VOICE2TEXT AI

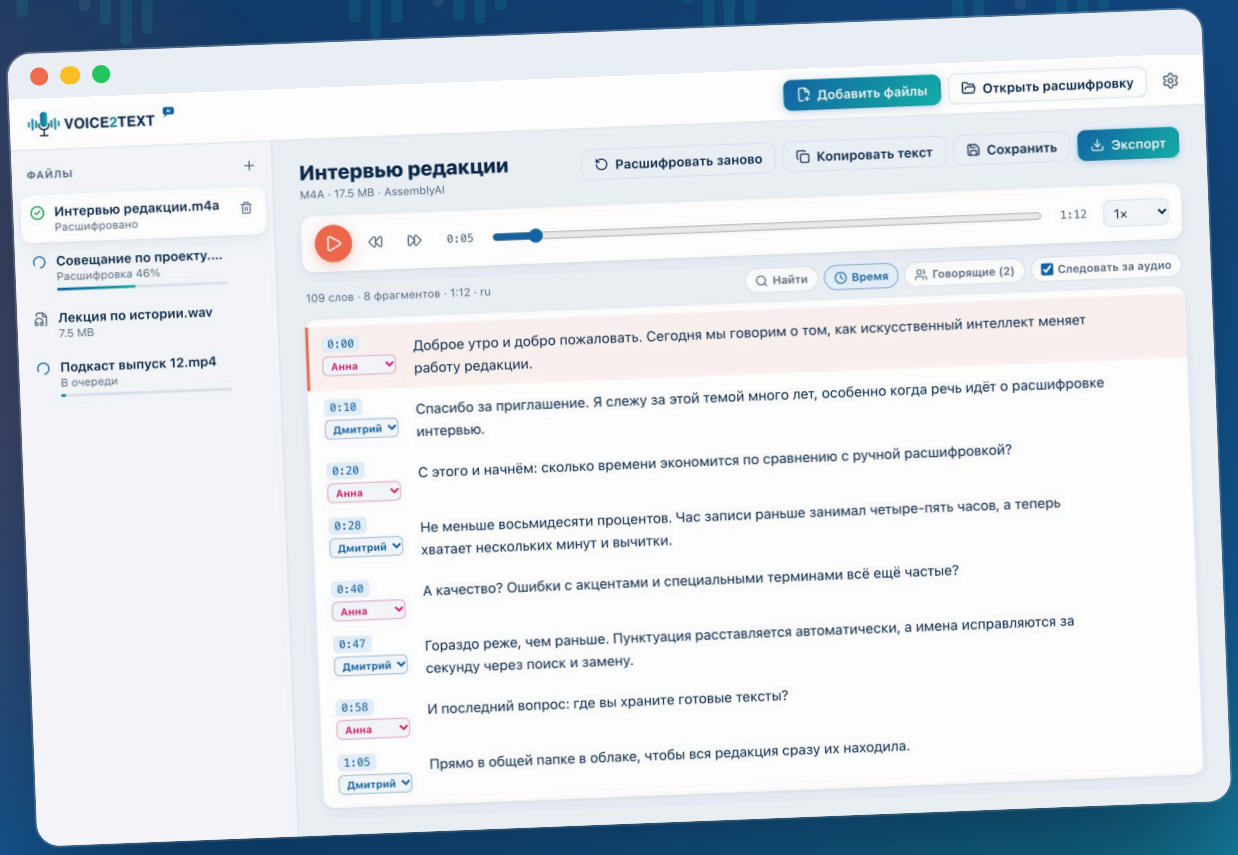
РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ

Voice2Text

От голоса к тексту — на любом языке.

Расшифровывайте аудио- и видеофайлы с помощью искусственного интеллекта, исправляйте текст и сохраняйте его куда угодно.

Офлайн-Whisper | 12 форматов | Облако | 9 языков



Версия 1.1 | macOS · Windows · Linux | Русский

Содержание

01	Добро пожаловать в Voice2Text	3
02	Установка	4
03	Главное окно	5
04	Выбор движка и языка	6
05	Онлайн-сервисы: аудио покидает компьютер	8
06	Во время расшифровки	10
07	Правка расшифровки	11
08	Сохранение и экспорт	13
09	Сохранение в облако	15
10	Настройки и языки	17
11	Сочетания клавиш	19
12	Частые вопросы	20

01 Добро пожаловать в Voice2Text

> Voice2Text превращает аудио- и видеозаписи в текст. Интервью, совещания, лекции, подкасты, голосовые сообщения: перетащите файл в окно, выберите движок ИИ — и через несколько минут текст готов к правке и сохранению.

Возможности

- **Любой формат.** MP3, M4A, WAV, FLAC, OGG, AAC, AIFF, видео MP4, MOV, MKV, WebM и многие другие; с установленным ffmpeg — практически все.
- **Даже без интернета.** Whisper работает прямо на компьютере: аудио никуда не уходит.
- **Кто говорит.** AssemblyAI и Deepgram разделяют голоса, и каждому собеседнику можно дать имя.
- **Удобная правка.** Текст разбит на фразы, связанные с аудио: один щелчок — и вы снова слышите нужное место.
- **Любой формат на выходе.** Word, PDF, ODT, RTF, TXT, Markdown, HTML, субтитры SRT и VTT, CSV, JSON и проект, который можно открыть снова.
- **Сразу в облако.** Google Drive, OneDrive, Dropbox, Box, Proton Drive, iCloud и другие — через их папки на компьютере.
- **Девять языков.** Интерфейс на итальянском, английском, французском, испанском, немецком, китайском, японском, русском и арабском; переводы — это XML-файлы, которые можно исправлять.

Системные требования

macOS 11 или новее (Mac с Apple или Intel), 64-битная Windows 10/11 или 64-битный Linux (Ubuntu, Debian, Fedora и производные). Для офлайн-Whisper желательно не менее 8 ГБ памяти; онлайн-сервисам нужно только подключение.

02 Установка

> Скачайте файл для вашей системы и установите его как любую другую программу.

macOS

Откройте `Voice2Text_1.1.1_universal.dmg` и перетащите значок Voice2Text в папку **Программы**. При первом запуске macOS может сообщить, что разработчик не проверен: откройте **Системные настройки** → **Конфиденциальность и безопасность** и выберите **Всё равно открыть**, либо щёлкните приложение правой кнопкой и выберите **Открыть**. Это нужно сделать один раз.

Windows

Запустите `Voice2Text_1.1.1_x64-setup.exe`. Если появится синее окно Windows SmartScreen, нажмите **Подробнее**, а затем **Выполнить в любом случае**. Права администратора не нужны.

Linux

В Ubuntu и Debian установите пакет `.deb` (двойной щелчок или `sudo apt install ./Voice2Text_1.1.1_amd64.deb`), в Fedora — пакет `.rpm`. Можно также сделать `.AppImage` исполняемым (`chmod +x`) и запустить его.

> **Необязательно: ffmpeg.** Voice2Text сам читает самые распространённые форматы. Для редких (OPUS, WMA, AMR, AVI…) установите ffmpeg: `brew install ffmpeg` на Mac, `winget install ffmpeg` в Windows, `sudo apt install ffmpeg` в Linux. Он находится автоматически.

03 Главное окно

> Вся работа идёт в одном окне: слева список файлов, в центре выбранный файл.

