

DOI <https://doi.org/10.51647/kelm.2022.6.31>

MOŻLIWOŚCI WYKORZYSTANIA W ANALIZIE KRYMINALNEJ TECHNOLOGII TRANSKRYBACJI (NA PRZYKŁADZIE NARODOWEJ POLICJI UKRAINY)

Liliia Korostelova

*Inspektor Wydziału Wsparcia Programów Informatyzacji i Pracy z Otwartymi Danymi Biura Wsparcia
Informacyjno-Analitycznego Głównego Urzędu Narodowej Policji w Obwodzie Ługańskim
(Siewierodoneck, Ukraina)*

ORCID ID: 0000-0002-4782-1354

rozaenko@ukr.net

Dmytro Hudenko

*kierownik Departamentu Analizy Karnej
Narodowej Policji Ukrainy (Kijów, Ukraina)*

ORCID ID: 0000-0002-9789-2361

sciencehdm@gmail.com

Adnotacja. W artykule omówiono innowacyjne narzędzie do automatycznego rozpoznawania materiałów audio do formatu tekstowego. Zidentyfikowano problematyczne kwestie, które pojawiają się podczas wdrażania wspomnianego narzędzia. Opisano główne etapy procesu transkrypcji do formatu tekstowego.

Zaproponowano i rozważono koncepcje wykorzystania tego narzędzia dla jednostek analizy kryminalnej Policji Narodowej Ukrainy.

Przeprowadzono praktyczną analizę wykorzystania oprogramowania testowego, które rozpoznaje pliki audio w formatach takie jak format MP3, który jest najbardziej rozpowszechnionym i najpopularniejszym formatem cyfrowego kodowania informacji dźwiękowych, obsługuje popularne systemy operacyjne. Zbadano możliwość rozpoznawania formatu WAV (format sygnału audio) i formatu MPEG, który koduje wideo i powiązany z nim dźwięk do późniejszego przechowywania z przepływnością 1,5 Mb/s (ISO/IEC 11172).

Ustalono opinię autorską dotyczącą dalszego wykorzystania oprogramowania testowego „Converting audio files into text” do analiz analitycznych w jednostkach analizy kryminalnej.

Słowa kluczowe: analiza kryminalna, transkrypcja plików audio, Python, analiza danych.

POSSIBILITIES OF USING TRANSCRIPTION TECHNOLOGY IN CRIME ANALYSIS (IN NATIONAL POLICE OF UKRAINE AS AN EXAMPLE)

Liliia Korostelova

*Inspector of the Department of Support for Informatization Programs and Open Data Maintenance of
Department of Information and Analytical Support of the Main Department of the National Police in
Luhansk region ((Severodonetsk, Ukraine)*

ORCID ID: 0000-0002-4782-1354

rozaenko@ukr.net

Dmytro Khudenko

*Head of Crime Analysis Department of the National Police of Ukraine
(Kyiv, Ukraine)*

ORCID ID: 0000-0002-9789-2361

sciencehdm@gmail.com

Abstract. The article considers an innovative tool for the automatic recognition of audio materials into text format. Problematic issues arising during the implementation of the specified tool are identified. The main stages of the process of transcription into text format are described.

Concepts of using the specified tool for crime analysis units of the National Police of Ukraine were proposed and considered.

Practical analysis of the use of the test software product, which recognizes audio files in formats such as the MP3 format, which is the most common and popular format of digital encoding of sound information supported by popular operating systems, was conducted. The possibility of recognizing the WAV format (audio signal format) and the MPEG format which encodes video and the associated sound for further storage with a bit rate of 1.5 Mb/s (ISO/IEC 11172) has been studied.

The author's opinion on the further use of the test software product "Converting audio files into text" for analytical intelligence in crime analysis units is defined.

Key words: crime analysis, transcription of audio files, Python, data analysis.

