

	СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ «Технології обробки великих даних» Рівень вищої освіти: <u>Другий (магістерський)</u> Спеціальність: <u>ЕЗ Комп'ютерні науки</u> Рік навчання: <u>1-й, семестр 1-й</u> Кількість кредитів ECTS: <u>4 кредитів</u> Назва кафедри: <u>Комп'ютерних наук та цифрової економіки</u> Мова викладання: <u>українська</u>
Лектор курсу	к.т.н, старший викладач Титарчук Євгеній Олександрович
Контактна інформація лектора (e-mail)	etitarchuk@gmail.com

ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Навчальна дисципліна «Технології обробки великих даних» є вибірковою компонентою ОПП.

Загальний обсяг дисципліни 120 год.: лекції – 24 год.; практичні заняття – 0 год., самостійна робота – 74 год.

Формат проведення: лекції, практичні заняття, консультації.

Підсумковий контроль – залік.

Основні положення навчальної дисципліни можуть застосовуватися під час вивчення таких дисциплін: «Інтелектуальні методи та моделі в Big Data», «Серверне програмування», .

Призначення навчальної дисципліни

Дисципліна «Технології обробки великих даних» спрямована надати здобувачам здатність проектувати архітектуру розподілених систем зберігання й опрацювання великих обсягів даних, застосовувати паралельні обчислення, документоорієнтовані бази даних і хмарні платформи, забезпечувати узгодженість, надійність та ефективність потоків даних.

Мета вивчення навчальної дисципліни

Метою навчальної дисципліни «Технології обробки великих даних» є формування у здобувачів освіти системи теоретичних знань і практичних навичок, необхідних для проєктування, розроблення та супроводу рішень для збирання, зберігання, передавання й аналізу великих даних, а саме: ознайомлення з принципами архітектури розподілених систем, моделями

паралельної обробки, екосистемами для роботи з великими даними, механізмами когерентності кешу, документоорієнтованими базами даних, хмарними платформами, інженерією даних і операційним супроводом процесів, а також розвиток здатності розв'язувати комплексні практичні завдання та оцінювати сучасні тенденції галузі.

Завдання вивчення дисципліни

Завданням дисципліни є ознайомлення з теоретичними засадами та практичними підходами до обробки великих даних; формування знань про властивості, джерела, обсяги та особливості таких даних, архітектуру розподілених систем зберігання на прикладі розподіленої файлової системи GFS, модель MapReduce та її реалізації, а також екосистему Hadoop і засоби Apache Spark. Курс передбачає вивчення принципів розподіленого зберігання й опрацювання даних, механізмів узгодженості кешу, особливостей документоорієнтованих баз даних, можливостей хмарних платформ для роботи з великими даними, підходів до інженерії даних та операційного управління процесами їх обробки. Особливу увагу приділено проектуванню надійних, масштабованих і ефективних рішень, організації потоків даних, автоматизації обробки та забезпеченню якості результатів. Курс спрямований на підготовку студентів до практичного застосування технологій великих даних під час розв'язання прикладних завдань, а також на розвиток їхніх аналітичних, системних і проблемно-розв'язувальних навичок з урахуванням перспектив розвитку галузі та сучасних тенденцій.

ПЕРЕЛІК КОМПЕТЕНТНОСТЕЙ, ЯКИХ НАБУВАЄ ЗДОБУВАЧ ПРИ ВИВЧЕНІ ДИСЦИПЛІНИ ВІДПОВІДНО ДО ОСВІТНЬОЇ ПРОГРАМИ

У результаті вивчення навчальної дисципліни здобувач повинен сформувати такі програмні компетентності:

інтегральну компетентність (ІК):

Здатність розв'язувати задачі дослідницького та/або інноваційного характеру у сфері комп'ютерних наук.

загальні компетентності (ЗК):

ЗК1. Здатність до абстрактного мислення, аналізу та синтезу.

ЗК2. Здатність застосовувати знання у практичних ситуаціях.

спеціальні (фахові, предметні) компетентності (СК):

СК1. Усвідомлення теоретичних засад комп'ютерних наук.

СК5. Здатність розробляти, описувати, аналізувати та оптимізувати архітектурні рішення інформаційних та комп'ютерних систем різного

призначення.

СК7. Здатність розробляти програмне забезпечення відповідно до сформульованих вимог з урахуванням наявних ресурсів та обмежень.

СК9. Здатність розробляти та адмініструвати бази даних та знань.

ПРОГРАМНІ РЕЗУЛЬТАТИ НАВЧАННЯ ВІДПОВІДНО ДО ОСВІТНЬОЇ ПРОГРАМИ

РН9. Розробляти алгоритмічне та програмне забезпечення для аналізу даних (включно з великими).

РН10. Проектувати архітектурні рішення інформаційних та комп'ютерних систем різного призначення.

РН12. Проектувати та супроводжувати бази даних та знань.