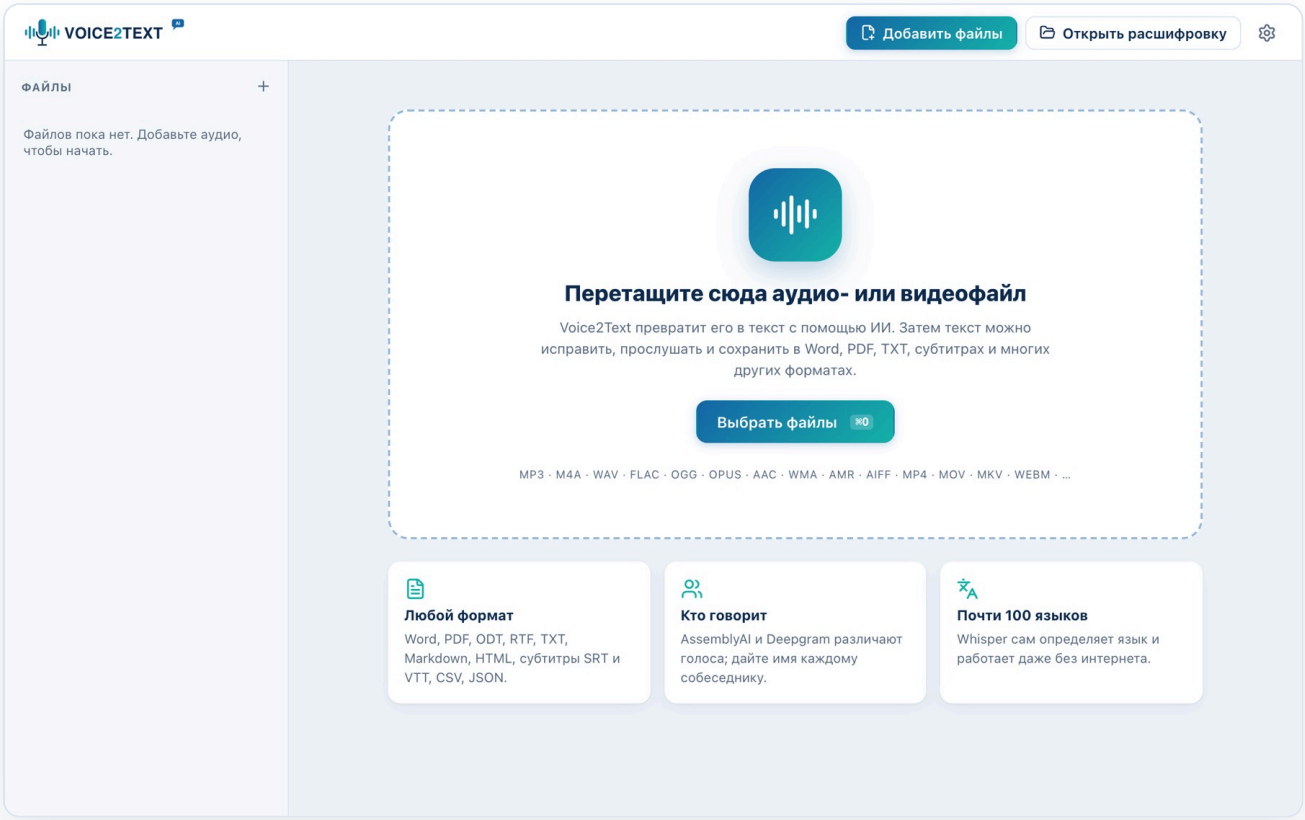


Рисунок 1. Начальный экран: перетащите файл в рамку или нажмите «Выбрать файлы».

- **Верхняя панель**: **Добавить файлы**, **Открыть расшифровку** (сохранённый проект `.v2t`) и шестерёнка **Настроек**.
- **Список файлов** слева: у каждого файла видны состояние (готово, в очереди, расшифровка, расшифровано, ошибка) и ход работы. Точка после имени означает несохранённые изменения.
- **Центральная область**: настройки расшифровки, ход работы или текст для правки.

Добавление файлов

Перетащите в окно один или несколько аудио- или видеофайлов либо нажмите **Добавить файлы** (⌘O / Ctrl+O). Файлов может быть сколько угодно: они расшифровываются по очереди. Перетащив файл `.v2t`, вы снова откроете сохранённую расшифровку.

04 Выбор движка и языка

> Перед расшифровкой выберите, кто будет работать: Whisper на вашем компьютере или онлайн-сервис.

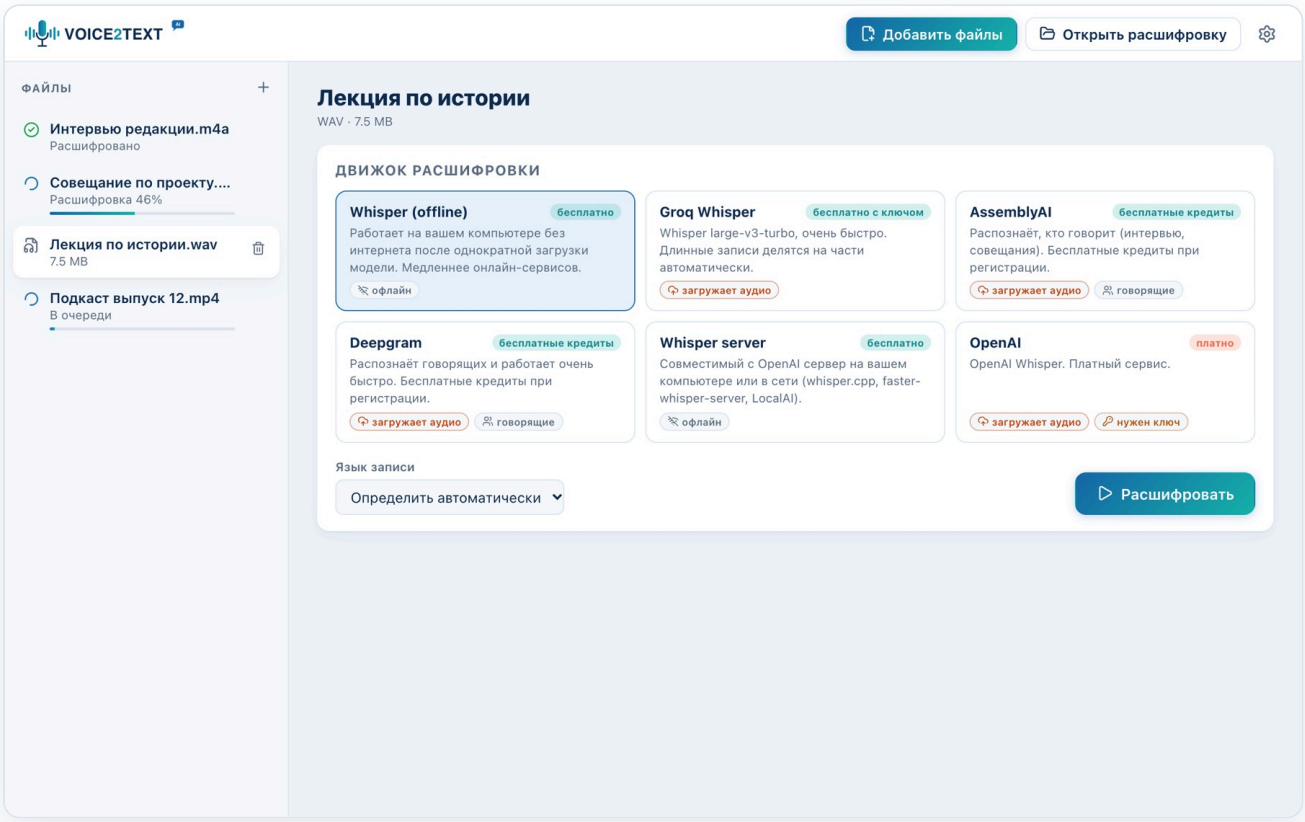


Рисунок 2. Выбор движка и языка для выбранного файла.

