

4. Deren, M. Essential Deren: collected writings of film. Documentext. Kingston, New York, 2005. – 265с.

УДК 004.772/.773:004.4'277

Урбанович Ю. П., Агиев А. М., Железнова М. А.

**«VOICE-TO-TEXT CONVERTER» — ПЛАГИН ДЛЯ ПЕРЕВОДА
ГОЛОСОВЫХ СООБЩЕНИЙ**

Юлия Павловна Урбанович

учитель информатики и ИКТ

urbanovich91@yandex.ru

Андрей Михайлович Агиев

обучающийся 11 класса

agiev29@gmail.com

Мария Андреевна Железнова

обучающаяся 11 класса

mms00180@gmail.com

Муниципальное автономное общеобразовательное учреждение гимназия № 99,

Россия, г. Екатеринбург, Россия

**«VOICE-TO-TEXT-CONVERTER» - PLUGIN FOR TRANSLATING VOICE
MESSAGES**

Yulia Pavlovna Urbanovich

Maria Andreevna Zheleznova

Andrey Mikhailovich Agiev

Grammar school №99, Russia, Yekaterinburg, Russia

Аннотация. В статье рассматривается разработка плагина по переводу голосовых сообщений в текст и его реализация в социальной сети ВКонтакте.

Abstract. The article discusses the development of a plugin to translate voice messages into text and its implementation in the social network VKontakte.

Ключевые слова: голосовые сообщения, социальная сеть ВКонтакте.

Keywords: Voice messages, social network VKontakte.

Тенденция развития программного обеспечения подразумевает включение не только базовых функций, но и дополнительного функционала, созданного сторонними разработчиками. Сейчас расширения для браузера могут использоваться совершенно для разных целей: отслеживание проведенного времени в социальных сетях или перевод голосовых сообщений в текст для более комфортного общения.

Голосовые сообщения — явление не новое. Трудно сказать, кто именно их изобрёл, однако точно известно, что первыми популярными мессенджерами, которые вооружились аудиосообщениями, стали «Вотсап» и «Фейсбук мессенджер». Пользователям инновация понравилась, и теперь от голосовых практически нигде нет спасения.

Сегодня эта опция есть в «Телеграме», «ВКонтакте», «Вайбере», «Инстаграме», «Скайпе» и ряде других популярных сервисов. Потому что есть спрос. По данным Facebook голосовые сообщения — второй по популярности тип файлов, которые передают друг другу его пользователи. Тем не менее, противников у голосовых сообщений едва ли не больше, чем сторонников, их регулярно высмеивают и посвящают им мемы — прямо скажем, не всегда корректные.

Но что делать, если голосовое сообщение отправлено не в очень удобное время и прослушать его нет возможности?

Расширение для браузера «Voice-to-text-converter» способно помочь решить проблему — сложность прослушивания голосовых сообщений, когда физически это сделать невозможно.

Основной задача плагина «Voice-to-text converter» - перевод голосового сообщения в текст по нажатию кнопки пользователем.

Плагин для браузера состоит из трёх основных компонентов:

- Плагин для браузера, реализующий функционал работы с пользователем;

- Linux-сервер, реализующий функционал перевода голосового сообщения в текст.
- HTTP-сервер, реализующий функционал взаимодействия Linux-сервера и плагина (рисунок 1).

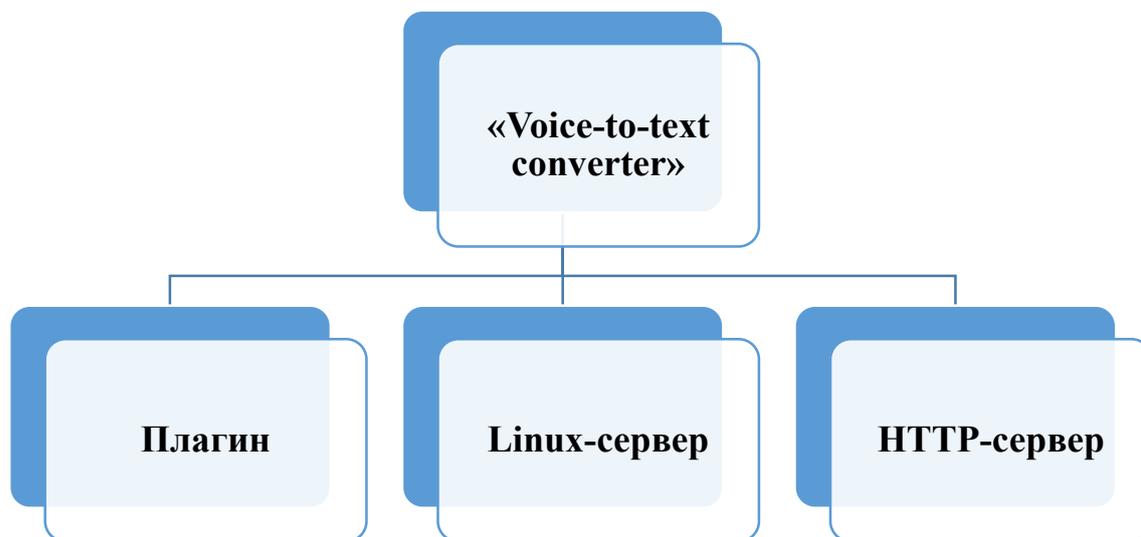


Рисунок 1 — Компоненты плагина для браузера «Voice-to-text converter»

Плагин «Voice-to-text converter» реализован с помощью языков программирования: Bash, JavaScript, PHP.