МОЖЛИВОСТІ ВИКОРИСТАННЯ У КРИМІНАЛЬНОМУ АНАЛІЗІ ТЕХНОЛОГІЇ ТРАНСКРІБАЦІЇ (НА ПРИКЛАДІ НАЦІОНАЛЬНОЇ ПОЛІЦІЇ УКРАЇНИ)

Лілія Коростельова

*інспектор відділу супроводження програм інформатизації та роботи із відкритими даними
Управління інформаційно-аналітичної підтримки Головного управління Національної поліції
в Луганській області (Севєродонецьк, Україна)*

ORCID ID: 0000-0002-4782-1354

rozaenko@ukr.net

Дмитро Худенко

*начальник Департаменту кримінального аналізу
Національної поліції України (Київ, Україна)*

ORCID ID: 0000-0002-9789-2361

sciencehdm@gmail.com

Анотація. У статті розглянуто інноваційний інструмент автоматичного розпізнання аудіо матеріалів у текстовий формат. Визначені проблемні питання, що виникають під час реалізації зазначеного інструмента. Описано основні етапи процесу транскрибації у текстовий формат.

Запропоновано та розглянуто концепції використання зазначеного інструменту для підрозділів кримінального аналізу Національної поліції України.

Проведено практичний аналіз використання тестового програмного продукту, який розпізнає аудіо файли форматів таких як формат MP3, що є найбільш поширеним і популярним форматом цифрового кодування звукової інформації, що підтримує популярні операційні системи. Вивчено можливість розпізнання формату WAV (звуковий формат сигналу) та формату MPEG, що кодує відео та пов'язаний з ним звук для подальшого зберігання з бітрейтом 1.5 Мб/с (ISO/IEC 11172).

Визначена авторська думка, щодо подальшого використання тестового програмного продукту «Converting audio files into text» для аналітичних розвідок у підрозділах кримінального аналізу.

Ключові слова: кримінальний аналіз, транскрибація аудіофайлів, Python, аналіз даних.

Introduction. In the 21st century, the concept of information acquires a new meaning. We live in the information age. Today, information is the most valuable resource that humanity possesses. Under the condition of its skillful use, in the hands of the performer, unlimited influence is concentrated on a specific individual or on a large social mass (Kosyk, 2018).

Indeed, the information explosion and the modernization of society have very radically changed the life of society. In modern Ukrainian conditions, there is an increasing need for the collection, processing, analysis and transmission of information using automated resources, in particular for law enforcement agencies. At the same time, the process of creating, processing, transmitting and storing information requires the criminal police in certain cases to carry out its preliminary analysis, diagnosis or recognition as well as to create a working copy of physical evidence or "work product" in the form of a file with a speech recording for the mentioned actions (Протокол Берклі з ведення розслідувань з використанням відкритих цифрових даних, 2020).

The issue of the possibility of using technology in crime analysis, which ensured the transformation of speech into text or transcription, became especially relevant in wartime conditions because it is due to an extremely systemic practical need. Open source or OSINT includes audio and video recordings of human conversations. In the conditions of the legal regime of martial law, new content of specific content appeared - recordings of radio intercepts, interviews of prisoners of war, public statements, etc., analysis of which helps to better understand the statements and to conduct search of necessary information.

The problem is diverse and lies in the fact that, In the light of the increase in the number of samples for the analytical research, the cognitive burden on the method of manual processing and one specific program is not universal. Moreover, the technology itself is not perfect. Although researchers are trying to solve the existing shortcomings as science improves. The essence of these shortcomings is different and is determined by objective factors, for example, a weak signal against the background of surrounding noise, so fast speech that the words seem to be swallowed, unclear or expressive speech, or the person's use of homonyms, or a strong accent. However, a number of shortcomings can be corrected by the user himself or by using a set of appropriate programs.

The main part of the scientific article. The purpose of our article is to study and characterize the possibility of using transcription technology in crime analysis (In National Police of Ukraine as an example)

Both foreign and domestic scientists, including T. Vyntsiuk (Винцюк, 2000), F. Dzhielinek (Джелинек, 1976), С. Во (С. Во, 2009), Itakura F. (Itakura, 1975), Hinton G. (Hinton, 2012) and others, paid attention to the study of transcription technology in various fields of science.

