



## Силабус освітнього компонента

Програма навчальної дисципліни

# Алгоритми і системи пошуку плагіату



Шифр та назва спеціальності  
124 – Системний аналіз

Інститут  
ННІ Комп'ютерних наук та інформаційних технологій

Освітня програма  
Системний аналіз і управління.

Кафедра  
Системного аналізу та інформаційно-аналітичних технологій (322)

Рівень освіти  
Бакалавр

Тип дисципліни  
Професійна, вибіркова

Семестр  
7

Мова викладання  
Українська,

## Викладачі, розробники



[ganna.sydorenko@khpi.edu.ua](mailto:ganna.sydorenko@khpi.edu.ua)

Кандидат технічних наук, доцент, доцент кафедри системного аналізу та інформаційно-аналітичних технологій НТУ «ХПІ».

Досвід роботи – 20 років. Автор понад 70 наукових та навчально-методичних праць. Провідний лектор з дисциплін: «Якість та тестування програмних систем», «Чисельні методи», «Сучасні методи математичного та комп'ютерного моделювання», «Технології розробки мобільних додатків», «Теорія масового обслуговування»

[Детальніше про викладача на сайті кафедри](#)

## Загальна інформація

### Анотація

Предметом вивчення навчальної дисципліни "Алгоритми і системи пошуку плагіату" є теоретичні знання стосовно сутності та об'єктивних закономірностей функціонування аналітичних компетентностей, необхідних під час планування і виконання дисертаційного дослідження для забезпечення адекватної комунікації наукової спільноти у сучасному світі.

### Мета та цілі дисципліни

Метою викладання навчальної дисципліни "Алгоритми і системи пошуку плагіату" є оволодіння теоретичними основами наукової комунікації, основами чинного законодавства та наказами ЗВО щодо академічної доброчесності і протидії плагіату, сучасними теоретичними основами визначення та аналізу основних наукометричних критеріїв; засвоєння методичних та теоретичних основ патентного пошуку, виявлення плагіату, несанкціонованих співпадінь, девіантної, псевдонаукової поведінки і спотворення результатів досліджень, порівняння гіпотез, вміння оцінювати результати аналізу за окремими критеріями та у взаємозв'язку з чинниками, що на них впливають; розуміння сучасних концепцій академічної доброчесності та методів протидії конфліктам інтересів і академічному плагіату.

## Формат занять

Лекції, лабораторні роботи, самостійна робота, консультації. Підсумковий контроль - іспит.

## Компетентності

Інтегральна компетентність: Здатність розв'язувати складні спеціалізовані задачі та практичні проблеми професійної діяльності видавництва та поліграфії або у процесі навчання, що передбачає застосування теорій і методів технічних, природничих, гуманітарних, соціальних наук і характеризується комплексністю та невизначеністю.

Загальні компетентності:

ЗК 1. Здатність вчитися і оволодівати сучасними знаннями.

ЗК 2. Знання та розуміння предметної області та розуміння професійної діяльності.

ЗК 3. Здатність застосовувати знання у практичних ситуаціях.

ЗК 4. Здатність приймати обґрунтовані рішення.

ЗК 8. Здатність працювати в команді.

ЗК 10. Здатність зберігати та примножувати моральні, культурні, наукові цінності і досягнення суспільства на основі розуміння історії та закономірностей розвитку предметної області, її місця у загальній системі знань про природу і суспільство та у розвитку суспільства, техніки і технологій, використовувати різні види та форми рухової активності для активного відпочинку та ведення здорового способу життя.

Спеціальні (фахові) компетентності:

СК10. Здатність розробляти експериментальні та спостережувальні дослідження і аналізувати дані, отримані в них.

## Результати навчання

РН11. Знати і вміти застосовувати на практиці системи управління базами даних і знань та інформаційні системи.

РН13. Проектувати, реалізовувати, тестувати, впроваджувати, супроводжувати, експлуатувати програмні засоби роботи з даними і знаннями в комп'ютерних системах і мережах.

## Обсяг дисципліни

Загальний обсяг дисципліни 120 год.: лекції – 32 год., лабораторні роботи – 16 год., самостійна робота – 72 год.

## Передумови вивчення дисципліни (пререквізити)

## Особливості дисципліни, методи та технології навчання

Дисципліну відрізняє суто прикладний характер лабораторних робіт та самостійних завдань, що передбачає застосування, як методів навчання, проектною і командною роботи, елементів гейміфікації, кейсових технологій тощо.