Движок	Стоимость	Особенности
Whisper (офлайн)	бесплатно	Работает без интернета; модель скачивается один раз (от 45 до 600 МБ). Медленнее онлайн-сервисов.
Groq Whisper	бесплатно с ключом	Очень быстро; длинные записи делятся на части автоматически.
AssemblyAI	бесплатные кредиты	Распознаёт, кто говорит. Идеально для интервью и совещаний.
Deepgram	бесплатные кредиты	Распознаёт говорящих и работает очень быстро.
Whisper server	бесплатно	Сервер Whisper на вашем компьютере или в сети (whisper.cpp, faster-whisper-server, LocalAI).
OpenAI	платно	Whisper от OpenAI.

Язык записи

В режиме **Определить автоматически** Voice2Text сам распознаёт язык речи. Если вы его знаете, выберите вручную: результат будет точнее, особенно в первые секунды.

Ключи онлайн-сервисов

Groq, AssemblyAI, Deepgram и OpenAI требуют личный ключ: в разделе **Настройки → Движки** есть поле для каждого и ссылка «получить ключ». Регистрация в Groq бесплатна; AssemblyAI и Deepgram дарят кредиты при регистрации.

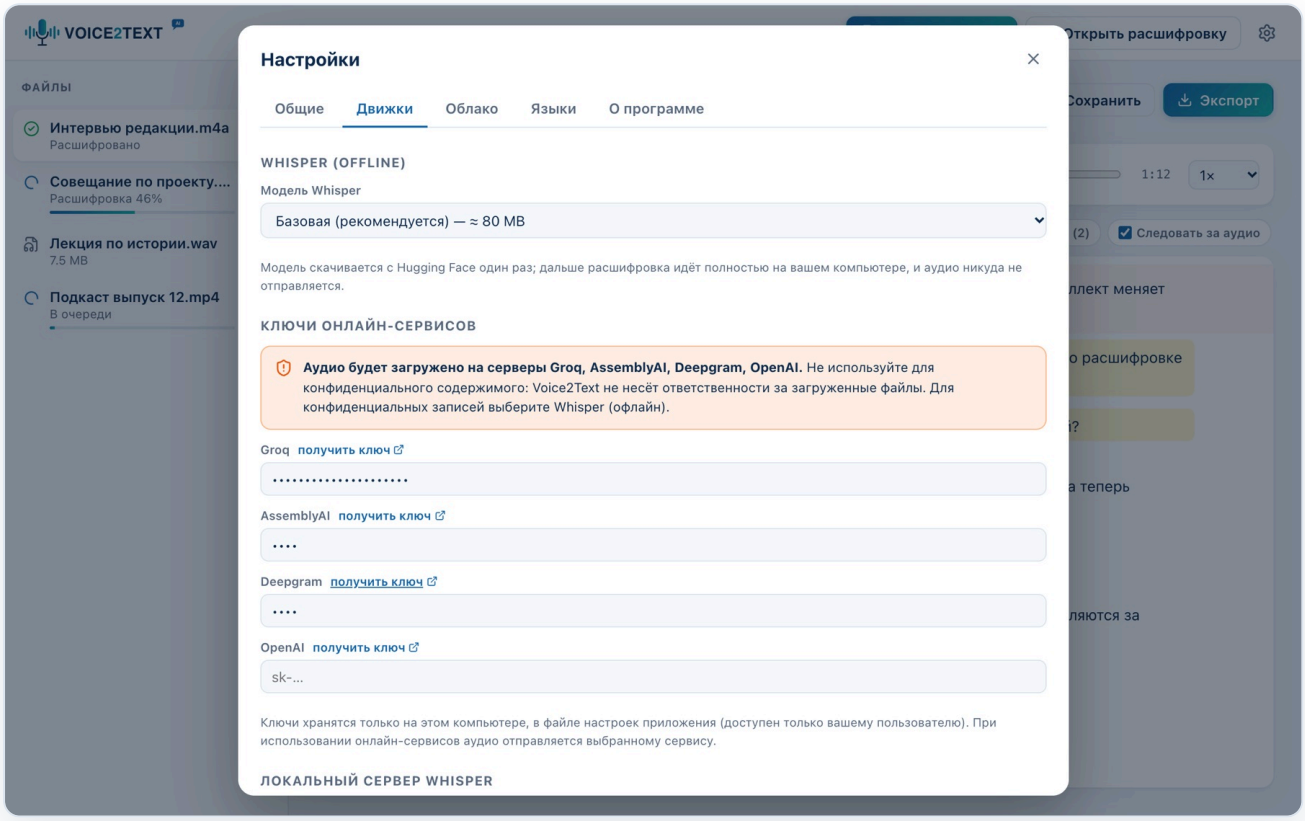


Рисунок 3. Настройки → Движки: модель Whisper и ключи сервисов.

Нажмите **Расшифровать**, чтобы начать. Если ждут несколько файлов, на боковой панели появится и кнопка **Расшифровать все**.

05 Онлайн-сервисы: аудио покидает компьютер

> При использовании онлайн-сервиса ИИ аудиофайл загружается в интернет. Voice2Text чётко предупреждает об этом и просит подтверждения.

