

Vekhova Alexandra Danilovna, Samara State Technical University, 244, Molodogvardeyskaya Street, Samara, 443100, Russian Federation; Student of Information Systems Department. Tel. +7 846 278-43-11. E-mail: aleksav2014@yandex.ru

References

1. Galkin M.A., Semikin G.J. Hardware and software for diagnostics of human cardiovascular system using method of optical plethysmography of high resolution. *Gumanitarnyj vestnik*, 2013, no. 12 (14), pp. 2. (In Russ.)
2. Vyalkov A.I. Information technologies in healthcare. *Upravlenie i ekonomika zdavoohraneniya*. Moscow: GEOTAR-Media, 2009, pp. 8–9. (In Russ.)
3. *World Health Organization Recommendations on Physical Activity and Sedentary Lifestyle: A Brief Review*. Geneva: World Health Organization, 2020. URL: <https://chocmp.ru/wp-content/uploads/2022/07/9789240014909-rus.pdf> (accessed: 15.11.2023). (In Russ.)
4. Ermolaeva E.L., Gribina G.A. The Importance of water for humans. *Mezhdunarodnyj studencheskij nauchnyj vestnik*, 2018, no. 6, pp. 53. (In Russ.)
5. Kovalzon V.M., Dolgikh V.V. Regulation of the wakefulness-sleep cycle. *Nevrologicheskij zhurnal*, 2016, no. 6, pp. 316–322. (In Russ.)
6. What is a sleep tracker and how does it work? URL: <https://www.askona.ru/blog/suggestions/chto-takoe-treker-sna-i-kak-on-rabotaet/> (accessed: 11.11.2023). (In Russ.)
7. Balanced diet: how many calories should i eat per day? URL: <https://profilaktika.ru/for-population/profilaktika-zabolevaniy/vse-o-pravilnom-pitanii/skolko-kaloriy-nuzhno-est-v-den/> (accessed: 12.11.2023). (In Russ.)
8. Healthy eating. URL: <https://www.who.int/ru/news-room/fact-sheets/detail/healthy-diet> (accessed: 12.11.2023). (In Russ.)
9. Physical activity. URL: <https://www.who.int/ru/news-room/fact-sheets/detail/physical-activity> (accessed: 13.11.2023). (In Russ.)
10. Physical activity awareness week. URL: <https://volosovocrb.ru/novosti/404-nedelya-informirovaniya-o-vazhnosti-fizicheskoy-aktivnosti.html> (accessed: 13.11.2023). (In Russ.)
11. Fitness and healthy eating trainer bot. URL: <https://t.me/gymbosbot> (accessed: 15.10.2023). (In Russ.)
12. Bot for conducting workouts. URL: https://t.me/gym_ru_bot (accessed: 15.10.2023). (In Russ.)
13. Bot providing advice on nutrition and workouts. URL: https://t.me/Sport_nutrition_Bot (accessed: 15.10.2023). (In Russ.)
14. Sleep quality improvement chatbot. URL: <https://t.me/sleepyrobot> (accessed: 15.10.2023). (In Russ.)

Received 22.11.2023

УДК 372.811.1

РАЗРАБОТКА ПРИЛОЖЕНИЯ ДЛЯ АНАЛИЗА КОММЕНТАРИЕВ АНГЛОЯЗЫЧНОГО YOUTUBE

Дукальская И.В., Киргизов А.Г.

Поволжский государственный университет телекоммуникаций и информатики, Самара, РФ

E-mail: i.dukalskaya@psuti.ru, kirgizoffand@yandex.ru

Данная статья представляет разработку инновационного приложения, цель которого – сбор и анализ комментариев на англоязычных видеороликах YouTube с целью выявления и систематизации наиболее часто употребляемых англоязычных аббревиатур. Это приложение не только облегчает процесс изучения английского языка, но и предоставляет глубокое понимание использования аббревиатур в контексте реальных диалогов, способствуя более эффективному их использованию в повседневной коммуникации. Позволяет взглянуть на английский язык со стороны носителя. Подробно рассматриваются особенности разработки приложения, такие как использование YouTube Data API, анализ текста комментариев с помощью регулярных выражений и сохранение данных для

дальнейшего изучения. Данная работа уделяет внимание вызовам при работе с API (Application Programming Interface), методам обработки текста и показывает возможности исследования английского языка в онлайн-среде.

