



Національний
транспортний
університет

ВК Інженерія інформаційно-пошукових систем

Рівень вищої освіти – *другий (магістерський)*

Дні занять, час занять, аудиторія: будуть доступні згідно розкладу за посиланням <https://kizt.ntu.edu.ua/to-the-student/rozklady-zanyat-i-sesij/>

Кафедра інформаційних систем і технологій

Лекційні заняття проводить

к.ф.-м.н., доц. Вітер Михайло Богданович

Контактна інформація	mbviter@gmail.com +380 (66) 941-39-05
----------------------	---

Адреса, номер аудиторії	м. Київ, вул. Івана Огієнка, 19, каб. 601a
-------------------------	--

Час консультацій	Щовівторка з 15:00 до 16:00
------------------	-----------------------------

Практичні заняття проводить

к.ф.-м.н., доц. Вітер Михайло Богданович

Контактна інформація	mbviter@gmail.com +380 (66) 941-39-05
----------------------	---

Аудиторія	м. Київ, вул. Івана Огієнка, 19, каб. 601a
-----------	--

Час консультацій	Щовівторка з 15:00 до 16:00
------------------	-----------------------------

Анотація курсу: «Інженерія інформаційно-пошукових систем» - вибіркова навчальна дисципліна, спрямована на формування інженерних умінь проєктувати, реалізовувати та оцінювати інформаційно-пошукові системи. У курсі розглядаються колекції документів, аналіз тексту, інвертовані індекси, моделі відповідності, API пошуку, оцінювання релевантності, продуктивність і гібридні сценарії пошуку.

Практичний фокус дисципліни: Apache Lucene, Elasticsearch або OpenSearch; підготовка колекції, налаштування аналізаторів, побудова індексу, реалізація запитів і пошукового API, розрахунок метрик precision/recall/MAP/MRR/nDCG, профілювання затримки та документування інженерних рішень.

Предметом вивчення є методи та інструменти інженерії інформаційно-пошукових систем: формування колекцій і вимог, аналіз тексту, індексація, пошук за полями, ранжування, фасетна навігація, пошукові API, оцінювання релевантності, масштабування та супровід індексу. Обсяг дисципліни - 3 кредити ЄКТС (90 годин): денна форма - 9 годин лекцій, 18 годин практичних занять і 63 години самостійної роботи; заочна форма - 2 години лекцій, 2 години практичних занять і 86 годин самостійної роботи.

Міждисциплінарні зв'язки: дисципліна спирається на результати попереднього навчання з програмування, структур даних, баз даних, комп'ютерних мереж та аналізу вимог до програмного забезпечення.

Програма навчальної дисципліни складається з таких модулів:

Модуль 1. Колекції, індексація та ранжування

Тема 1. Предметна область, колекції та архітектура пошукової системи.

Інформаційна потреба, документ, запит і релевантність; типи колекцій та метаданих; компоненти пошукового контуру; життєвий цикл документа від надходження до видачі; вимоги до повноти, затримки й пояснюваності результатів; межі застосування повнотекстового пошуку.

Тема 2. Аналіз тексту, токенизація та побудова інвертованого індексу.

Кодування і нормалізація тексту; токенизація, лематизація та стоп-слова; поля, позиції й частоти термінів; структура інвертованого індексу та postings; оновлення, сегменти й злиття; оброблення українськомовних і багатомовних документів.

Модуль 2. Пошукові сервіси та оцінювання якості

Тема 3. Моделі відповідності та ранжування результатів.

Булевий, векторний і ймовірнісний підходи; TF-IDF та BM25; ваги полів, фразові й нечіткі запити; сортування, фільтри та фасети; підсвічування збігів; компроміс між повнотою, точністю і затримкою.

Тема 4. Запити, фільтрація, фасетна навігація та пошуковий API.

Структура пошукового запиту й DSL; відповідність полям і типам даних; фільтри без зміни релевантності; фасети та агрегації; підказки, автодоповнення й виправлення помилок; пагінація, сортування та підсвічування; вимоги до REST-інтерфейсу пошуку.

Тема 5. Оцінювання якості пошуку та експериментальне налаштування.

Тестова колекція, інформаційні потреби та judgments; precision, recall, F1, MAP, MRR, nDCG і recall@k; офлайн та інтерактивне оцінювання; побудова базової лінії; аналіз помилок і нерелевантних результатів; відтворюваність експерименту.

Тема 6. Масштабування, кешування та гібридний пошук.

Інкрементальна індексація та конвеєри оновлення; сегментація колекції; кешування запитів і результатів; вимірювання затримки та пропускну здатності; відмовостійке оновлення індексу; поєднання лексичного BM25-пошуку з векторним пошуком як розширений сценарій; контроль пояснюваності й вартості.

