

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Національний аерокосмічний університет ім. М. Є. Жуковського
«Харківський авіаційний інститут»

Факультет систем управління літальних апаратів

Кафедра комп'ютерних наук та інформаційних технологій

Пояснювальна записка
до кваліфікаційної роботи
(тип кваліфікаційної роботи)

магістра

(освітній ступінь)

на тему «Автоматизація пошуку та візуалізації даних наукометричного аналізу
на основі моніторингу профілів дослідників»

XAI.302.366.24O.122.203103 ПЗ

Виконав: здобувач 6 курсу групи № 366
Спеціальність 122 «Комп'ютерні науки»

(код та найменування)

Освітня програма Комп'ютеризація
обробки інформації та управління

Кірілов В.Ю.

(прізвище й ініціали студента)

Керівник: проф. каф. 302, д.т.н.

Прохоров О.В.

(прізвище й ініціали)

Рецензент: Плехова Г.А.

(прізвище й ініціали)

Харків – 2024

Міністерство освіти і науки України
Національний аерокосмічний університет ім. М.Є. Жуковського
«Харківський авіаційний інститут»

Факультет систем управління літальних апаратів
 Кафедра комп'ютерних наук та інформаційних технологій
 Рівень вищої освіти другий (магістерський)
 Галузь знань 12 «Інформаційні технології»
 (код та найменування)
 Спеціальність 122 «Комп'ютерні науки»
 Освітня програма Комп'ютеризація обробки інформації та управління

ЗАТВЕРДЖУЮ
Завідувач кафедри
Олег ФЕДОРОВИЧ

(підпис) (ім'я та прізвище)
 «08» листопада 2024 р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

Кірілову Віталію Юрійовичу
 (прізвище, ім'я, по батькові)

1. Тема кваліфікаційної роботи Автоматизація пошуку та візуалізації даних наукометричного аналізу на основі моніторингу профілів дослідників

керівник кваліфікаційної роботи Прохоров О.В., д.т.н., проф. каф. 302,

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом Університету № 2050-уч. від «06» листопада 2024 року

2. Термін подання здобувачем кваліфікаційної роботи 15.12.2024

3. Вихідні дані до проекту (роботи) ресурси наукометричних баз Google Scholar, Scopus, WoS

4. Зміст пояснювальної записки (перелік завдань, які потрібно розв'язати) аналіз особливостей обробки даних ресурсів наукометричної інформації; огляд інформаційних ресурсів наукової інформації; дослідження методів аналізу ресурсів наукової та наукометричної інформації; розроблення алгоритмів для створення інформаційних масивів на базі ресурсу Google Scholar для побудови спільнот предметних областей та співавторів; здійснення програмної реалізації, проведення експериментальних досліджень та порівняння результатів

5. Перелік графічного матеріалу аналітика наукометричної інформації; мета та завдання дослідження; інформаційні ресурси наукової інформації; основні наукометричні ресурси; огляд інструментів аналізу наукометричних показників; кількісні методи наукометричного аналізу; бібліотеки для отримання даних із Google Scholar; розробка модулів пошуку та візуалізації даних наукометричного аналізу; аналіз та візуалізація даних про авторів; аналіз та візуалізація даних про групу вчених; формування спільнот для планування майбутніх досліджень

6. Консультанти розділів кваліфікаційної роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
з основної частини	Прохоров О.В. проф. каф. 302		
з спеціального розділу	Прохоров О.В. проф. каф. 302		
з економіки та управління	Малєєва Ю. А. доц. каф. 302		

Нормоконтроль _____ Олександр ПРОХОРОВ «__» _____ 20__ р.
(підпис) (імя та прізвище)

7. Дата видачі завдання « 06 » вересня 2024 р.

КАЛЕНДАРНИЙ ПЛАН

№ п/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів кваліфікаційної роботи	Примітка
1	АНАЛІЗ ОСОБЛИВОСТЕЙ ПОШУКУ ТА ОБРОБКИ ДАНИХ РЕСУРСІВ НАУКОМЕТРИЧНОЇ ІНФОРМАЦІЇ	06.09 – 20.09	
2	РЕСУРСИ ТА МЕТОДИ НАУКОМЕТРИЧНОГО АНАЛІЗУ	21.09 – 28.09	
3	РОЗРОБКА МОДУЛІВ ПОШУКУ ТА ВІЗУАЛІЗАЦІЇ ДАНИХ НАУКОМЕТРИЧНОГО АНАЛІЗУ	29.09 – 20.10	
4	МОДУЛЬ ФОРМУВАННЯ ТА АНАЛІЗУ СПІЛЬНОТ ДОСЛІДНИКІВ	21.10 – 20.11	
5	ЕКОНОМІЧНІ РОЗРАХУНКИ	20.11 – 25.11	