Ключевые слова: методы извлечения данных, сохранение данных, анализ комментариев, изучение английского языка, обработка текста, онлайн платформы, англоязычные аббревиатуры

Введение

В условиях современного мира, где культурные преграды сокращаются благодаря глобализации и интернету, изучение иностранного языка становится ключевой целью. Чтобы обогатить свой словарный запас и научиться понимать язык, мы ищем новые методы изучения. Один из перспективных подходов – анализ реальных диалогов и комментариев на англоязычных платформах, таких как, например, YouTube. Это не только открывает нам доступ к разнообразным языковым проявлениям, но и позволяет изучать язык в контексте реальных общений. В данной статье речь пойдет о разработке инновационного приложения, специализирующегося на анализе комментариев к видеороликам на YouTube. Это приложение предлагает возможность выявления и выделения наиболее часто используемых англоязычных сокращений. Оно не только облегчает процесс изучения английского языка, но и дает более глубокое понимание использования сокращений в реальных диалогах, что помогает более эффективно освоить их для повседневного использования.

Это приложение является не только полезным инструментом для изучения языка, но также открывает новые возможности для исследования и использования английского языка. Позволяет посмотреть на его некоторые аспекты новым взглядом [4; 5].

Особенности разработки приложения

YouTube Data API (Application Programming Interface) является основой для функционирования приложения. Это API предоставляет доступ к базе данных комментариев, оставленных на англоязычных видеороликах YouTube. Приложение использует ключ разработчика (*developer_key*), получаемый каждым пользователем индивидуально, и уникальный идентификатор видеоролика (*video_id*), чтобы понимать, с каким конкретным видеороликом идет работа. Благодаря всему этому осуществляются запросы к API с помощью библиотеки *googleapiclient* и обеспечивается эффективный и быстрый сбор информации о комментариях (рисунок 1) [7].

Одной из главных особенностей являются ограничения, установленные самим YouTube на частоту запросов. Необходимо контролировать

количество запросов к API, чтобы избежать блокировок или ограничений доступа к данным. Также важно учитывать, что некоторые данные могут быть ограничены, что может повлиять на полноту или доступность информации при использовании API.

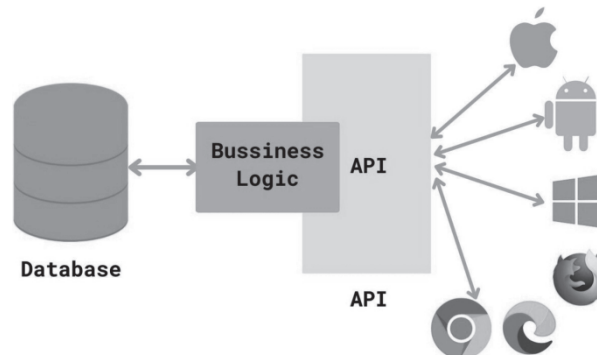


Рисунок 1. Схема API

Еще одной трудностью при работе с API является обработка большого объема данных, особенно при загрузке комментариев и ответов к ним. Необходимо эффективно организовывать код и процесс сбора данных, чтобы избежать возможных ошибок и обеспечить оптимальную производительность приложения.