Вивчення даної дисципліни формує у здобувачів освіти соціальні навички (soft skills): комунікативність (реалізується через: метод роботи в парах та групах, робота з інформаційними джерелами), робота в команді (реалізується через: метод проєктів), лідерські навички (реалізується через: робота в групах, метод проєктів, метод самопрезентації).

ПЛАН ВИВЧЕННЯ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

№ з/п	Назви теми	Форми організації навчання та кількість годин		Самостійна робота, кількість годин
		лекційні заняття	практичні заняття	
3-й семестр				
1	Тема 1. Вступ до великих даних	2	0	6
2	Тема 2. Архітектура розподілених систем зберігання на прикладі GFS	2	0	8
3	Тема 3. Модель MapReduce та її реалізації	4	0	8
4	Тема 4. Екосистема Hadoop	2	0	8
5	Тема 5. Когерентність кеша	2	0	6
6	Тема 6. Apache Spark	4	0	10
7	Тема 7. Документо-орієнтовані бази даних	2	0	8
8	Тема 8. Хмарні платформи для Big Data	2	0	6
9	Тема 9. Data Engineering та DataOps	2	0	8
10	Тема 10. Перспективи та сучасні тренди	2	0	6
Разом		24	0	74

Самостійна робота здобувача вищої освіти

Самостійна робота здобувача організовується шляхом видачі індивідуального переліку питань і практичних завдань з кожної теми, які не виносяться на аудиторне опрацювання та виконання індивідуального творчого завдання.

Самостійна робота здобувача є одним із способів активного, цілеспрямованого набуття нових для нього знань та умінь. Вона є основою його підготовки як фахівця, забезпечує набуття ним прийомів пізнавальної діяльності, інтерес до творчої роботи, здатність вирішувати наукові та практичні завдання.

Виконання здобувачем самостійної роботи передбачає, за необхідності, отримання консультацій або допомоги відповідного фахівця. Навчальний матеріал навчальної дисципліни, передбачений робочою програмою для засвоєння здобувачем у процесі самостійної роботи, виносяться на поточний і підсумковий контроль поряд з навчальним матеріалом, який опрацьовувався під час аудиторних занять. Організація самостійної роботи здобувачів передбачає: планування обсягу, змісту, завдань, форм і методів контролю самостійної роботи, розробку навчально-методичного забезпечення; виконання здобувачем запланованої самостійної роботи; контроль та оцінювання результатів, їх систематизацію, оцінювання ефективності виконання здобувачем самостійної роботи.

Індивідуальні завдання здобувач виконує самостійно під керівництвом

викладача згідно з індивідуальним навчальним планом.

У випадку реалізації індивідуальної освітньої траєкторії здобувача заняття можуть проводитись за індивідуальним графіком.

Види самостійної роботи

№ п/п	Вид самостійної роботи	Години	Термін виконання	Форма та метод контролю
1	Підготовка самостійних питань з тематики дисципліни (опрацювання теоретичних основ прослуханого лекційного матеріалу)	44	щотижнево / під час заліково-екзаменаційно ї сесії	Усне та письмове опитування, тестування, оцінювання конспекту
2	Підготовка до лекційних та практичних занять (робота з інформаційними джерелами: опрацювання першоджерел)	10	щотижнево	Усне опитування, оцінювання конспекту
3	Індивідуальні творчі практичні завдання (вирішення і письмове оформлення завдань, інших робіт графічного характеру; презентації за заданою проблемною тематикою)	10	1 раз на 2 тижні / під час заліково-екзаменаційно ї сесії	Спостереження за виконанням, обговорення, виступ з презентацією
4	Підготовка до контрольних робіт та тестування, виконання завдань різних видів контролю під час практичних занять (самостійне опрацювання тестів відповідно до теми практичного заняття; самостійне розв'язання практичних задач, ситуаційних вправ)	10	щотижнево / під час заліково-екзаменаційно ї сесії	Тестування у системі Moodle
Разом		74		

РЕКОМЕНДОВАНІ ДЖЕРЕЛА ІНФОРМАЦІЇ

Основні

1. Ghemawat S., Gobioff H., Leung S.-T. The Google File System. Proceedings of the 19th ACM Symposium on Operating Systems Principles. 2003. P. 20–43. URL: <https://research.google/pubs/the-google-file-system/>

2. Dean J., Ghemawat S. MapReduce: Simplified Data Processing on Large Clusters. Proceedings of the 6th Symposium on Operating Systems Design and Implementation (OSDI '04). 2004. P. 137–150. URL: <https://research.google/pubs/mapreduce-simplified-data-processing-on-large-clusters/>

3. Apache Hadoop Documentation / Apache Software Foundation. URL: <https://hadoop.apache.org/docs/stable/>

4. Apache Spark Documentation. Apache Software Foundation.

URL: <https://spark.apache.org/docs/latest/>

5. MongoDB Manual. MongoDB.

URL: <https://www.mongodb.com/docs/manual/>

Додаткові

1. Apache Airflow Documentation. Apache Software Foundation.

URL: <https://airflow.apache.org/docs/> (дата звернення: 09.08.2026).

