

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ЧОРНОМОРСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ПЕТРА МОГИЛИ

**Попель Олександр Олександрович**

УДК 004.93

**СТВОРЕННЯ СИСТЕМИ АВТОМАТИЧНОГО ФОРМУВАННЯ БАНКУ  
МЕДІАДАНИХ**

Галузь знань 12 «Інформаційні технології» за спеціальністю

122 «Комп'ютерні науки»

122 – ДР.А – 403.21830303

Автореферат

дипломної роботи на здобуття освітньої кваліфікації

«бакалавр комп'ютерних наук»

Миколаїв – 2020

Дипломна робота є рукопис.

Робота виконана в Чорноморському національному університеті імені Петра Могили Міністерства освіти і науки України на кафедрі Інтелектуальних інформаційних систем

Науковий керівник:

Канд. пед. наук, доц. каф. ІПЗ,  
Кірей Катерина Олександрівна

Рецензент:

Заст. декана, доц. каф. ІПЗ, к.т.н.,  
Давиденко Євген Олександрович

Захист відбудеться « 23 » червня 2020 р. о 9<sup>30</sup> год. на засіданні екзаменаційної комісії (ауд. 2-403) у Чорноморському національному університеті імені Петра Могили за адресою: 54003, м. Миколаїв, вул. 68-ми Десантників, 10.

З дипломною роботою можна ознайомитися в бібліотеці Чорноморського національного університету імені Петра Могили за адресою: 54003, м. Миколаїв, вул. 68-ми Десантників, 10.

Автореферат представлений « \_\_\_\_ » червня 2020 р.

Секретар  
екзаменаційної комісії,  
викладач кафедри ІС

Скакодуб Олександр Сергійович



## ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

### **Актуальність теми.**

Невпинний розвиток технологій на протязі останніх двох століть призвів до того, що кількість інформації, доступної кожній людині зростає експоненційно. Від численних оцифрованих книг, газет та журналів, кінофільмів та музичних творів, до створених вже в цифровому форматі статей або відео – до всього цього можна отримати доступ завдяки Інтернету.

Така загальнодоступність інформації дала змогу ще більше пришвидшити розвиток освіти та технологій, а також дати змогу невідомим раніше авторам та їх публікаціям чи творам отримати заслужену аудиторію, не обмежену бажаннями та цілями наукових журналів, літературних видавництв чи музичних студій. І хоча, беззаперечно, це є черговим кроком людства у правильному напрямку, як і будь-який прогрес, це викликає і пов'язані проблеми. Коли інформації забагато, знайти потрібну часто нетривіальна, а інколи і майже неможлива задача. На щастя, задача пошуку тексту була вирішена ще десятиліттями тому, завдяки бурхливому розвитку пошукових систем. Індексація гіпертексту займає значно менше як обчислювальних ресурсів, так і місця у пам'яті ніж індексація інших типів інформації. Тому донедавна пошук додаткової інформації про, наприклад, медіа-файли у Інтернеті, залежав від наявності пов'язаної з ними текстової інформації. Але така інформація не завжди була доступною і це перетворювало пошук метаданих до цих медіа-файлів у довгий і складний процес.

В той час як кінофільми завжди включають інформацію про себе, не кожен музичний твір містить навіть слова, які змогли б допомогти ідентифікувати його шляхом пошуку тексту в Інтернеті. Але навіть наявність у творі текстової інформації, що може допомогти його ідентифікувати, не є достатньою у випадку наявності достатньо великої медіатеки. В сучасних масштабах кількості інформації, ручний пошук часто не є здійсненим. Тому, як це часто буває, технологічний прогрес почав вирішення проблеми, яку сам і створив. Розвиток систем штучного інтелекту, комп'ютерного бачення та слухання дозволив машинам сприймати медіа-інформацію схожим чином до того, як її сприймають люди. І, завдяки набагато

кращим (та швидшим) здібностям до пошуку інформації, ніж у людей, комп'ютери нарешті можуть вирішити проблему ідентифікації майже будь-якої інформації та пошуку її метаданих.

Незважаючи на швидкий розвиток стрімінгових сервісів, у багатьох людей залишається потреба у локальному зберіганні власної аудіотеки. Інтернет-покриття сьогодні ще є не повсюди, а навіть там де є, мобільне – часто лімітоване. До того ж, не можна бути впевненим у тому, що потрібна інформація залишиться незмінною чи взагалі залишиться у стрімінгового сервісу. Численні вилучення альбомів зі Spotify чи блокування соціальної мережі «Вконтакте» призвели до того, що багато людей втратили доступ до інформації, яку довірили зовнішньому сервісу і є чудовим прикладом того, що зберігання власних медіа-бібліотек є актуальним і на сьогодні.

Аудіо-файли можуть втратити пов'язані метадані через багато чинників – від пошкодження файлу до неякісного програмного забезпечення чи недобросовісних дистриб'юторів, що не додають таку інформацію до файлів, які розповсюджують. Саме тому пошук такої інформації є необхідним та актуальним, а завдяки розвитку інформаційних технологій – і можливим.