Кроме того, важно учитывать обновления и изменения в API, так как YouTube периодически вносит изменения в свои правила и функционал. Это может потребовать соответствующих обновлений в коде приложения для поддержания его работоспособности и соответствия актуальным требованиям платформы. Для решения данной проблемы был использован такой способ создания приложения как версионирование. При каждом обновлении приложения выпускается новая версия. Устаревшие методы помечаются как *deprecated*, а новые вносятся в документацию. Тем самым мы получаем новый функционал, не нарушая работу старого.

Таким образом, использование YouTube Data API предоставляет важные возможности для приложения, но требует внимательного подхода к управлению запросами, обработке данных и поддержанию соответствия изменениям в API, чтобы обеспечить надежную и эффективную работу приложения.

Одним из ключевых направлений деятельности приложения является глубокий анализ текста комментариев, оставленных на англоязычных видео-

роликах на платформе YouTube. Основной целью является извлечение и систематизация англоязычных аббревиатур и сокращений из этих комментариев, с целью выявления наиболее часто встречающихся языковых элементов. Для достижения этой цели прибегаем к применению методов обработки текста и регулярных выражений, что позволяет нам точно выделять и анализировать аббревиатуры, используемые в комментариях.

Процесс анализа текста комментариев включает в себя создание специализированных алгоритмов, способных эффективно выявлять и извлекать аббревиатуры из текстового контента. Уделяется внимание разработке и оптимизации этих алгоритмов, чтобы обеспечить точность и полноту извлечения аббревиатур, а также эффективность обработки больших объемов текста.

Основными инструментами для извлечения аббревиатур и сокращений являются регулярные выражения, позволяющие проводить гибкий и точный поиск заданных шаблонов текста. Используются разнообразные шаблоны регулярных выражений с учетом различных вариантов написания и использования аббревиатур, что позволяет нам оптимально охватить разнообразие аббревиатур, применяемых в комментариях на YouTube.

Однако анализ текста комментариев не лишен сложностей. Разнообразие стилей и контекстов использования аббревиатур, а также их многообразие и динамичность могут представлять вызовы при точной классификации и интерпретации. Это требует постоянного совершенствования методов обработки текста и алго-

ритмов извлечения данных для более точного и всестороннего анализа.

Таким образом, приложение активно исследует и анализирует текстовый контент комментариев, используя методы обработки текста и регулярные выражения для точного выделения англоязычных аббревиатур и сокращений, несмотря на некоторые сложности и особенности данного процесса.

Проанализированные данные сохраняются в файл в формате CSV (формат Microsoft Excel). В этом файле содержится информация о комментариях, включая `etag`, `id`, текст комментария, автор и другие сведения. Эти данные могут быть использованы для дальнейшего изучения и анализа.

Процесс разработки

Процесс сбора комментариев с англоязычных видеороликов на YouTube начинается с использования YouTube Data API в рамках разработанного приложения на языке программирования Python.

Инициализация API и отправка запросов: первым шагом в процессе сбора комментариев является инициализация YouTube Data API с помощью ключа разработчика (DEVELOPER_KEY). Посредством вызова метода `youtube.commentThreads().list()` с указанием параметров, таких как `part`, `maxResults`, `pageToken`, и `videoId` для определения идентификатора видеоролика (VIDEO_ID), осуществляется запрос к API для получения корневых комментариев.

Циклический процесс загрузки комментариев: после отправки первого запроса на получение комментариев для указанного видеоролика про-

1	textOriginal
2	Book an English class on LiveXP for \$0.99: https://bit.ly/livexp_antonio
3	Use my discount code ANTONIO30 to get 30% off. 🎁
4	I am lowkey watching your all videos . Is that good Anton?
5	Your explanation was legit! 🤩
6	Thanks
7	Nice ☐
8	Love it♥
9	The origin of to bail is the act of jumping out of an airplane, with a parachute we hope.
10	You are excellent
11	I really like all your videos. You should be a teacher If you are not already one
12	♥♥♥♥
13	Thanks 😊
14	I like your short hair!
15	I super thanks you.
16	Your videos are absolutely amazing - congrats!
17	Thanks for sharing these useful slang words your voice is so nice and clear .I really love your videos.