## Програма навчальної дисципліни

### Теми лекційних занять

**Тема 1 Основи академічної доброчесності. Інтелектуальна власність та об'єкти авторського права в освіті та науці.**

Інформаційно-цифрові компетентності. Цінності академічної доброчесності (АД). Правове забезпечення АД. АД в університетах світу. Забезпечення АД в НТУ ХПІ. Поняття інтелектуальної власності. Об'єкти авторського права в освіті та науці. Охорона авторського права і суміжних прав. Право інтелектуальної власності в цифровому середовищі. Ліцензії Creative Commons. Визначення можливих запозичень у текстах.

**Тема 2. Плагіат та його різновиди. Запобігання плагіату.**

Академічний плагіат та його різновиди. Системи перевірки текстів на запозичення. Практичні поради щодо запобігання плагіату. Цитування. Перефразування. Узагальнення. Загальні правила оформлення посилань в освітніх та наукових роботах. ДСТУ 8302:2015. Бібліографічне посилання. Загальні положення та правила складання.

### **Тема 3. Джерела освітньої та наукової інформації**

Основні джерела освітньої та наукової інформації. Поняття відкритого доступу до наукової і освітньої інформації. Повнотекстові, реферативні, бібліографічні, наукометричні бази даних. Джерела даних. Організація та збереження даних. Алгоритми написання та оформлення різних видів наукових текстів.

### **Тема 4. Алгоритми інтелектуальної обробки текстів**

Основні визначення та класифікація методів інтелектуальної обробки текстів. Визначення можливих запозичень у текстах Text Mining. Метод Naive Bayes. Метод шинглів.

### **Тема 5. Методи статистичного аналізу текстів.**

Основні визначення та класифікація методів статистичного аналізу текстів. Аналіз методів виявлення дублікатів. Метод шинглів. Опис методів пошуку та виявлення плагіатів. Кореляційні методи виявлення плагіату.

### **Тема 6. Кластеризація документів на основі близькості термінів**

Загальні поняття кластеризації документів. Пошук документів і кластеризація. Міри якості кластерів та кластерного розбиття. Кластерний аналіз на зважених графах. Опис алгоритму кластеризації. Методика проведення експерименту.

### **Тема 7. Перевірка зображень на плагіат**

Основні визначення та класифікація плагіату у зображеннях. Чинне законодавство та методи боротьби з плагіатом. Відповідальність за плагіат. Якими правами має фотограф? Як підвищити унікальність фото.

### **Тема 8. Перевірка на плагіат аудіо- та відео файлів.**

Засоби пошуку плагіату в аудіо- та відео файлів. Методи, що використовуються для пошуку плагіату в аудіо- та відео файлів. Інструменти перевірки на оригінальність файлів.

### **Тема 9. Плагіат у програмному коді**

Існуючі методи пошуку плагіату в програмному коді. Огляд текстових методів; структурних та семантичних методів. Метод текстового аналізу. Метод на основі аналізу стилю програмування. Метод на основі метрик програми. Метод на основі ідентифікаційних міток.

## **Теми практичних занять**

Практичні заняття в рамках дисципліни не передбачені.

## **Теми лабораторних робіт**

### **Лабораторна робота 1**

Перевірка запозиченої текстової роботи на запозичення за допомогою будь-якої антиплагіатної системи.

### **Лабораторна робота 2**

Перевірка власної текстової роботи на запозичення за допомогою будь-якої антиплагіатної системи..

### **Лабораторна робота 3**

Поняття та визначення ентропії текстового файлу.

### **Лабораторна робота 4**

Алгоритми інтелектуальної обробки текстів. Порівняння двох текстових файлів та визначення схожості.

### **Лабораторна робота 5**

Застосування методів для оцінки подібності між послідовностями маркерів у кількох текстових файлах. . Метод шинглів.

### **Лабораторна робота 6**

Розробка та програмування методів кластеризації для виявлення плагіату та особливостей автора у текстовому файлі.

### **Лабораторна робота 7**

Виявлення плагіату у зображеннях.

## Лабораторна робота 8

Використання та/або розробка та програмування методів пошуку плагіату в аудіо- та відео файлів.

## Самостійна робота

Самостійна робота передбачає поглиблене опанування наступних тем:

Тема 1. Методологічні засади академічної доброчесності.

Тема 2. Хешування з урахуванням місцевості.

Тема 3. Пошукові системи в мережі Інтернет та оптимізація пошукових запитів з метою виявлення плагіату.

Тема 4. ШІ як засіб для уникання плагіату.

Тема 5. Методи для оцінки подібності між послідовностями маркерів у кількох системах.

Тема 6. Аналіз методів та дієвих систем виявлення плагіату навчальних кодових програм.

Тема 7. Виявлення копіювання відео.