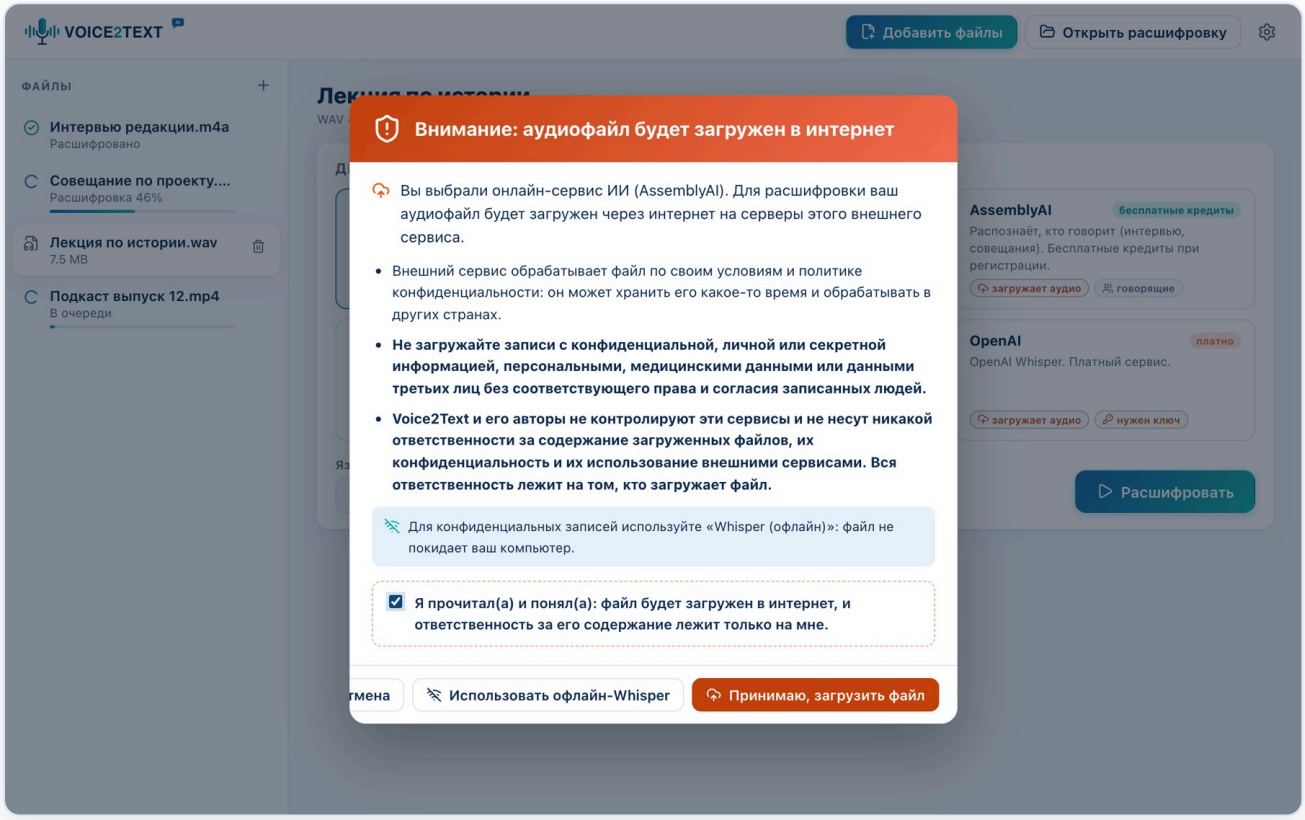


Рисунок 4. Обязательное предупреждение перед первой загрузкой в онлайн-сервис.

При первом использовании Groq, AssemblyAI, Deepgram или OpenAI появляется это предупреждение. Чтобы продолжить, отметьте **«Я прочитал(а) и понял(а)»**; можно также сразу перейти на **офлайн-Whisper**.

Аудиофайл загружается на серверы внешнего сервиса, который обрабатывает его по своим условиям и политике конфиденциальности: он может хранить файл какое-то время и обрабатывать его в других странах.

Не загружайте записи с конфиденциальной информацией, личными или секретными сведениями, персональными, медицинскими данными или данными третьих лиц без соответствующего права и согласия записанных людей.

Voice2Text и его авторы не несут никакой ответственности за содержание загруженных файлов, их конфиденциальность и их использование внешними сервисами. Вся ответственность лежит на том, кто загружает файл.

Согласие действует для конкретного сервиса и сохраняется с датой. Пока выбран онлайн-сервис, об этом напоминает оранжевая рамка. В разделе **Настройки → О программе** можно перечитать предупреждение и снова включить его показ.

> **Конфиденциальная запись?** Используйте **Whisper (офлайн)**: расшифровка идёт полностью на вашем компьютере.

06 Во время расшифровки

> Voice2Text в любой момент показывает, как идёт работа.

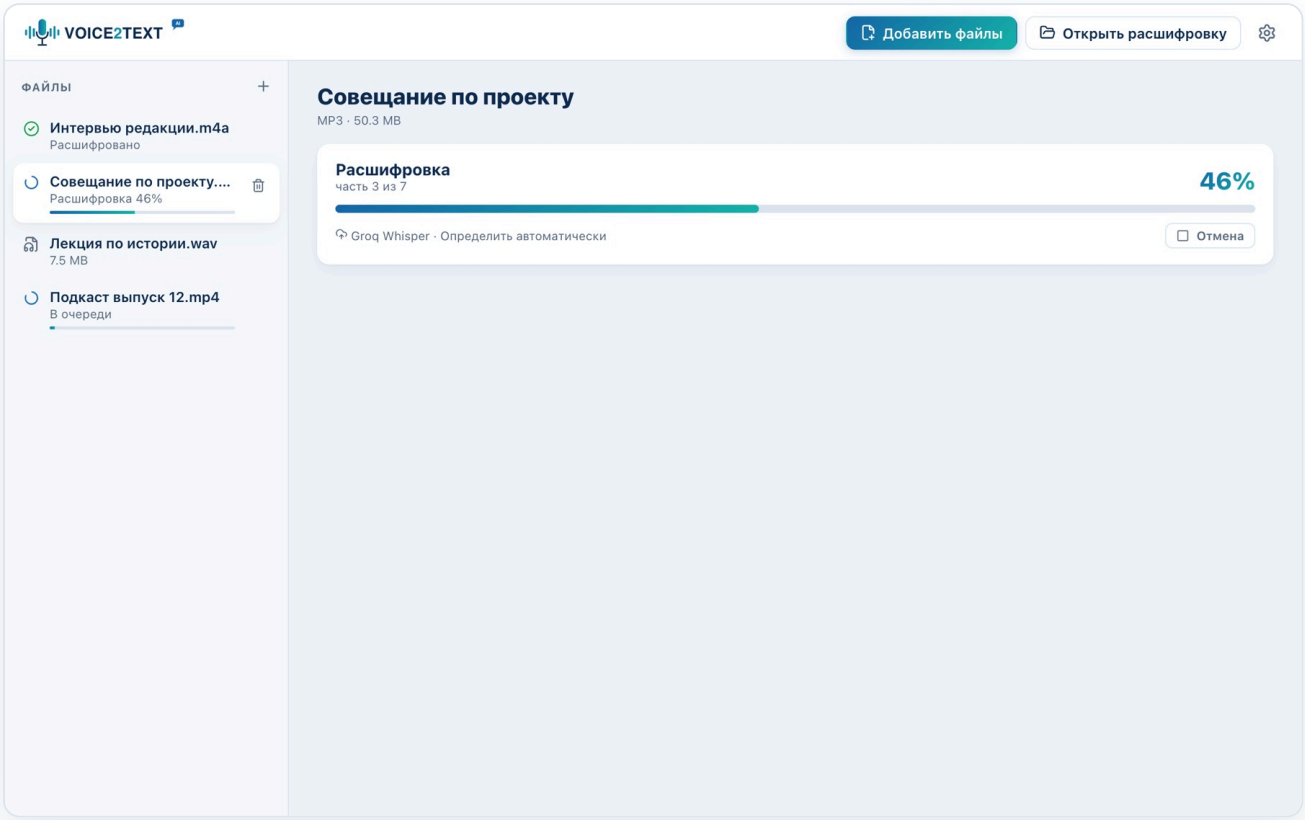


Рисунок 5. Расшифровка с помощью Groq: ход работы и обрабатываемая часть.

- **Чтение аудио**: файл декодируется на компьютере.
- **Подготовка модели** (только офлайн-Whisper, в первый раз): выбранная модель скачивается и сохраняется.
- **Отправка аудио** и **Расшифровка**: для длинных файлов видна текущая часть («часть 3 из 7»).
- **Ожидание сервиса**: на бесплатных тарифах при превышении лимита запросов Voice2Text ждёт и повторяет попытку сам.

Отменить можно в любой момент кнопкой **Отмена**; тем временем можно работать с другими файлами — когда текст будет готов, придёт уведомление.

07 Правка расшифровки

Текст разбит на фразы, у каждой есть время и, если распознан, говорящий.