Рисунок 2. Пример сохранения комментариев

исходит циклический процесс, в котором используется параметр `nextPageToken` для получения последующих страниц комментариев. Для этого повторяем запросы, указывая `pageToken` равным значению `nextPageToken`, чтобы получить следующую порцию комментариев. Это продолжается до тех пор, пока `nextPageToken` не станет равным `None`, что означает достижение конца списка комментариев.

Сохранение комментариев: полученные комментарии сохраняются в переменную `items`. Каждый комментарий представлен в формате JSON и содержит различные атрибуты, такие как идентификаторы, текст комментария, информацию об авторе и времени публикации.

Дополнительная загрузка ответов на комментарии (далее – «реплаев»): помимо извлечения корневых комментариев, приложение проверяет наличие реплаев к каждому комментарию. Если у корневого комментария есть реплаи, используется аналогичный процесс для загрузки реплаев на основании их `parentId`, чтобы получить дополнительные комментарии, относящиеся к ним.

Сохранение данных в файл CSV: по завершении сбора комментариев и их реплаев приложение сохраняет полученные данные в файл формата CSV. Для этого используется модуль `csv` для создания файла `youtuberresults.csv` и записи всех собранных комментариев и реплаев в этот файл с определенной структурой и разделителями между данными (рисунок 2) [9].

Таким образом, процесс сбора комментариев с англоязычных YouTube-видеороликов осуществляется путем систематического использования YouTube Data API, опираясь на различные параметры запросов и обработку полученных данных для сохранения их в файле для последующего анализа и обработки.

Пример использования

Для того чтобы протестировать приложение, было выбрано 3 видео с платформы YouTube по различной тематике. Для процесса анализа текста был написан скрипт, который скачивает список аббревиатур и сленга с OpenApi NLTK, ищет количество вхождений слов во всех комментариях, считает их соотношение к общему тексту [8; 10]. Пример работы кода с тестовым видео:

По результату первого анализа:

Slang Words, their Percentage, and Counts:

Omg: 0.03% (1 occurrences)

Lol: 0.03% (1 occurrences)

bro: 0.03% (1 occurrences)

Dude: 0.03% (1 occurrences)

Total Words: 3,095

По результату второго анализа:

lol: 0.11% (4 occurrences)

Bro: 0.03% (1 occurrences)

idk: 0.03% (1 occurrences)

Lol: 0.06% (2 occurrences)

mate: 0.03% (1 occurrences)

bro: 0.06% (2 occurrences)

Brb: 0.03% (1 occurrences)

dude: 0.03% (1 occurrences)

lmao: 0.03% (1 occurrences)

OMG: 0.03% (1 occurrences)

Idk: 0.03% (1 occurrences)

Dude: 0.03% (1 occurrences)

Total Words: 3,556

По результату третьего анализа:

Slang Words, their Percentage, and Counts:

lol: 0.01% (1 occurrences)

CHILL: 0.01% (1 occurrences)

dude: 0.02% (2 occurrences)

chill: 0.05% (5 occurrences)

Lol: 0.01% (1 occurrences)

OMG: 0.02% (2 occurrences)

Omg: 0.01% (1 occurrences)

Props: 0.01% (1 occurrences)

Total Words: 10,530

В данной статистике видно, что большей популярностью пользуются сокращения для неформальной речи. Далее мы можем узнать их значение, посмотреть примеры использования и начать употреблять их в своей речи, тем самым улучшая свой уровень английского. Таким же образом можно анализировать технические и профессиональные видео, собирая тематическую лексику [1; 2; 3].

Заключение

Это приложение, совмещающее в себе технологии обработки текста и доступ к обширной базе данных комментариев на YouTube, открывает перед пользователями уникальную возможность погружения в англоязычную среду. Анализ комментариев на основе современных методов обработки текста и регулярных выражений позволяет выделить, систематизировать и изучить разнообразные аббревиатуры, активно используемые в реальных диалогах пользователей на этой платформе [6].