## Література та навчальні матеріали

### Основна література

1. Willson-Metzger A. C. Ethically speaking : academic librarians' perceptions of information Privacy, intellectual freedom, and the USA Patriot Act : [дис. докт. філос. у галузі освіти : 12. 2011]. Norfolk: Old Dominion University, 2021. 124 p.
2. Програмне забезпечення для перевірки наукових текстів на плагіат : інформаційний огляд / автори-укладачі: А. Р. Вергун, Л. В. Савенкова, С. О. Чуканова ; редколегія: В. С. Пашкова, О. В. Воскобойнікова-Гузєва, Я. Є. Сошинська, О. В. Башун; Українська бібліотечна асоціація. Київ : УБА, 2019. Електрон. вид. 1 електрон. опт. диск (CD-ROM). 37 с.
3. Міжнародні правила цитування та посилання в наукових роботах : методичні рекомендації / автори-укладачі: О. Боженко, Ю. Корян, М. Федорець ; редколегія: В. С. Пашкова, О. В. Воскобойнікова-Гузєва, Я. Є. Сошинська, О. М. Бруй ; Науково-технічна бібліотека ім. Г. І. Денисенка Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» ; Українська бібліотечна асоціація. – Київ : УБА, 2019. Електрон. вид. 1 електрон. опт. диск (CD-ROM). 117 с.
4. an, Wuwei; Xu, Wei (2018). Neural Network Models for Paraphrase Identification, Semantic Textual Similarity, Natural Language Inference, and Question Answering. Proceedings of the 27th International Conference on Computational Linguistics ([Santa Fe, New Mexico, USA: Association for Computational Linguistics](#)): 3890–3902.
5. Breanna Devore-McDonald, Emery D. Berger. Mossad: Defeating Software Plagiarism Detection, 2020. [Електронний ресурс] Режим доступу до ресурсу: <https://arxiv.org/pdf/2010.01700.pdf> (дата звернення: 10.05.2022).
6. Plagiarism [Електронний ресурс] Режим доступу до ресурсу : <https://www.ox.ac.uk/students/academic/guidance/skills/plagiarism> (дата звернення: 18.05.2022)
7. How to use JPlag. [Електронний ресурс] Режим доступу до ресурсу: <https://github-wiki-see.page/m/jplag/JPlag/wiki/1.-How-to-Use-JPlag> (дата звернення: 04.05.2022).

### Додаткова література

1. Тарануха В.Ю. Інтелектуальна обробка текстів: [навчальний посібник]. – Київ: електронна публікація на сайті факультету, 2014. – 80 с.
2. Навчання студентів академічній доброчесності у бібліотеці ВНЗ : методичні поради / автори-укладачі: Л. В. Савенкова, С. О. Чуканова ; редколегія: В. С. Пашкова, О. В. Воскобойнікова-Гузєва, Я. Є. Сошинська ; Українська бібліотечна асоціація. – Київ : УБА, 2016. – Електрон. вид. – 1 електрон. опт. диск (CD-ROM). – 39 с. – ISBN 978-966-97569-4-7.
2. Harris R. A. The Plagiarism Handbook: strategies for preventing, detecting, and dealing with plagiarism. Ryzczak Publishing, 2001. 208 p.
3. Gilmore B. Plagiarism : Why it happens and how to prevent it. Portsmouth, NH: Hienemann, 2008. 144p.

## Система оцінювання

### Критерії оцінювання успішності студента та розподіл балів

Бали нараховуються за наступним співвідношенням:

- лабораторні роботи: 20 % семестрової оцінки;
- самостійна робота: 20 % семестрової оцінки;
- поточні контрольні роботи: 20 % семестрової оцінки;
- іспит: 40 % семестрової оцінки.

### Шкала оцінювання

Сума балів	Національна оцінка	ECTS
90–100	Відмінно	A
82–89	Добре	B
75–81	Добре	C
64–74	Задовільно	D
60–63	Задовільно	E
35–59	Незадовільно (потрібне додаткове вивчення)	FX
1–34	Незадовільно (потрібне повторне вивчення)	F

## Норми академічної етики і політика курсу

Студент повинен дотримуватися «Кодексу етики академічних взаємовідносин та доброчесності НТУ «ХПІ»: виявляти дисциплінованість, вихованість, доброзичливість, чесність, відповідальність. Конфліктні ситуації повинні відкрито обговорюватися в навчальних групах з викладачем, а при неможливості вирішення конфлікту – доводитися до відома співробітників дирекції інституту. Нормативно-правове забезпечення впровадження принципів академічної доброчесності НТУ «ХПІ» розміщено на сайті: <http://blogs.kpi.kharkov.ua/v2/nv/akademichna-dobrochesnist/>

## Погодження

Силабус погоджено

25.08.2023

Завідувач кафедри  
Юрій ДОРОФЄЄВ

25.08.2023

Гарант ОП  
Сергій КОВАЛЕНКО