Для стимулювання науково-дослідницького й творчого інтересу здобувачів вищої освіти і здобуття ними навичок наукової діяльності доречно заохочувати їх до виконання додаткових видів робіт в певних проєктах, і враховувати цю діяльність при підведенні підсумків роботи здобувачів вищої освіти у семестрі. Так, наприклад, здобувачам вищої освіти може бути запропоновано виконання рефератів за темами дисципліни у вигляді презентацій та захистити її на практичному занятті. Найкращі рекомендуються до виступу на науково-практичній конференції.

Індивідуальні види робіт не обмежуються і приймаються пропозиції щодо вдосконалення курсу (як лекційної, так і практичної складової).

Методи контролю: Методи поточного контролю: виконання практичних, ситуаційних, тестових та інших завдань, опитування за темами, короткі виступи та презентації з тематики дисципліни.

Методи модульного контролю: письмова контрольна робота, підсумкове тестування.

Джерела для вивчення курсу –

1. Силабус навчальної дисципліни.

2. Електронний ресурс бібліотеки НТУ: <http://lib.ntu.edu.ua/catalog/login.html>

Віртуальне середовище навчання (Google classroom):

Оцінювання

Контроль протягом семестру (60 балів)					Залік	Сума балів
Модуль 1		Модуль 2		Модуль 3 (ІЗ)		
Тема 1	Тема 2	Тема 3	Теми 4–6			
Для денної форми здобуття вищої освіти: - практичні завдання за темами 1–6 - по 5 балів (усього 30); - модульна контрольна робота № 1 - 15 балів; - модульна контрольна робота № 2 - 15 балів; - поточний контроль разом - 60 балів.					40	100
Для заочної форми здобуття вищої освіти: - письмові/дистанційні завдання за темами 1–6 - по 5 балів (усього 30); - контрольна робота - 30 балів; - поточний контроль разом - 60 балів.						

Самостійна робота

Написання та захист реферату (тематика погоджується із викладачем курсу) у вигляді доповіді та/або презентації, написання тез, статті, участь у міжнародних, всеукраїнських та/або інших заходах та/або конкурсах (за тематикою навчальної дисципліни) оцінюються додатково до 10 балів.

Підготування до семінарських занять та контрольних заходів здійснюється під час самостійної роботи здобувачів з можливістю консультування з викладачем у визначений час консультацій або за допомогою електронного листування (електронна пошта, месенджери).

Додаткові бали до поточного контролю здобувач освіти може отримати, пройшовши навчальний курс у вигляді неформальної освіти з отриманням сертифікату в межах предмету вивчення дисципліни та пройшовши процедуру визнання згідно Тимчасове положення про порядок визнання результатів навчання, набутих студентами Національного транспортного університету у неформальній/інформальній освіті

Критерії оцінювання

Підсумкові бали з навчальної дисципліни визначаються як сума балів, отриманих здобувачем протягом семестру відповідно Додатку 1 до Положення про організацію освітнього процесу у Національному транспортному університеті

Політика несвоєчасного проходження контрольних заходів.

Поточний та підсумковий контролю проводяться згідно з графіком освітнього процесу та встановленими Науково-методичною радою НТУ графіками. У випадку неявки здобувача вищої освіти на підсумковий контроль з поважних причин можливе індивідуальне проведення контролю в узгоджений з викладачем термін за наявності дозволу директорату Навчально-наукового інституту управління, технологій та правових наук.

Повторне складання екзамену у випадку отримання незадовільної оцінки допускається не більше двох разів: один раз – викладачу, другий – комісії, яка створюється з директором Навчально-наукового інституту управління, технологій та правових наук.

При поданні індивідуального завдання пізніше встановленого терміну без поважної причини оцінка буде знижена на 10 %. Технічні проблеми (поломка обладнання, проблеми з друком) не є поважною причиною для несвоєчасного подання індивідуального завдання.

Політика перескладання. Упродовж тижня після оголошення результатів поточного контролю здобувач освіти може звернутися до оцінювача за роз'ясненням і/або з незгодою щодо отриманої оцінки. У випадку незгоди з рішенням оцінювача щодо результатів підсумкового контролю здобувач освіти може звернутися до оцінювача з незгодою щодо отриманої оцінки у день її оголошення. Повторне проходження підсумкового контролю з метою підвищення позитивної оцінки не допускається.

Політика відвідування та / або активності. Відвідування навчальних занять є обов'язковим для здобувача освіти. Невиконання здобувачем освіти завдань, що визначені індивідуальним навчальним планом практичних занять, через відсутність на заняттях є підставою для прийняття рішення про недопущення до підсумкового контролю. За рішенням директорату Навчально-наукового інституту управління, технологій та правових наук передбачається надання можливості виконати пропущені завдання за індивідуальним графіком (але не пізніше, ніж до завершення підсумкового контролю).