**Метою дипломної роботи** є створення системи автоматичного формування банку медіаданих задля вирішення проблеми ідентифікації невідомого аудіо за допомогою відкритих та вільно доступних програмних компонентів, веб-сервісів та баз медіаданих.

### **Практичне значення отриманих результатів.**

Практичне значення дипломної роботи полягає у дослідженні вільних та відкритих засобів аудіо-ідентифікації та побудови локальних банків медіаданих, а також у створенні готової до використання широким колом зацікавлених користувачів системи, яка якісно виконує поставлену задачу.

**Структура дипломної роботи.** Пояснювальна записка до дипломної роботи складається із вступу, 5-и розділів, висновків, додатку. Загальний обсяг роботи складає 102 сторінки, 34 рисунків, 1 таблиці та 30 посилань на літературні джерела.

## ОСНОВНИЙ ЗМІСТ РОБОТИ

**У вступі** описується стан предметної сфери, тобто сфери ідентифікації невідомого аудіо, та існуючі в ній проблеми, такі як велика кількість невідомих медіа файлів без розпізнавальних метаданих. Коротко розглядаються причини виникнення вищезгаданих проблем та існуючі технології, спрямовані на їх вирішення. Формується мета дипломної роботи, визначаються її предмет та об'єкт. Пояснюються складнощі, з якими зустрічаються люди, на яких впливають описані проблеми та чітко визначається актуальність роботи на основі цього. Наостанок, коротко описується порядок виконання кожного з розділів дипломної роботи, та мета кожного з них.

**У першому розділі** детально досліджується питання актуальності автоматичного формування локальних сховищ медіаданих, перевіряються наявні готові рішення та алгоритми за допомогою яких вони працюють, визначаються недоліки цих рішень та формується перелік вимог до нової розроблюваної системи.

Для чіткої постановки задачі пропонуються загальні способи вирішення деяких з проблем, що дозволяють базувати подальші рішення, пов'язані з загальним дизайном системи, на чітко визначеній наперед основі.

**У другому розділі** розглядаються алгоритми, що є основою головної частини створюваної системи – акустичного фінгерпринтера. Визначаються проблеми реального світу з якими він мусить бути здатним впоратись для виконання поставленої задачі і аналізуються методи, за допомогою яких він ці проблеми вирішує. У першому підрозділі розглядається стан вхідних даних системи, побудова якої є задачею дипломної роботи, а саме – стиснутих музичних файлів. Також передбачається ускладнення, які певні типи даних можуть викликати, шляхом дослідження ефектів різних видів стиснення на оригінальні аудіо файли. У другому підрозділі розглядаються принципи роботи алгоритму фінгерпринтингу Chromaprint та описуються математичні моделі, які він використовує для створення надійних і унікальних відбитків музичних файлів, а також методи кодування результату у компактну форму репрезентації для ресурсо-ефективного зберігання та пошуку

збігів у відбитках. Також перевіряється робота алгоритму на практиці, використовуючи тестові вхідні дані, закодовані у lossy та lossless формати а також дані що є окремими сутностями незважаючи на свою схожість (інструментальні версії композицій та версії з вокалом). Робиться висновок, що в цілому, алгоритм чудово проявляє себе у різних складних умовах і його можна вважати придатним для використання у вирішенні основної задачі дипломної роботи.

**У третьому розділі** розробляються та будуються функціональна та процесна моделі системи, детально описується кожна з виділених підсистем проекту дипломної роботи і чітко визначаються їх задачі та загальна структура. Також проводиться загальний огляд програмно-технічного забезпечення, придатного для використання у розробці створюваної системи, визначаються загальні рекомендації для вибору компонентів, що використовуватимуться у системі та проводиться первинний відбір можливих компонентів за наявності очевидних функціональних переваг чи недоліків у останніх. До того ж, розглядаються мови програмування, що використовуються для розробки програмних систем даного типу і визначаються такі, що відповідають вимогам, поставленим у першому розділі.

**У четвертому розділі** виконується практична частина поставленої задачі, а саме – створюється програмна реалізація системи автоматичного формування банку медіаданих. Для цього спочатку обираються програмні інструменти та базові програмні засоби, такі як медіа-бібліотеки, графічний інструментарій та мова програмування. Потім, перед розробкою програмної реалізації обирається технологію розробки ПЗ, що найкраще відповідає особливостям проекту. Далі розробляється та описується програмна реалізація системи та демонструється її робота. Наостанок розробляється інструкція користувача з ціллю зробити програму доступною для будь-кого.

**У розділі з охорони праці** досліджуються рекомендації та правила при роботі з ПК, задля збереження здоров'я та безпеки працівників, що проводять більшість робочого часу за персональними комп'ютерами.

## ЗАГАЛЬНІ ВИСНОВКИ

Метою даної дипломної роботи було створення системи автоматичного формування банку медіаданих задля вирішення проблеми ідентифікації невідомого аудіо за допомогою відкритих та вільно доступних програмних компонентів, веб-сервісів та баз медіаданих. В результаті виконання дипломної роботи, її мету було досягнуто повністю. Створена програмна система дозволяє будь кому створювати власне локальне сховище медіаданих за допомогою сучасних алгоритмів, заснованих на технологіях комп'ютерного бачення, що використовуються для ідентифікації невідомого аудіо, дозволяючи додати його до персонального банку медіаданих.

