C8C2C100