The legal basis of technology at the national level is in the process of formation and is currently formed by separate provisions contained in the Constitution of Ukraine dated 28.06.1996, the Convention on the Protection of Human Rights and Fundamental Freedoms dated 04.11.1950, the Code of Criminal Procedure of Ukraine

dated 13.04.2012, Laws of Ukraine dated 18.02.1992 "On operative investigative activity", dated 06.01.2010 "On personal data protection", dated 10.02.1992 "On information", Decree of the Cabinet of Ministers of Ukraine dated 02.12.2020 No. 1556-r "On approval of the Concept of the development of artificial intelligence in Ukraine", etc.

(Of course, there are a number of ready-made software solutions, for example, the American company Nuance has already developed a number of ready-made software solutions for example the platform for application in such spheres as medicine, law, finance, journalism, construction, security and automotive sector.

According to the order of the National Police of Ukraine dated December 29, 2019 No. 1354 "On approval of the Regulation on the Department of Crime Analysis of the National Police of Ukraine", the Department of Crime Analysis develops, implements and applies new methods and areas of crime analysis aimed at increasing the effectiveness of combating crime.

One of the innovative methods for the needs of crime analysis is natural language processing technology. This technology is one of the most promising areas of development and data analysis.

According to current definitions, natural language processing can be defined as software and computer applications designed to apply computational techniques to the analysis and synthesis of natural language and speech. The main purpose of such applications is to implement language processing, just like a person does, for the analysis in a broad sense of texts formed with the help of computer technology (Ставицький, Нескородева, 2021).

In general, speech recognition is understood as the process of converting a speech signal into a text stream (Розпізнання мовлення, 2022). In our case, we are not talking about the answer to the question of which language the segment of the speech signal belongs to.

Research material and methods. Research in the recognition of audio and video files occupies a special place among artificial intelligence because the results in the field of processing and analysis of the symbol stream are applied in many expert systems.

However, relying only on practical aspects, it is not always possible to comprehensively and optimally assess the entire problem as a whole. Thus, the problems of transmission, storage, processing and analysis of the stream pose tasks, the purpose of which is to build a compressed representation of a whole class of streams, taking into account the fact that the individual characteristics of a person require compactness, informativeness and novelty in everything. In this connection, a number of tasks arise in the field of linguistic analysis and closure of information flows (cryptography, shorthand, etc.) (Хорошко, 2011).

The introduction of modern tools for the recognition of audio files into text in law enforcement activities, particularly in crime analysis units, has a promising direction of research

Transcription of audio into text contributes to an innovative approach, simplifies the indexing of the provided information, and the user perceives the content in a more convenient way for himself. Several tools can be used for fast and high-quality speech recognition of audio or video into text. It is also easier for search robots to perceive information in the form of printed text than in video format. To improve indexing, it is best to convert audio files to text. In connection with the beginning of a full-scale military invasion by Russia into our country, there was a need to use automatic tools for recognizing audio files in text.

Results and their discussion. The Department of Crime Analysis of the National Police of Ukraine created a test software product based on the Python programming language (version 3.9), namely "Converting audio files into text". This program recognizes audio files in formats such as MP3 format, which is the most common and popular format for digital encoding of sound information, which supports popular operating systems. And also has the ability to recognize the WAV format (audio signal format) and the MPEG format, which encodes video and the associated sound for further storage with a bit rate of 1.5 Mb/s (ISO/IEC 11172).