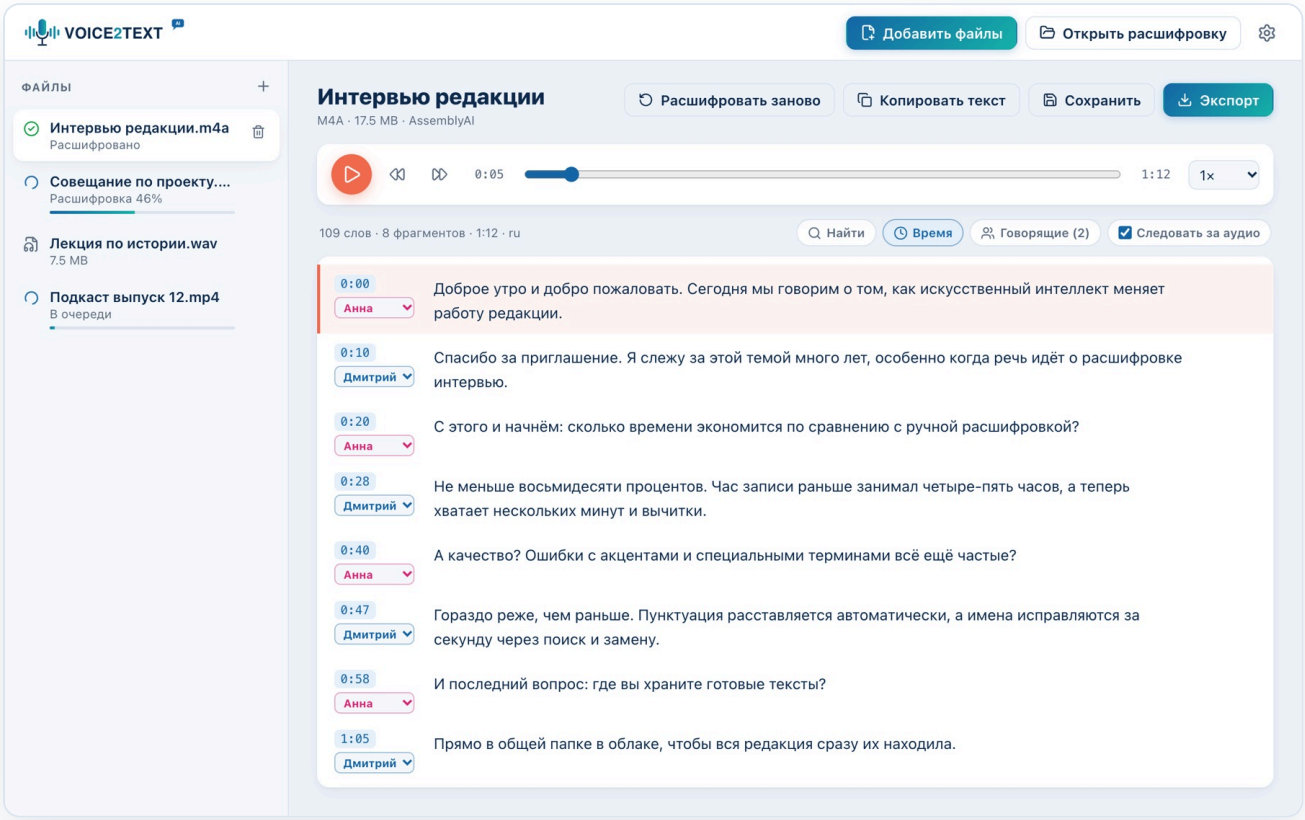


Рисунок 6. Редактор: сверху аудиоплеер, ниже редактируемые фразы; воспроизводимая фраза выделена.

Прослушивание

Плеер сверху воспроизводит исходное аудио: **Пробел** запускает и ставит на паузу (когда вы не печатаете), кнопки перематывают на 5 секунд, скорость можно выбрать от 0,75× до 2×. Щёлкните время фразы, чтобы прослушать её; с параметром **Следовать за аудио** воспроизводимая фраза всегда остаётся на экране.

Правка

- Щёлкните фразу и печатайте: изменения применяются сразу.
- **Enter** делит фразу в месте курсора; **Shift+Enter** добавляет перенос строки внутри фразы.
- **Backspace** в начале фразы объединяет её с предыдущей; кнопка **Объединить** соединяет её со следующей.
- Корзина удаляет фразу (например, шум, распознанный как слово).

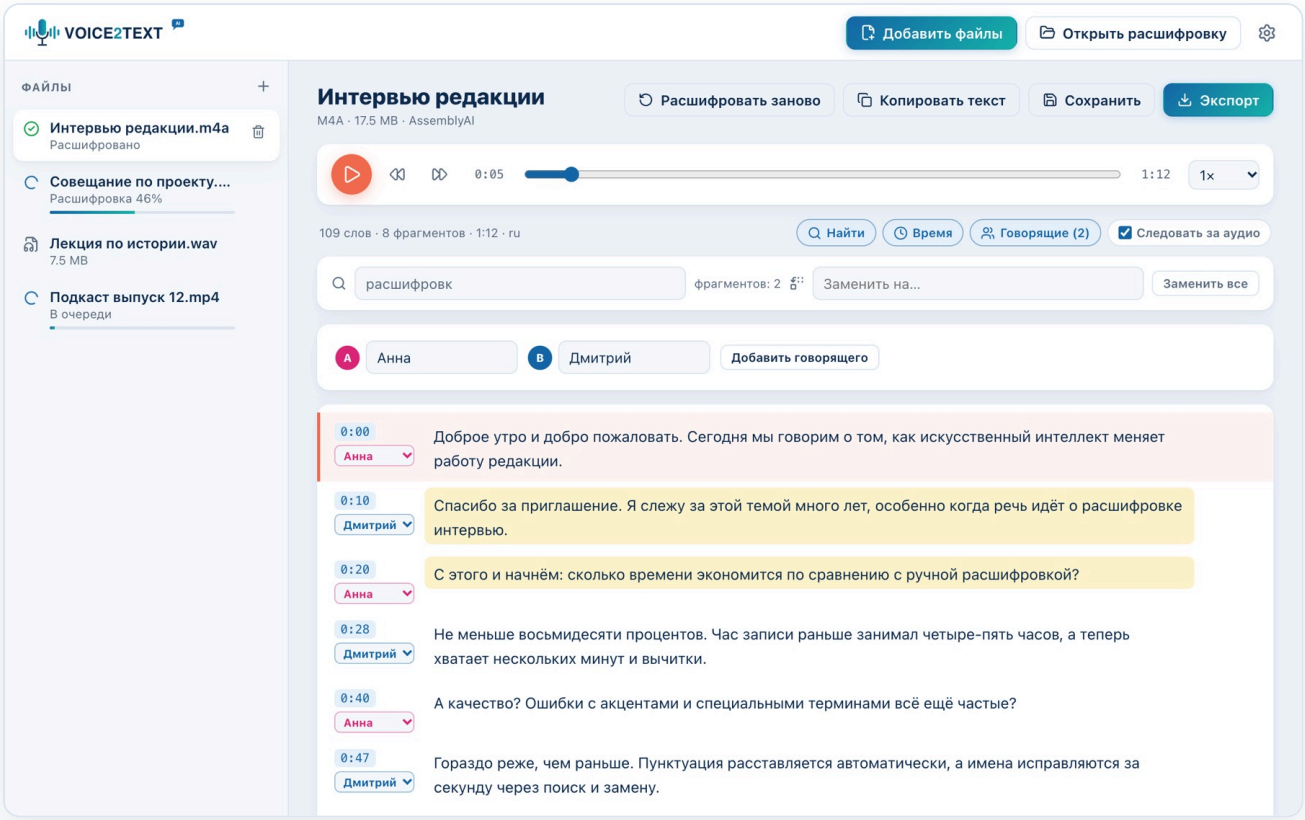


Рисунок 7. Поиск и замена, имена говорящих.