Здобувач _____ Віталій КІРІЛОВ
(підпис) (імя та прізвище)

Керівник кваліфікаційної роботи _____ Олександр ПРОХОРОВ
(підпис) (імя та прізвище)

УДК 004.8

РЕФЕРАТ

Автоматизація пошуку та візуалізації даних наукометричного аналізу на основі моніторингу профілів дослідників – Кірілов Віталій, кваліфікаційна робота магістра, Харків, Національний аерокосмічний університет ім. М. Є. Жуковського «ХАІ», кількість сторінок, враховуючи додатки 70, кількість таблиць 14, кількість рисунків 22, кількість джерел літератури 15.

КЛЮЧОВІ СЛОВА: співавтори, наукові дослідження, публікації, наукометричні показники, h-індекс, Google Scholar, бібліометрія, цитування.

ЦІЛЬ: дослідження та розробка моделей і алгоритмів автоматизації пошуку та проведення наукометричного аналізу бібліометричних даних для групи дослідників.

ОБ’ЄКТ ДОСЛІДЖЕННЯ: процес аналізу ресурсів наукометричної інформації.

ПРЕДМЕТ ДОСЛІДЖЕННЯ: методи, моделі та засоби аналізу та візуалізації наукометричної інформації.

ЗАДАЧІ:

1. Аналіз особливостей обробки даних ресурсів наукометричної інформації.
2. Огляд інформаційних ресурсів наукової інформації.
3. Дослідження методів аналізу ресурсів наукової та наукометричної інформації.
4. Розроблення алгоритмів для створення інформаційних масивів на базі ресурсу Google Scholar для побудови спільнот предметних областей та співавторів.
5. Здійснення програмної реалізації, проведення експериментальних досліджень та порівняння результатів.

МЕТОДИ ДОСЛІДЖЕННЯ: статистичні методи, методи аналізу наукометричних даних та показників, методи візуалізації, ресурсу Google Scholar Citations.

ОТРИМАНІ РЕЗУЛЬТАТИ: розроблено програмне забезпечення, що дозволяє вирішувати завдання вилучення комплексної бібліометричної інформації для заданого списку авторів, автоматизації пошуку та візуалізації даних про цитування, співпрацю, вплив та планування майбутніх досліджень і робочих груп для наукових проєктів та грантів.

UDC 004.8

ABSTRACT

Automation of search and visualization of scientometric analysis data based on monitoring of researchers' profiles – Vitaly Kirilov, Master Thesis, Kharkiv, National Aerospace University “KhAI”, amount of pages 70, amount of tables 14, amount of figures 22, amount of literature's sources 15.

KEYWORDS: co-authors, scientific research, publications, scientometric indicators, h-index, Google Scholar, bibliometrics, citations.

PURPOSE OF RESEARCH: research and development of models and algorithms for automating the search and scientometric analysis of bibliometric data for a group of researchers.

OBJECT OF RESEARCH: the process of analyzing scientometric information resources.

SUBJECT OF RESEARCH: methods, models, and tools for analyzing and visualizing scientometric information.

TASKS:

1. Analysis of the specifics of data processing in scientometric information resources.
2. Review of information resources for scientific information.
3. Study of methods for analyzing resources of scientific and scientometric information.
4. Development of algorithms for creating information arrays based on Google Scholar to build communities of subject areas and co-authors.
5. Implementation of software, conducting experimental research, and comparing results.

METHODS OF RESEARCH: statistical methods, methods for analyzing scientometric data and indicators, visualization methods, Google Scholar Citations resource.

RESULTS: software has been developed that enables solving tasks of extracting comprehensive bibliometric information for a given list of authors, automating the search and visualization of data on citations, collaboration, impact, and planning future research and working groups for scientific projects and grants.

