

Поговори со мною, робот!

тест редакции

Окончание. Начало в №№ 5...9/2021

Арсений Ворошилов

Эта статья – завершающая в цикле материалов о тестировании интернет-ресурсов, обеспечивающих преобразование текста в речь, а проще говоря, озвучивание. Позади пять разных сервисов, ниже приводится результат тестирования сервиса Zvukogram (zvukogram.com). Дальше продолжать тестирование вряд ли есть смысл, поскольку, как показывает опыт, сервисы во многом схожи, а различаются в основном ассортиментом голосов и поддерживаемых языков, немного – интерфейсом, а также дополнительными возможностями и финансовой моделью. Думаю, тех шести вариантов, что были рассмотрены в журнале (включая и Zvukogram), вполне достаточно, чтобы составить представление о спектре предложений и сделать выбор в пользу того или иного сервиса либо продолжить поиски самостоятельно.

Итак, Zvukogram относится к тем сервисам, которые не предоставляют бесконечной возможности бесплатного пользования с ограничением размера одновременно озвучиваемого текста. Вместо этого для старта каждому новому пользователю, зарегистрировавшемуся или анонимному, предоставляется кредит в 5 токенов. Этого достаточно для преобразования текста объемом примерно 5 тыс. символов при использовании обычного голоса. Выбор голоса категории Pro сокращает объем текста в 10 раз.

Но, в общем-то, даже столь небольшого кредита вполне достаточно для того, чтобы протестировать сервис и понять, подходит ли он для решения тех или иных задач.

Начну со стартового интерфейса, куда пользователь попадает, зайдя на ресурс. Интерфейс удобен и вполне функционален. Кроме того, в нем есть дополнительные инструменты, которых не было в предыдущих сервисах. На экране есть традиционное окно для ввода текста, а над ним расположены органы настройки. Есть также ссылка на разные голоса, которые можно сравнить и оценить различия между обычными и уровня Pro. Это удобно, поскольку позволяет провести сравнение, не тратя токены. Там же, в окне сравнения голосов, можно поиграть настройками и проверить, как звучит тот или иной голос на разной скорости, с разной высотой тона. Но что еще интереснее, так это возможность менять эмоциональную окраску озвучивания. В интерфейсе этот параметр называется «Настроение». Действительно, робот эмоционально по-разному читает один и тот же текст, когда «Настроение» установлено в значение «Криминальная хроника» и, например, «Новости спорта». Вариантов «настроений» вполне достаточно.

Немало и поддерживаемых языков – 32. Для многих из них широк спектр мужских и женских голосов. Но не для всех. В частности, для иврита есть всего три голоса, а для казахского – вообще один. Но, думаю, это зависит от спроса. К тому же сервисы такого типа находятся в постоянном развитии, так что при необходимости голоса будут добавлены.

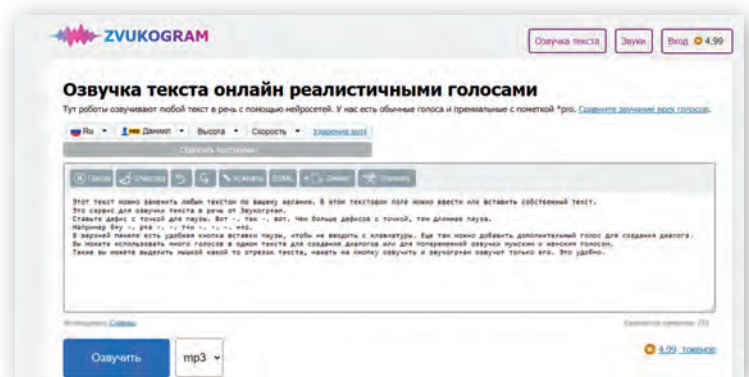
Теперь о настройках, расположенных непосредственно над окном ввода текста. Здесь можно выбрать язык, голос, задать высоту тона и скорость чтения. Тут же находится ссылка на подсказку о том, как работает SSML и как использовать МФА – международный фонетический алфавит. При использовании этой технологии привычный по предыдущим сервисам значок «+» для постановки ударения здесь не работает – придется освоить более сложные инструменты. Но это отдельная тема, требующая отдельной статьи. Тем же, кто не хочет осваивать SSML, остается старый добрый «+», поддерживаемый большинством голосов данного сервиса.

В верхней части текстового окна находятся дополнительные инструменты, позволяющие вставлять в текст паузы, обрезать аудиотрек, формировать диалоги, усиливать те или иные участки текста, использовать SSML. Удобно, что здесь же есть кнопки Undo/Redo. Ну а под окном находится закладка выбора формата формируемого звукового файла. Здесь три варианта – mp3, wav и ogg.

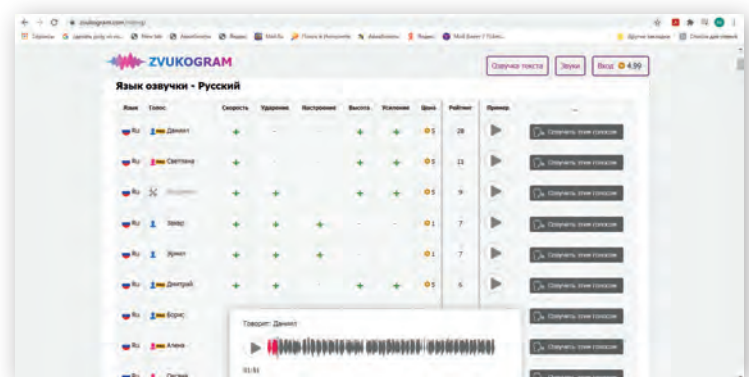
Голоса Zvukogram звучат вполне естественно, даже обычные, а не Pro. Ну а версии Pro, конечно же, по качеству заметно выше. Потому и

цена на них в 10 раз больше. Однако тарифы здесь тоже доступные, так что если требуется быстро и с неплохим качеством озвучить текст, не привлекая профессионального диктора, то этот сервис окажется не только полезным, но и экономически обоснованным.

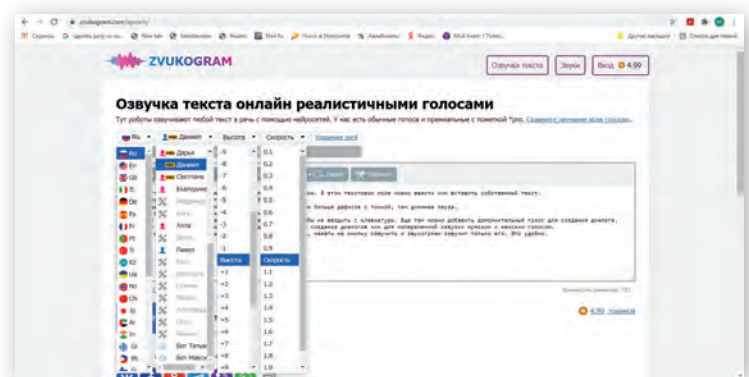
В завершение цикла хочу сказать, что несмотря на богатство и определенное разнообразие сервисов преобразования текста в речь подобные сервисы пока не могут сравниться с профессионалами, обладающими хорошо поставленной речью и обширным опытом озвучивания медиаконтента. Роботы неплохо подходят для озвучивания нейтральных текстов, чаще специализированных, не несущих почти никакой эмоциональной окраски. Но нет сомнения, что технология будет развиваться, а значит, качество станет расти, а сфера применения – расширяться.



Стартовая страница сервиса Zvukogram



Сравнение звучания разных голосов



Основные настройки озвучивания