Поиск и замена

Найти (⌘F / Ctrl+F) выделяет фразы с искомым словом; **Заменить все** разом исправляет неверно записанное имя во всей расшифровке.

Говорящие

С AssemblyAI и Deepgram у каждой фразы есть метка (Говорящий A, Говорящий B...). На панели **Говорящие** впишите настоящее имя: оно появится везде, в том числе в экспортированных файлах. Говорящего фразы можно сменить в меню рядом со временем, добавить новых или назначить их вручную, даже если движок их не различил.

08 Сохранение и экспорт

> Сохраните проект, чтобы продолжить правку позже, и экспортируйте текст в нужном формате.

Проект .v2t

Сохранить (`⌘S` / `Ctrl+S`) создаёт файл `.v2t` с расшифровкой, правками и именами говорящих. Открыть его снова можно через **Открыть расшифровку** (`⇧⌘O` / `Ctrl+Shift+O`) или перетащив в окно. `⇧⌘S` сохраняет под другим именем.

Экспорт

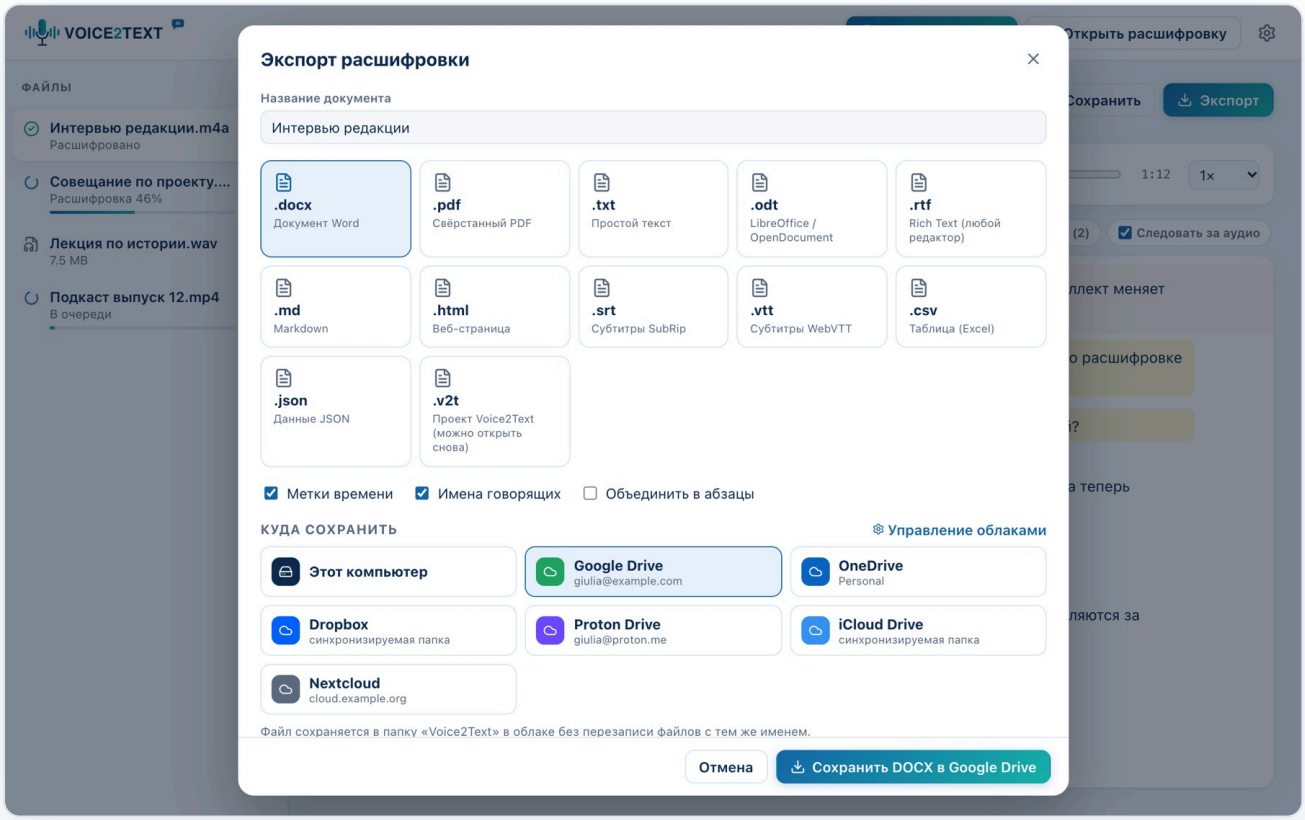


Рисунок 8. Окно экспорта: формат, параметры и место сохранения.

Нажмите **Экспорт** (`⌘E` / `Ctrl+E`), введите название документа и выберите формат:

Формат	Для чего
.docx Word · **.odt** LibreOffice · **.rtf**	редактируемые документы с заголовком и сведениями о записи
.pdf	документы для печати или отправки; создаётся через системное окно печати («Сохранить как PDF»)
.txt · **.md** · **.html**	простой текст, Markdown, веб-страница
.srt · **.vtt**	субтитры для видео (YouTube, Premiere, DaVinci, VLC…)
.csv · **.json**	таблицы (Excel, Numbers) и данные для других программ
.v2t	проект, который можно открыть снова, в том числе из облака

Параметры определяют, включать ли **метки времени** и **имена говорящих** и **объединять ли фразы в абзацы**. Арабский и другие языки с письмом справа налево верстаются правильно.

09 Сохранение в облако

> Voice2Text сохраняет файлы прямо в ваше облако через папки, которые приложения сервисов создают на компьютере. Не нужно подключать аккаунты или вводить пароли.

Как это работает

После установки приложения Google Drive, OneDrive, Dropbox, Box, **Proton Drive**, iCloud Drive, pCloud, MEGA, Nextcloud или Sync.com на компьютере появляется синхронизируемая папка. Voice2Text находит её сам: в окне **Экспорт**, в разделе **Куда сохранить**, выберите облако и нажмите кнопку. Файл записывается во вложенную папку `Voice2Text`, и приложение сервиса сразу загружает его.

- Файлы никогда не перезаписываются: если файл с таким именем уже есть, новый получит имя «имя (2)».
- В облако можно сохранить любой формат, включая проект `.v2t`.
- PDF создаётся через окно печати: выберите там «Сохранить как PDF» и папку облака.