Thanks to the flexibility of the "Converting audio files into text" program, it can be easily supplemented with various key tags as well as further develop the technology of multi-level voice analysis to determine the gender of the person and emotional state.

```
import os
import shutil
from pathlib import Path
import subprocess
import math
from typing import List
import speech_recognition as sr
from pydub import AudioSegment
from audio_models import create, read, insert_tags_for_text
from translate_text import translate
from utilities.utils import my_str
os.environ["PATH"] += os.pathsep + R"C:\Users\User\Desktop\ffmpeg-2022-07-21-git-f7d510b33f-full_build\
bin"
def convert_audio_to_wav(audio_file: Path, out_dir: Path) -> Path:
    """Convert audio to wav."""
    wav_file = out_dir / audio_file.with_suffix('.wav').name
    file_format = audio_file.suffix[1:].lower()
    if file_format == 'wav': # just copying
```

```

    shutil.copy(audio_file, wav_file)
else:
    if file_format == 'mp3':
        print('*mp3 !' + 20*'.')
        audio = AudioSegment.from_mp3(audio_file)
    else:
        audio = AudioSegment.from_file(audio_file, format=file_format)
        audio.export(wav_file, format="wav")
print(F"{wav_file=}, \n {str(wav_file.resolve())}")
return wav_file
class SplitWavTimed():
    """Split wav-file into timed chunks.
    Inspired by:
    https://stackoverflow.com/questions/37999150/how-to-split-a-wav-file-into-multiple-wav-files
    """
    def __init__(self, audio_file_path: Path, folder_to: Path, chunk_duration_in_mins=4):
        self.audio_file_path = audio_file_path
        self.folder_to = folder_to
        self.chunk_duration_in_mins = chunk_duration_in_mins
        self.audio = AudioSegment.from_wav(self.audio_file_path)

```

In this case, we used the Python speech recognition library for this transcription process. It supports several speech recognition engines such as CMU Sphinx, Google Cloud Speech API, Microsoft Bing Voice Recognition and IBM Speech to Text. It should be noted that some of these mechanisms require the use of API tokens. In this project, we use Google speech recognition and API key by default.

The text recognition library has a Recognizer class that contains a set of built-in functions for settings and speech recognition. First of all we import the library and configure the recognizer, and then we import the audio file for decoding.

From here, we can directly perform the transcription using the `recognize_google` function, which takes an input audio file and outputs the transcription as a conclusion. However, depending on the quality of the audio file, there are several other features that can be used to improve your audio file and therefore improve your transcription.

A speech recognition mechanism can have multiple transcription results available to different recognition methods. The Recognizer class automatically selects the one with the highest reliability score. You can also check other transcription results by adding the function `show_all=True` to the `.recognize_google` variable.

In addition to the original recognition of audio files into text format, the specified program has the function of translating the original text into Ukrainian using the free version of Google Translate, in the specified case, the `language="id-ID Ukraine"` variable `recognize_google` function was added as well as support for searching for key tags was added (for various types of criminally illegal acts).

Audio file recognition can simultaneously decode up to 10 audio files of different sizes and formats. Within the framework of the research, 7 audio files with MP3, WAV, and MPEG formats were analyzed, with the largest size of 42MB each.

Currently, within the framework of the research, audio files of intercepted conversations of the Russian military were used, which in the conditions of war is a relevant area of work, but in the future, with the possibility of adding various key tags, the software product has prospects. As a result of the conducted research, the pattern of high-quality 89-99% recognition of audio files into text with automated downloading into Microsoft Excel format and automatic search for key tags in the text was revealed.

Conclusion. "It can be concluded that") To summarize "Converting audio files into text" or transcoding technology has the necessary range of capabilities for law enforcement agencies, in particular for conducting analytical research by the crime analysis units of the National Police of Ukraine. The specified software product in the test version has a simple and flexible interface and is quite easy to use with various audio files. The effectiveness of using such a product should be considered under the condition of an integrated approach, that is, with the use of specialized software, for example, Nuance.

References:

1. Iryna Kosyk. (2018). Informatsiinyi vybukh i hlobalizatsiia svitovoi polityky. [Information explosion and globalization of world politics.]. URL: https://enigma.ua/articles/informatsiyniy_vibukh_i_globalizatsiya_svitovoi_politiki [in Ukrainian].
2. Protokol Berkli z vedennia rozsliduvan z vykorystanniam vidkrytykh tsyfrovyykh danykh. [The Berkeley Protocol for Conducting Investigations Using Open Digital Data]. (2020) <https://www.law.berkeley.edu/wp-content/uploads/2022/03/Berkeley-Protocol-Ukrainian.pdf> [in Ukrainian].
3. Vynsiuk T.N. (2000). Obraznyy kompiuter Sovremennye problemy v kompiuternykh naukakh. [Visual computer Contemporary problems in computer science]. *Sovremennye problemy v kompiuternykh naukakh*. P. 5-14. [in Ukrainian].
4. Dzhelynek F. (1976). Raspoznavanye nepreryvnoi rechy s pomoshchiu statystycheskykh metodov. [Recognition of continuous speech using statistical methods Statistical theory of speech recognition]. *Statystycheskaia teoriya raspoznavaniya rechy*. Vol 4, P. 131-160. [in Ukrainian].