ЗМІСТ

ВСТУП.....	7
1 АНАЛІЗ ОСОБЛИВОСТЕЙ ПОШУКУ ТА ОБРОБКИ ДАНИХ РЕСУРСІВ НАУКОМЕТРИЧНОЇ ІНФОРМАЦІЇ.....	9
1.1 Концепція відкритої науки та доступ до наукової інформації	9
1.2 Класифікація інформаційних ресурсів наукової інформації	10
1.3 Огляд інструментів аналізу наукометричних показників.....	13
1.4 Постановка задачі.....	15
Висновки за розділом.....	16
2 РЕСУРСИ ТА МЕТОДИ НАУКОМЕТРИЧНОГО АНАЛІЗУ.....	18
2.1 Основні наукометричні ресурси.....	18
2.2 Кількісні методи наукометричного аналізу.....	22
2.2.1 Індекс Гірша	22
2.2.2 Імпакт-фактор	23
2.2.3 Показники Google Scholar.....	24
2.3 Бібліотеки для отримання даних із Google Scholar.....	24
Висновки за розділом.....	28
3 РОЗРОБКА МОДУЛІВ ПОШУКУ ТА ВІЗУАЛІЗАЦІЇ ДАНИХ НАУКОМЕТРИЧНОГО АНАЛІЗУ.....	29
3.1 Отримання даних з Google Scholar	29
3.2 Аналіз та візуалізація даних про автора	32
3.3 Аналіз та візуалізація даних про групу вчених	37
Висновки за розділом.....	43
4 МОДУЛЬ ФОРМУВАННЯ ТА АНАЛІЗУ СПІЛЬНОТ ДОСЛІДНИКІВ.....	44
4.1 Формування спільнот для планування майбутніх досліджень	44
Висновки за розділом.....	49
5 ЕКОНОМІЧНИЙ РОЗДІЛ.....	50
5.1 Мета розділу.....	50
5.2 Розрахунок собівартості та визначення тривалості створення програмного продукту (перший варіант)	50
5.3 Розрахунок собівартості та визначення тривалості створення програмного продукту (другий варіант).....	52
5.4 Висновки за розділом	56
ВИСНОВКИ.....	60
ПЕРЕЛІК ПОСИЛАНЬ	61
ДОДАТОК А. Графічні матеріали	63

ВСТУП

Актуальність теми дослідження.

Сучасне наукове співтовариство стикається з новими викликами, які спрямовані на залучення більшої кількості індикаторів для оцінки, систематизації та узагальнення наукової діяльності. Одним із широко визнаних у світі підходів, що використовуються для формування загальної картини певної галузі науки чи науки країни в цілому, оцінки публікаційної активності, успіху вченого, наукового колективу, проекту або окремої установи, є використання наукометричних показників. Для організації роботи із наукометричним інструментарієм розроблено ряд ресурсів, що представляють таку інформацію: переліки публікацій, видань, галузева приналежність, відповідні показники, що обраховуються автоматично. Таким чином, актуальність теми дослідження обумовлена необхідністю розробки моделі, що забезпечує ефективне використання даних про співробітників для підтримки ухвалення ефективних та своєчасних управлінських рішень. Роботу присвячено актуальному завданню, що дозволяє вирішувати задачі вилучення комплексної бібліометричної інформації для заданого списку авторів, автоматизації пошуку та візуалізації даних про цитування, співпрацю, вплив та планування майбутніх досліджень і робочих груп для наукових проектів та грантів.

Мета дослідження – дослідження та розробка моделей і алгоритмів автоматизації пошуку та проведення наукометричного аналізу бібліометричних даних для групи дослідників.

Для досягнення мети необхідно виконати наступні **завдання**:

1. Аналіз особливостей пошуку та обробки даних ресурсів наукометричної інформації.
2. Огляд інформаційних ресурсів наукової інформації.
3. Дослідження методів аналізу ресурсів наукової та наукометричної інформації.
4. Розроблення алгоритмів для створення інформаційних масивів на базі ресурсу Google Scholar для побудови спільнот предметних областей та співавторів.
5. Здійснення програмної реалізації, проведення експериментальних досліджень та порівняння результатів.

Об'єкт дослідження – процес аналізу ресурсів наукометричної інформації.

Предмет дослідження – методи, моделі та засоби аналізу та візуалізації наукометричної інформації.

Практичне значення отриманих результатів. Розроблено програмне забезпечення, що дозволяє вирішувати завдання вилучення комплексної бібліометричної інформації для заданого списку авторів, автоматизації пошуку та візуалізації даних про цитування, співпрацю, вплив та планування майбутніх досліджень і робочих груп для наукових проєктів та грантів.

Структура та обсяг кваліфікаційної роботи магістра. Кваліфікаційна робота магістра складається зі вступу, 5 розділів, висновків, додатків та переліку посилань. Повний обсяг кваліфікаційної роботи магістра становить 70 сторінок, в тому числі: 22 рисунки; 14 таблиць; 1 додаток на 8 сторінках; перелік посилань з 15 найменувань на 2 сторінках.