Рассмотрим подробнее реализацию каждого компонента плагина.

1. Плагин для браузера, реализован средствами JavaScript и дополнительного модуля jQuery.

При разработке плагина большое внимание уделялось подходу AJAX, который обеспечил фоновый обмен данными браузера с веб-сервером.

Для создания плагина необходим файл «manifest.json» (манифест). Манифест веб-приложения предоставляет информацию о приложении (такую как имя, авторство, иконку и описание) в формате JSON-файла. Цель манифеста — установить веб-приложение на домашний экран устройства, предоставляя пользователю более быстрый доступ и больше возможностей (рисунок 2).

```
1 {
2   "manifest_version": 2,
3
4   "name": "расширение тест 3",
5   "description": "Пзволяет перевести аудио сообщение в текст.",
6   "version": "1.0",
7
8   "content_scripts": [
9     {
10      "matches": ["*://vk.com/im*"],
11      "js": ["scripts.js"]
12    }
13  ]
14 }
```

Рисунок 2 — Программный код файла «manifest.json»

Основная js программа выполняет поиск аудиосообщений на странице сайта «vk.com/im/» и выводит кнопки у найденных сообщений для взаимодействия с пользователем, при нажатии которых отправляется ссылка на аудио на http сервер.

2. Linux-сервер, реализованный средствами BASH.

BASH (Bourne again shell) — одна из наиболее популярных современных разновидностей командной оболочки UNIX. Bash представляет собой командный процессор, работающий в интерактивном режиме в текстовом окне. Как и все Unix-оболочки, он поддерживает автодополнение имён файлов и каталогов, подстановку вывода результата команд, переменные, контроль над порядком выполнения, операторы ветвления и цикла [2].

Для перевода аудио файла в текст используется Google Speech API.

Google Cloud Speech-to-Text API позволяет разработчикам конвертировать аудио в текст, применяя средства нейронных сетей в API. Он может обрабатывать потоковое или предварительно записанное аудио в режиме реального времени, используя технологию машинного обучения Google [5].

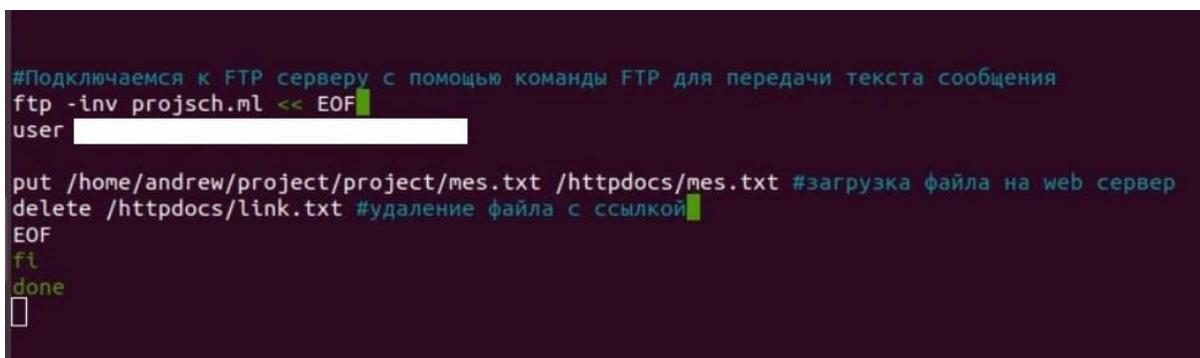
Плагин «Voice-to-text converter» работает на бесконечном цикле: при обнаружении файла на ftp сервере, программа получает его и отправляет его на сервер Google API для его конвертации, сконвертированный JSON файл программа преобразовывает в txt файл, а затем удаляет все «хвосты».

3. HTTP-сервер, реализованный средствами PHP.

HTTP-сервер — это сервер, работающий по HyperText Transfer Protocol. Используется для обмена произвольными данными между компьютерами по сети Интернет. Эта технология является одной из самых востребованных для скачивания и загрузки данных с/на удаленные серверы, рассредоточенные по всему миру [3].

Для реализации возможности передачи файла на сервер, а также сохранения ссылки сообщения в уже существующий файл на веб-сервере используются PHP файлы, находящиеся на сервере.