Это приложение не только обогащает словарный запас изучающих английский язык, но и стимулирует интерес к языку через интерактивное и практическое изучение. Анализ аббревиатур, наиболее часто встречающихся

в реальных комментариях, предоставляет не только знания о сленге и сокращениях, но и глубокое понимание контекста и специфики языка, используемого в повседневной коммуникации.

Эффективное использование современных методов анализа текста и доступ к обширной базе данных комментариев на YouTube делает это приложение значимым инструментом в контексте образования и изучения языка. Объединение технологии и образования позволяет приложению предложить новый и инновационный подход к изучению английского языка, делая этот процесс более интересным, доступным и эффективным для всех желающих.

Литература

1. Peters P. The Cambridge Dictionary of English Grammar. Cambridge: Cambridge University, 2004. 512 p.
2. Adams M. Slang: The People's Poetry. Oxford: Oxford University, 2009. 240 p.
3. Joshi M. Abbreviations in English. Independently published, 2019. 120 p.
4. 79 Incredible English Slang Words to Help you Understand Native Speakers. URL: <https://storylearning.com/blog/english-slang-words> (дата обращения: 20.11.23).
5. Slang Words: List of 100 Common Slang Words & Phrases you Need to Know! URL: <https://eslforums.com/slang-words/> (дата обращения: 20.11.23).
6. Re – Regular Expression Operations. URL: <https://docs.python.org/3/library/re.html> (дата обращения: 20.11.23).
7. Добавьте функциональность YouTube в свое приложение. URL: <https://developers.google.com/youtube/v3?hl=ru> (дата обращения: 22.11.23).
8. NLTK Documentation. Natural Language Toolkit. URL: <https://www.nltk.org> (дата обращения: 22.11.23).
9. CSV File Reading and Writing. URL: <https://docs.python.org/3/library/csv.html> (дата обращения: 23.11.23).
10. How to Learn Big Data? Step by Step Roadmap. URL: <https://www.mltut.com/how-to-learn-big-data-step-by-step/> (дата обращения: 24.11.23).

Получено 15.01.2024

Дукальская Ирина Владимировна, к.филол.н., доцент, доцент кафедры иностранных языков Поволжского государственного университета телекоммуникаций и информатики (ПГУТИ). 443090, Российская Федерация, г. Самара, Московское шоссе, 77. Тел. +7 927 204-77-93. E-mail: i.dukalskaya@psuti.ru

Киргизов Андрей Геннадьевич, студент кафедры программной инженерии ПГУТИ. 443090, Российская Федерация, г. Самара, Московское шоссе, 77. Тел. +7 932 861-85-21. E-mail: kirgizoffand@yandex.ru

DEVELOPMENT OF AN APPLICATION FOR ANALYZING COMMENTS ON ENGLISH-LANGUAGE YOUTUBE

Dukalskaya I.V., Kirgizov A.G.

Povolzhskiy State University of Telecommunications and Informatics, Samara, Russian Federation

E-mail: i.dukalskaya@psuti.ru, kirgizoffand@yandex.ru

This article introduces the development of an innovative application aimed at analyzing comments on English-language YouTube videos to identify and systematize the most frequently used English abbreviations. This application not only facilitates the process of learning English, but also provides a deep understanding of the use of abbreviations in the context of real dialogues, contributing to a more effective use of them in everyday communication. It provides a native speaker's perspective on the English language. Details the features of developing an application, such as using the YouTube Data API, analyzing comment text using regular expressions and saving data for further study. This paper focuses on the challenges of working with the API (Application Programming Interface), text processing techniques and shows the possibilities of exploring English in an online environment.