2. Bigtable: A Distributed Storage System for Structured Data / F. Chang et al. Proceedings of the 7th USENIX Symposium on Operating Systems Design and Implementation (OSDI'06). 2006. P. 205–218.

URL: <https://research.google/pubs/bigtable-a-distributed-storage-system-for-structured-data/>

3. Resilient Distributed Datasets: A Fault-Tolerant Abstraction for In-Memory Cluster Computing / M. Zaharia et al. Proceedings of the 9th USENIX Symposium on Networked Systems Design and Implementation (NSDI '12). 2012. P. 15–28. URL: <https://www.usenix.org/conference/nsdi12/technical-sessions/presentation/zaharia>

4. Scaling Memcache at Facebook / R. Nishtala et al. Proceedings of the 10th USENIX Symposium on Networked Systems Design and Implementation (NSDI '13). 2013. P. 385–398. URL: <https://www.usenix.org/conference/nsdi13/technical-sessions/presentation/nishtala> (дата звернення: 09.08.2026).

5. Lakehouse: A New Generation of Open Platforms that Unify Data Warehousing and Advanced Analytics / M. Armbrust, A. Ghodsi, R. Xin, M. Zaharia. Proceedings of the 11th Conference on Innovative Data Systems Research (CIDR '21). 2021. URL: https://www.cidrdb.org/cidr2021/papers/cidr2021_paper17.pdf

Інформаційні ресурси

1. Apache Hadoop Documentation / Apache Software Foundation.

URL: <https://hadoop.apache.org/docs/stable/>

2. Apache Spark Documentation / Apache Software Foundation.

URL: <https://spark.apache.org/docs/latest/>

3. MongoDB Manual / MongoDB.

URL: <https://www.mongodb.com/docs/manual/>

4. Apache Airflow Documentation / Apache Software Foundation.

URL: <https://airflow.apache.org/docs/>

СИСТЕМА ОЦІНЮВАННЯ ТА ВИМОГИ ДО КОНТРОЛЮ ЗНАНЬ ЗДОБУВАЧІВ ВИЩОЇ ОСВІТИ

У кінці семестру, здобувач вищої освіти може набрати до 60% підсумкової оцінки за виконання всіх видів робіт, що виконуються протягом семестру, до 10% за показники наукової, інноваційної, навчальної, виховної роботи та студентської активності і до 30% підсумкової оцінки – за результатами підсумкового контролю.

Розподіл балів за видами навчальної діяльності

№ п.п.	Вид навчальної діяльності	Бали
Атестація 1.		
1	Участь у дискусіях на лекційних заняттях	6
2	Участь у роботі на практичних заняттях	0
3	Виконання контрольних робіт, тестування	5
4	Підготовка самостійних питань з тематики дисципліни	19
	Всього за атестацію 1	30
Атестація 2.		
1	Участь у дискусіях на лекційних заняттях	6
2	Участь у роботі на практичних заняттях	0
3	Виконання контрольних робіт, тестування	5
4	Підготовка самостійних питань з тематики дисципліни	19
	Всього за атестацію 2	30
	Показники наукової, інноваційної, навчальної, виховної роботи та студентської активності	10
	Підсумкове тестування	30
Разом		100

У кінці семестру, здобувач вищої освіти може набрати до 60% підсумкової оцінки за виконання всіх видів робіт, що виконуються протягом семестру, до 10% за показники наукової, інноваційної, навчальної, виховної роботи та студентської активності і до 30% підсумкової оцінки – за результатами підсумкового контролю.

Якщо здобувач упродовж семестру за підсумками контрольних заходів набрав (отримав) менше половини максимальної оцінки з навчальної дисципліни (менше 35 балів), то він не допускається до заліку. Крім того, обов'язковим при мінімальній кількості балів за підсумками контрольних заходів є виконання індивідуальної творчої роботи (презентації).

Під час виконання навчальних завдань, а також завдань поточних та підсумкових контрольних заходів не допустимо порушення академічної доброчесності. Презентації та виступи мають бути авторськими та оригінальними, інформація про результати власної навчальної (наукової,

творчої) діяльності – достовірною; у разі використання ідей, розробок, тверджень, відомостей мають бути посилення на джерела інформації з дотриманням норм законодавства про авторське право і суміжні права.

Програма навчальної дисципліни передбачає врахування результатів неформальної та інформальної освіти при наявності підтверджуючих документів як окремі кредити вивчення навчальних дисциплін.

Переведення балів внутрішньої 100-бальної шкали в національну здійснюється у відповідності до шкали.

Відповідність шкал оцінок якості засвоєння навчального матеріалу

Сума балів за всі види навчальної діяльності	Оцінка ECTS	Оцінка за національною шкалою	
		для екзамену	для заліку
90-100	A	відмінно	зараховано
82-89	B	добре	
75-81	C		
66-74	D	задовільно	
60-65	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