В ході виконання дипломної роботи було досліджено основу проекту – сферу аудіо-ідентифікації, розглянуто існуючі рішення поставленої задачі та визначено їх недоліки і теоретичні способи їх вирішення. Сформовано список вимог до створюваної системи. Після цього було детально досліджено математичні моделі та методи як вхідних даних створюваної системи, так і алгоритмів, що мають стати її частиною. Далі було розроблено моделі системи, придатної для вирішення сформульованої проблеми, з урахуванням раніше встановлених вимог. Останньою частиною даного етапу стало практичне проектування створюваної системи та загальний огляд програмно-технічного забезпечення, придатного для її реалізацію. Заключним етапом основної частини дипломної роботи став вибір базових програмних засобів та бібліотек для використання у програмній реалізації системи, вибір технології розробки ПЗ, яка найбільше підходить для розробки програмної реалізації даної системи, суто програмна реалізація системи і створення інструкції користувача для збільшення доступності системи якомога більшої кількості користувачів.

Результатом роботи стала повністю функціональна система автоматичного формування банку медіаданих, що дозволяє будь кому автоматично організовувати власне локальне сховище метаданих медіафайлів.



## АНОТАЦІЯ

**Попель Олександр Олександрович.** Створення автоматизованої системи формування банку медіаданих. – На правах рукопису.

Дипломна робота на здобуття освітньої кваліфікації «бакалавр комп'ютерних наук та інформаційних технологій» в галузі знань 12 «Інформаційні технології» за спеціальністю 122 «Комп'ютерні науки та інформаційні технології»

Чорноморський національний університет імені Петра Могили, Миколаїв

Дипломна робота містить теоретичні дані про сферу аудіо-ідентифікації і використання її засобів для формування локальних банків медіаданих та опис практичної реалізації викладених ідей.

Об'єктом дослідження є медіа файли, що не містять жодної інформації про їх походження у вигляді текстових метаданих.

Предметом дослідження є автоматична комп'ютеризована ідентифікація таких файлів, що дозволяє наповнити їх інформацією, достатньою для визначення походження та каталогізації ідентифікованих файлів.

Робота складається з чотирьох розділів та спеціальної частини з охорони праці.

У першому розділі досліджено та проаналізовано сферу автоматичної аудіо-ідентифікації, визначено основні методи та принципи роботи широко використовуваних алгоритмів і технік даної сфери. Розглянуто існуючі рішення, що дозволяють вирішити поставлену проблему, сформовано списки їх недоліків та визначено набір вимог для нової системи.

У другому розділі детально описано процес створення акустичного відбитку обраним алгоритмом, який потім порівнюється з існуючими відбитками у базі даних для ідентифікації невідомого аудіо.

Третій розділ відведений на створення моделей програмної системи, здатної вирішувати поставлену проблему і опису програмних компонентів, що реалізують частини необхідного функціоналу.

У четвертому розділі описано, порівняно та обрано програмні засоби, що використовуватимуться у створюваній системі, визначено найбільш ефективну

технологію розробки ПЗ для даного проекту, описано програмну реалізацію і розроблено інструкцію користувача створеної програмної системи.

У спеціальній частині з охорони праці визначено необхідні умови для безпечної роботи з комп'ютерними системами в приміщеннях офісного типу.

Робота складається з 105 сторінок, з них 68 складає основна частина, що містить 34 рисунки, 2 формули, 1 таблицю і 30 джерел.

*Ключові слова: медіа, аудіо, автоматизація, розпізнавання аудіо, аудіо-фінгерпринтинг, комп'ютерне бачення, медіа теги.*

## ABSTRACT

Popel O. O. Creation of an automated system for organizing a media data bank.

This thesis has theoretical data about audio-identification domain and its usage in local media data banks formation, and a description of a practical implementation of proposed ideas.

The object of the research is media files that doesn't contain any information about their origin in for of textual metadata.

The subject of the research is an automatic computer-based identification of such files that allows filling them with information that is enough for determining their origin and cataloging.

The research consists of four chapters and a special chapter on labor safety.

In the first chapter, the domain of automatic audio-identification is researched and analyzed, the main methods and working principles of widely used algorithms and techniques are defined. Existing solutions are reviewed, the lists of their drawbacks are formed and a set of requirements for a new system is created.

The second chapter consists of the description of the process of acoustic fingerprint creation, which then gets compared to existing fingerprints in a database for identifying an unknown audio.

The third chapter is allotted to creation of models of a software system that is able to solve the task and the description of software components that implement parts of the required functions.

The fourth chapter describes, compares and selects software tools that will be used in the system, determines the most effective software development process for this project, describes the software implementation of the system and develops a user manual for the created software system.

The special chapter on labor safety determines the safe working conditions for working with computer systems in office-type buildings.

The thesis consists of 105 pages, of which 58 are the primary research, that has 29 pictures, 2 formulas, 1 table and 17 references.

*Keywords: media, audio, automation, audio identification, audio-fingerprinting, computer vision, media tags.*