Keywords: *data extraction methods, data storage, comment analysis, English language learning, text processing, online platforms, English-language abbreviations*

DOI: 10.18469/ikt.2023.21.3.11

Dukalskaya Irina Vladimirovna, Povolzhskiy State University of Telecommunications and Informatics, 77, Moskovskoe shosse, Samara, 443090, Russian Federation; Associate Professor of Foreign Languages Department, PhD in Philology. Tel. +7 927 204-77-93. E-mail: i.dukalskaya@psuti.ru

Kirgizov Andrey Gennadievich, Povolzhskiy State University of Telecommunications and Informatics, 77, Moskovskoe shosse, Samara, 443090, Russian Federation; Student of Software Engineering Department. Tel. +7 932 861-85-21. E-mail: kirgizoffand@yandex.ru

References

1. Peters P. *The Cambridge Dictionary of English Grammar*. Cambridge: Cambridge University, 2004, 512 p.
2. Adams M. *Slang: The People's Poetry*. Oxford: Oxford University, 2009, 240 p.
3. Joshi M. *Abbreviations in English*. North Charleston: Independently published, 2019, 120 p.
4. 79 Incredible English Slang Words to Help you Understand Native Speakers. URL: <https://storylearning.com/blog/english-slang-words/> (accessed: 20.11.23).
5. Slang Words: List of 100 Common Slang Words & Phrases you Need to Know! URL: <https://eslforums.com/slang-words/> (accessed: 20.11.23).
6. Re – Regular Expression Operations. URL: <https://docs.python.org/3/library/re.html> (accessed: 20.11.23).
7. Add YouTube functionality to your app. URL: <https://developers.google.com/youtube/v3?hl=ru> (accessed: 22.11.23) (In Russ).
8. NLTK Documentation. Natural Language Toolkit. URL: <https://www.nltk.org> (accessed: 22.11.23).
9. CSV File Reading and Writing. URL: <https://docs.python.org/3/library/csv.html> (accessed: 23.11.23).
10. How to Learn Big Data. URL: <https://www.mltut.com/how-to-learn-big-data-step-by-step/> (accessed: 24.11.23).

Received 15.01.2024

ТЕХНОЛОГИИ РАДИОСВЯЗИ, РАДИОВЕЩАНИЯ И ТЕЛЕВИДЕНИЯ

УДК 004.722.45

ИНТЕГРИРОВАННАЯ КОСМИЧЕСКАЯ СЕТЬ: РЕЗУЛЬТАТЫ ПЕРВИЧНОГО МОДЕЛИРОВАНИЯ МАРШРУТИЗИРУЕМОЙ МЕЖСПУТНИКОВОЙ СЕТИ

Кустов Н.Д.

Сибирский государственный университет науки и технологий имени академика М. Ф. Решетнева,

Красноярск, РФ

E-mail: kustovnd@sibsau.ru

В работе приведены первые результаты, полученные при моделировании маршрутизируемой межспутниковой сети на основе входных данных космической миссии ReshUCube-3. В рамках миссии планируется запуск группировки из четырех малых космических аппаратов типа CubeSat. Для обеспечения концепции интегрированной космической сети на физическом уровне планируется использование технологии LoRa, а в качестве протокола маршрутизации на сетевом уровне – Routing Protocol for Low-Power and Lossy Networks. Модель учитывает орбитальные параметры спутниковой группировки, конфигурацию антенных систем космических аппаратов и наземной станции, параметры приемо-передающих радиомодулей. На основе данных параметров проведен анализ межспутниковых каналов связи. Определены максимальная достигаемая дистанция между спутниками, энергетика каналов, среднесуточное время покрытия с учетом существующих маршрутов и т.д. Планируется, что предложенная модель послужит основой в процессе проектирования и эксплуатации межспутниковой сети.

Ключевые слова: интегрированная космическая сеть, межспутниковая связь, ReshUCube, CubeSat, LoRa, Routing Protocol for Low-Power and Lossy Networks