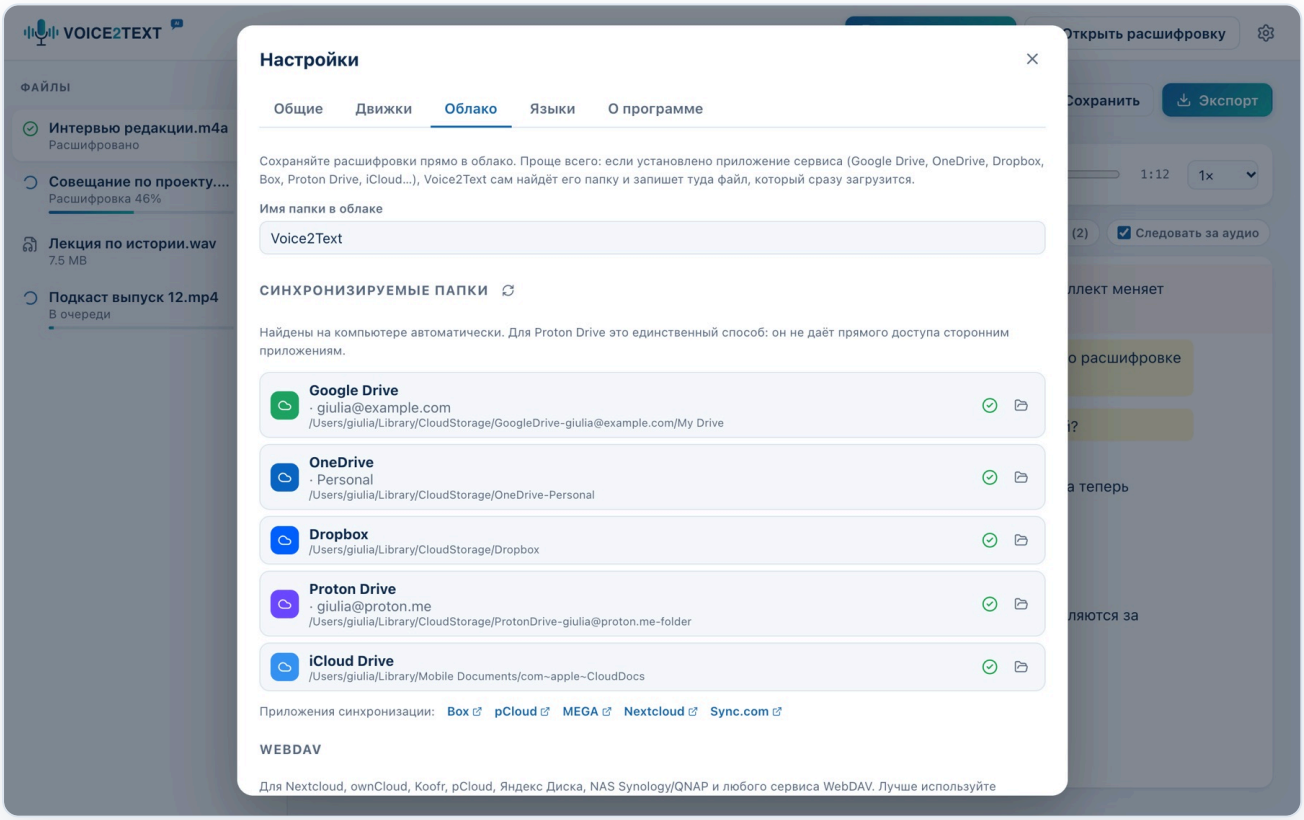


Рисунок 9. Настройки → Облако: найденные на компьютере папки и подключение WebDAV.

Настройки → Облако

- **Имя папки**: вложенная папка в облаке (по умолчанию `Voice2Text`).
- **Синхронизируемые папки**: найденные облака; значок папки открывает её. Если сервиса нет, ссылки ниже ведут к приложениям для установки.
- **Добавить другую папку**: для других облаков или сетевого диска (NAS).

- **WebDAV**: для Nextcloud, ownCloud, Koofr, pCloud, Яндекс Диска и NAS Synology/QNAP без установленного приложения. Выберите сервис, введите адрес, пользователя и пароль приложения, затем нажмите **Проверить подключение**.

> **Proton Drive** не даёт другим приложениям прямого доступа к аккаунту, поэтому Voice2Text использует его синхронизируемую папку со сквозным шифрованием.

10 Настройки и языки

Шестерёнка справа вверху (`⌘,` / `Ctrl+,`) открывает настройки.

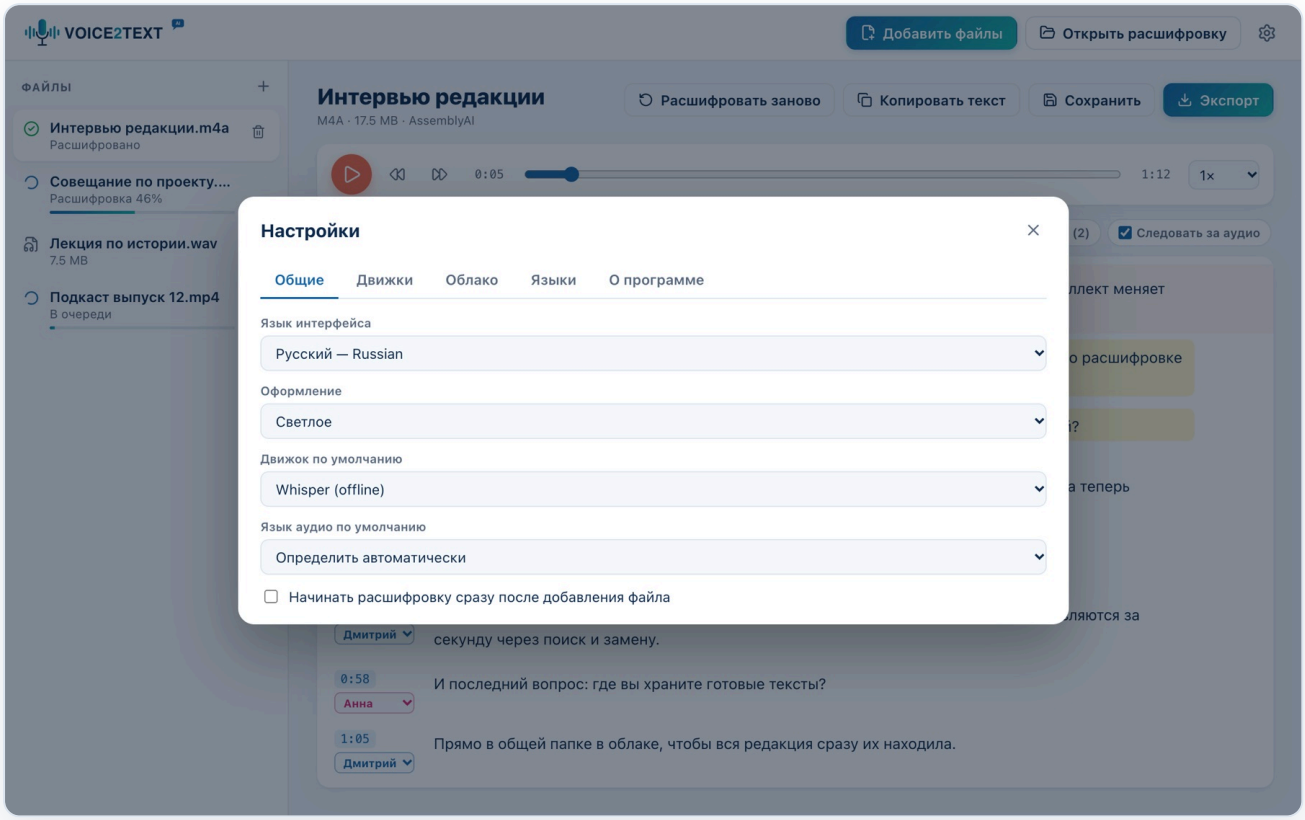


Рисунок 10. Настройки → Общие.