5. Bo C. (2009). Design and realization of an intelligent access control system based on voice recognition. *Communication, control and Management. IEEE International Conference*. P. 229–232. [in English].
6. Itakura F. (1975). Minimum prediction residual principle applied to speech recognition. *IEEE*. P. 67-72. [in English].
7. Hinton G. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *Signal Processing Magazine. IEEE*. P. 82-97. [in English].
8. Stavytskyi I. O., Neskorodieva A. R. (2021). Rozrobka servisu rozpoznavannia movlennia Bambusa. [Development of a speech recognition service]. *Visnyk studentskoho naukovooho tovarystva DonNU imeni Vasylia Stusa*, Vyp. 13, P. 294-297. [in Ukrainian].
9. Rozpoznavannia movlennia. [Speech recognition] (2022) https://uk.wikipedia.org/wiki/%D0%A0%D0%BE%D0%B7%D0%BF%D1%96%D0%B7%D0%BD%D0%B0%D0%B2%D0%B0%D0%BD%D0%BD%D1%8F_%D0%BC%D0%BE%D0%B2%D0%BB%D0%B5%D0%BD%D0%BD%D1%8F [in Ukrainian].
10. Khoroshko V.O. (2011). Analiz symvolichnykh informatsiynykh potokiv dlia vyznachennia indyvidualnykh osoblyvostei liudyny. [Analysis of symbolic information flows to determine the individual characteristics of a person]. *Suchasna spetsialna tekhnika*. Vol. 1 (24). P. 5-11[in Ukrainian].

Список використаних джерел:

1. Ирина Косік. Інформаційний вибух і глобалізація світової політики. URL: https://enigma.ua/articles/informatsiyniy_vibukh_i_globalizatsiya_svitovoi_politiki (дата звернення 28.10.2022)
2. Протокол Берклі з ведення розслідувань з використанням відкритих цифрових даних. URL: <https://www.law.berkeley.edu/wp-content/uploads/2022/03/Berkeley-Protocol-Ukrainian.pdf> (дата звернення 28.10.2022)
3. Винцюк Т. Н. Образный компьютер. *Современные проблемы в компьютерных науках. Львовская политехника*. Львов. 2000. С. 5-14
4. Джелинек Ф. Распознавание непрерывной речи с помощью статистических методов. *Статистическая теория распознавания речи*. 1976. Т.64. С.134-160.
5. Bo C. (2009) Design and realization of an intelligent access control system based on voice recognition. *Communication, control and Management. IEEE International Conference*. 229–232 [in English].
6. Itakura F. (1975) Minimum prediction residual principle applied to speech recognition. *IEEE*., 67-72 [in English].
7. Hinton G. (2012) Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *Signal Processing Magazine. IEEE*. 82-97[in English].
8. Ставицький І. О., Нескорodieва А. Р. Розробка сервісу розпізнавання мовлення Bambusa. *Вісник студентського наукового товариства ДонНУ імені Василя Стуса*. 2021. Том 1. Вип.13. С.294-297.
9. Розпізнавання мовлення. URL: https://uk.wikipedia.org/wiki/%D0%A0%D0%BE%D0%B7%D0%BF%D1%96%D0%B7%D0%BD%D0%B0%D0%B2%D0%B0%D0%BD%D0%BD%D1%8F_%D0%BC%D0%BE%D0%B2%D0%BB%D0%B5%D0%BD%D0%BD%D1%8F (дата звернення 28.10.2022)
10. Хорошко В.О. Аналіз символічних інформаційних потоків для визначення індивідуальних особливостей людини. *Сучасна спеціальна техніка*. 2011. №1 (24). С. 5-11.