Передача текстового файла с сервера Linux происходит с помощью стандартного консольного http клиента в Linux.



```
#Подключаемся к FTP серверу с помощью команды FTP для передачи текста сообщения
ftp -inv projsch.nl << EOF
user [REDACTED]

put /home/andrew/project/project/mes.txt /httpdocs/mes.txt #загрузка файла на web сервер
delete /httpdocs/link.txt #удаление файла с ссылкой
EOF
ft
done
█
```

Рисунок 3 — Программный код, осуществляющий передачу файла на http сервер

Работа плагина «Voice-to-text converter» выглядит следующим образом:

Пользователь, установивший плагин, заходит в личные сообщения в ВКонтакте, при наличии голосовых сообщений в личной переписке появляется кнопка для перевода голосового сообщения. По нажатию на данную кнопку, директория сервера ВКонтакте, где лежит сам аудиофайл, передается в php файл, с помощью которого ссылка аудиофайла сохраняется на http сервере.

В это же время на Linux-сервере работает программа, которая после обнаружения файла, получает его и с помощью API Google переводит его в текст, отправляя обратно на http сервер. Плагин получает файл и выводит текст в модальном окне браузера (рисунок 4).

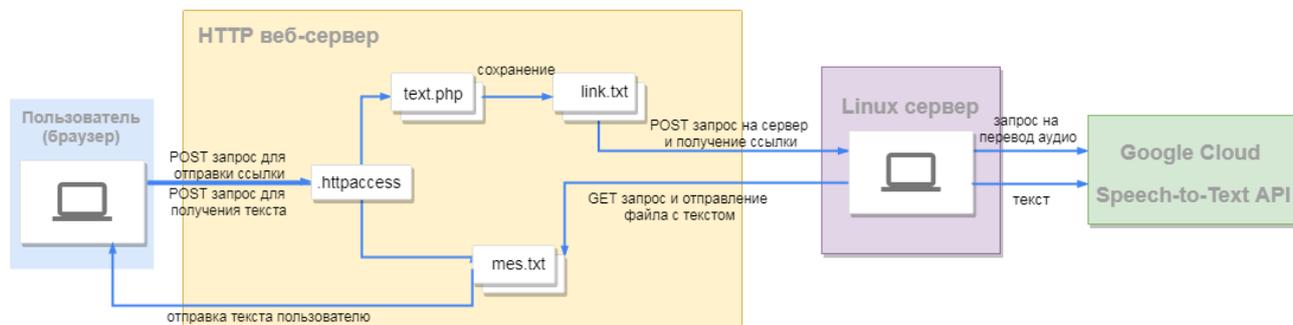


Рисунок 4 — Схема работы плагина

Для реализации плагина на своем персональном компьютере, определены следующие технические требования:

1. Браузеры семейства Chromium (Google Chrome, Yandex browser, Mozilla Firefox, Opera).
2. Объем памяти не менее 2 МБ.

Голосовые сообщения выглядят не самым логичным поворотом и как будто возвращают нас к более «старым» типам общения — долгим объяснениям голосом вместо пары строк текста. И тем не менее многие выбирают именно их. Число голосовых сообщений очень быстро выросло, особенно за последние два года. Например, в 2019 году количество голосовых сообщений увеличилось больше чем на пятьдесят процентов.

Для пользователей, которые активно используют социальные сети, данный плагин станет незаменимым помощником, в ситуациях, когда человек может находиться в транспорте, на совещании, в толпе или любом другом месте, где слушать голосовое сообщение просто дискомфортно.

Список литературы

1. Блог компьютерных советов it-doc. – URL: <https://it-doc.info/>.
2. Журнал TJournal — Википедия. – URL: <https://tjournal.ru/>.
3. Научно-популярный журнал Хабр. – URL: <https://habr.com/ru/>.

4. Основы JavaScript. – URL: https://developer.mozilla.org/ru/docs/Learn/Getting_started_with_the_web/JavaScript_basics.

5. Что такое API в Веб-приложениях и зачем он нужен. – URL: <https://mkdev.me/posts/chto-takoe-api-v-veb-prilozheniyah-i-zachem-on-nuzhen>.

6. Что такое API?. – URL: <https://habr.com/ru/sandbox/52599>.

7. Что такое PHP?. – URL: <http://php.net/manual/ru/intro-what-is.php>.

УДК [78.02:004]+78.071.5

Фильчаков С. И., Буторина Н. И.

**МУЗЫКАЛЬНО-КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ:
ТРАКТОВКИ ПОНЯТИЯ, ПРОБЛЕМЫ, ВОЗМОЖНОСТИ**

Сергей Иванович Фильчаков,

преподаватель

sergeipc@bk.ru

Школа искусств, Новоуральск

Наталья Иннокентьевна Буторина

кандидат педагогических наук, доцент

nainnrgppri@mail.ru

*ФГАОУ ВО «Российский государственный профессионально-педагогический
университет», Россия, Екатеринбург*

**MUSIC-COMPUTER TECHNOLOGIES: INTERPRETATIONS
OF CONCEPTS, PROBLEMS, OPPORTUNITIES**

Sergey Ivanovich Filchakov

Children's Art School, Novouralsk

Natalia Innokentievna Butorina

Russian State Vocational Pedagogical University, Russia, Ekaterinburg