- **Общие**: язык интерфейса, светлое, тёмное или автоматическое оформление, движок и язык аудио по умолчанию, автозапуск расшифровки.
- **Движки**: модель офлайн-Whisper (крошечная, базовая, малая, большая turbo), место, занятое моделями, ключи сервисов, локальный сервер Whisper, наличие ffmpeg.
- **Облако**: синхронизируемые папки и WebDAV (предыдущая глава).
- **Языки**: файлы перевода интерфейса.
- **О программе**: версия, лицензия, предупреждение об онлайн-сервисах.

Языки — это XML-файлы

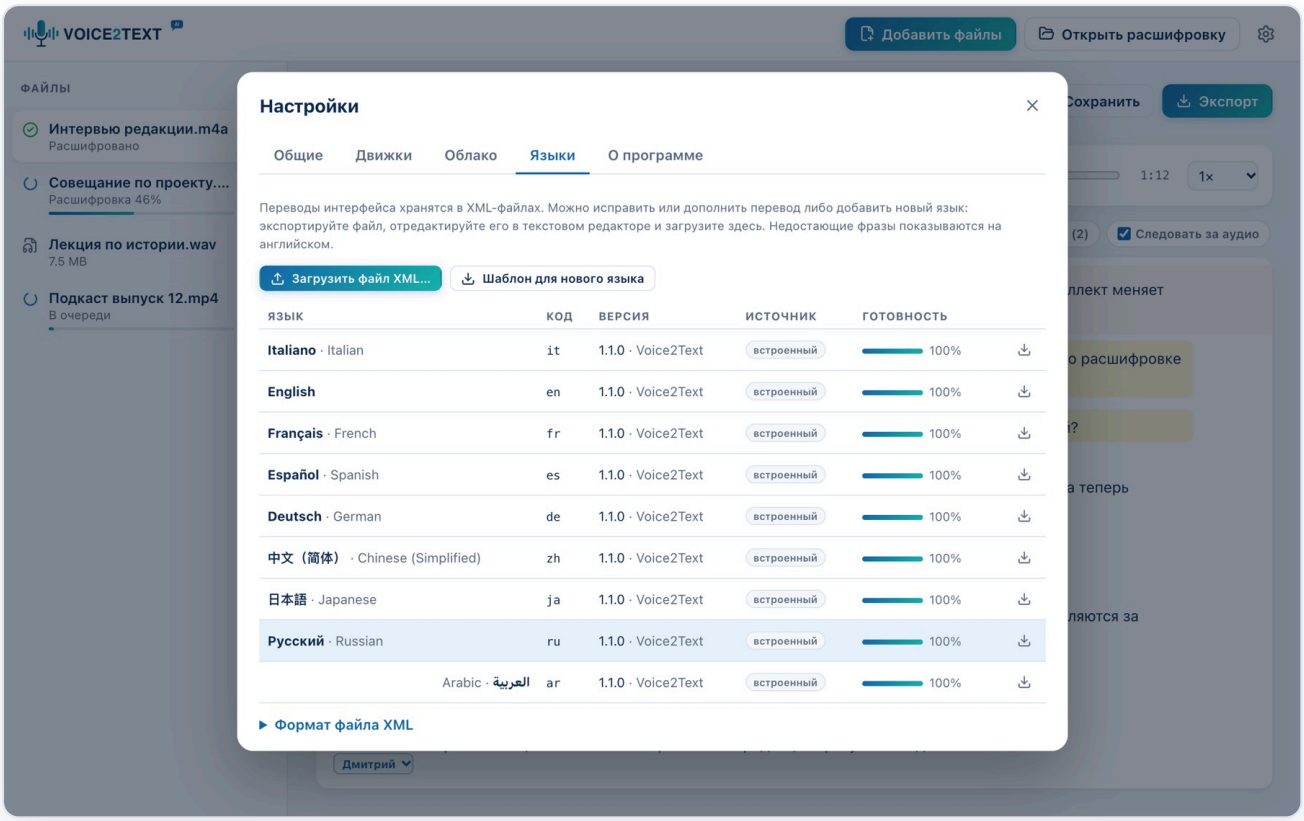


Рисунок 11. Настройки → Языки: доступные языки, готовность перевода и инструменты для XML-файлов.

Каждый язык интерфейса — это XML-файл. Чтобы исправить перевод или добавить новый:

1. Экспортируйте файл языка (стрелка в строке) или скачайте **Шаблон для нового языка**.
2. Откройте его в текстовом редакторе и переводите только текст между тегами, не меняя атрибуты `key` и заполнители в фигурных скобках, например `{n}`.
3. Нажмите **Загрузить файл XML...**: язык сразу появится в списке и в меню языков.

```
<language code="pt" name="Português" english="Portuguese" dir="ltr" version="1.0">
  <string key="app.addFiles">Adicionar ficheiros</string>
  <string key="progress.piece">parte {i} de {n}</string>
</language>
```

Файл с тем же кодом, что и встроенный язык, заменяет его (для исправлений); к оригиналу всегда можно вернуться. Недостающие фразы показываются на английском; столбец **Готовность** показывает долю переведённого. Для языков с письмом справа налево укажите `dir="rtl"`.

11 Сочетания клавиш

> В Windows и Linux используйте **Ctrl** вместо ⌘.

Клавиши	Действие
⌘O	Добавить аудио- или видеофайлы
⇧⌘O	Открыть сохранённую расшифровку (.v2t)
⌘S · ⇧⌘S	Сохранить · Сохранить как
⌘E	Экспорт
⌘F	Поиск и замена
⌘,	Настройки
Пробел	Воспроизведение / пауза (вне текста)
Enter · **Shift+Enter**	Разделить фразу · Перенос строки во фразе
Backspace в начале фразы	Объединить с предыдущей фразой
Esc	Выйти из текста · закрыть открытое окно

12 Частые вопросы

> Ответы на самые распространённые вопросы.

Формат моего файла не читается.
Установите ffmpeg (глава «Установка») или используйте AssemblyAI либо Deepgram — они принимают любой формат, потому что получают исходный файл.

Офлайн-Whisper работает медленно.
Выберите модель **Базовая** или **Крошечная** в разделе Настройки → Движки. «Большая turbo» самая точная, но без видеокарты работает долго. Для длинных файлов Groq гораздо быстрее.

Текст на другом языке или переведён.
Укажите язык записи вместо автоматического определения и нажмите **Расшифровать заново**.

«Неверный ключ сервиса».
Проверьте ключ в разделе Настройки → Движки: скопируйте его заново с сайта сервиса, без пробелов.

Моё облако не отображается.
Убедитесь, что приложение сервиса установлено и синхронизируется, затем нажмите **Искать снова** в разделе Настройки → Облако. Или добавьте папку вручную.

Плеер пишет «аудио недоступно».
Исходный аудиофайл был перемещён или переименован после сохранения проекта. Текст остаётся доступным; чтобы снова слушать, верните файл на место.

Где хранятся мои ключи и настройки?
Только на вашем компьютере, в файле настроек приложения, доступном лишь вашему пользователю. Они никогда не отправляются в Voice2Text.



VOICE2TEXT

От голоса к тексту — на любом языке.

Voice2Text — программа overminder. Лицензия PolyForm Strict 1.0.0: некоммерческое использование разрешено; распространение, изменение и производные работы запрещены.

© 2026 overminder · Версия 1.1