Академічна доброчесність.

Дотримання академічної доброчесності засновується на ряді положень та принципів академічної доброчесності, що регламентують діяльність здобувачів вищої освіти та викладачів НТУ:

Кодекс етики академічних взаємовідносин та доброчесності Національного транспортного університету

Положення про систему забезпечення академічної доброчесності педагогічними, науково-педагогічними, науковими працівниками та здобувачами вищої освіти в Національному транспортному університеті

Порушеннями академічної доброчесності є:

- академічний плагіат;
- фальсифікація;
- списування;
- обман;
- неправомірна вигода;
- хабарництво.

Порушення Кодексу етики академічних взаємовідносин та доброчесності Національного транспортного університету є серйозним порушенням, навіть якщо воно є ненавмисним.

При проходженні контролю (поточного або підсумкового) особа, яка проходить контроль, не має права використовувати будь-яку зовнішню (сторонню) допомогу. Якщо оцінювач підозрює особу, що проходить контроль, у використанні недозволених допоміжних засобів, він має право запропонувати їй вчинити дії, які б спростували підозру. У разі відмови від вчинення дій зі спростування підозри, списування, використання недозволених допоміжних засобів чи зовнішньої допомоги (обману) результат оцінюється як «незадовільно».

Поведінка в аудиторії. Ноутбуки та портативні пристрої можна використовувати ВИКЛЮЧНО з навчальною метою за вказівкою викладача. Неправомірне використання ноутбуків чи портативних пристроїв вважатиметься порушенням дисципліни, викладач має право ініціювати дії з унеможливлення їх використання.

В аудиторії забороняється вживання їжі, напоїв (за винятком води). Здобувачі вищої освіти та викладачі повинні дотримуватися етичних норм поведінки.

Здобувачам вищої освіти з обмеженими можливостями або особливими потребами слід до початку семестру звернутися до директорату Навчально-наукового інституту управління, технологій та правових наук та обговорити питання організації навчання.

При виникненні у здобувача вищої освіти проблем зі здоров'ям, які можуть заважати навчанню, слід звернутися до медичного закладу та повідомити про це директорат Навчально-

наукового інституту управління, технологій та правових наук.

Скарги, пропозиції, зауваження та повідомлення про наявність конфліктних ситуацій в рамках освітніх програм здобувачі можуть надсилати на електронну адресу general@ntu.edu.ua або скористатися скринькою довіри, яка розміщена при вході в університет.

Електронна скринька для звернень до психологічної служби: philosophy@ntu.edu.ua.

Зв'язок з викладачем: e-mail: mbviter@gmail.com; тел.: +380 (66) 941-39-05.

Рекомендована література:

1. Apache Software Foundation. Apache Lucene Core 10.5.1: documentation and release resources. [Online]. URL: <https://lucene.apache.org/core/> (дата звернення: 15.08.2026).
2. Elastic. Elasticsearch Reference: similarity settings and BM25. [Online]. URL: <https://www.elastic.co/docs/reference/elasticsearch/index-settings/similarity> (дата звернення: 15.08.2026).
3. Apache Solr. Solr Reference Guide: indexing, querying and relevance. [Online]. URL: <https://solr.apache.org/guide/solr/latest/> (дата звернення: 15.08.2026).
4. OpenSearch Project. Getting started with semantic and hybrid search. [Online]. URL: <https://docs.opensearch.org/latest/vector-search/ai-search/hybrid-search/index/> (дата звернення: 15.08.2026).
5. scikit-learn developers. TfidfVectorizer API reference. Version 1.9.1. [Online]. URL: https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html (дата звернення: 15.08.2026).
6. NIST. Text REtrieval Conference (TREC): evaluation resources. [Online]. URL: <https://trec.nist.gov/> (дата звернення: 15.08.2026).
7. Robertson S., Zaragoza H. The Probabilistic Relevance Framework: BM25 and Beyond. Foundations and Trends in Information Retrieval, 2009. URL: <https://doi.org/10.1561/15000000019> (дата звернення: 15.08.2026).
8. Manning C. D., Raghavan P., Schütze H. Introduction to Information Retrieval. Cambridge University Press, 2008. URL: <https://nlp.stanford.edu/IR-book/> (дата звернення: 15.08.2026).

Додаткова інформація

Детальнішу інформацію щодо методів навчання, практичних занять, форм оцінювання, самостійної роботи та повного списку літератури наведено у Робочій програмі навчальної дисципліни.