



Installation and User Instructions

Sum of Sounds

**Text-to-Speech converter with mathematical formulae
and LaTeX support**

May 2020
Version 1.0



Table of Contents

PREFACE	3
Description of the User	3
Obtaining Documentation and Information	3
Internet	3
Description of the product	4
Purpose of the Product	4
Process Overview	4
Running the software	5
Installing the software from source	5
3.2 Direct access through the web	5
Using the Sum of Sounds Text-to-Speech converter	6
Uploading a file or source LaTeX and listening to the generated audio	6
Changing the settings	9



1 PREFACE

1.1 Description of the User

This user manual is intended for the person using the Sum of Sounds web application on a modern internet browser. The intended user need not have any special skills to use this product except that they be able to navigate and use an internet browser. The application is free to use for anyone that has access to the service once it is hosted online. The specific users this application is targeted at are those in the scientific community especially blind professors, students, and research scientists that could benefit from having their scientific literature read out by the computer.

1.2 Obtaining Documentation and Information

1.2.1 Internet

The latest version of the user manual along with other development documentation is available at the following address: <https://bilalsiraj.github.io/sumOfSounds/>



2 Description of the product

2.1 Purpose of the Product

The Sum of Sounds Text-to-Speech converter is a tool to get PDF documents, images or source LaTeX read out by a computer generated voice. It is meant to be used by those with visual impairments who do research in mathematics and read scientific articles and books which usually contain a lot of formulas. It has a simple, minimal user interface and does not distract the user with unnecessary functionalities.

2.2 Process Overview

It is quite simple to use the Sum of Sounds Text-to-Speech converter. The user is presented with two options on the main page of the application. They can choose to either upload a file which could be either a PDF document or an image containing text or formulas, or they can enter source LaTeX into a textbox. To upload a file a user simply click the “Browse” button which will open up a file browser through which they can select the file to upload. After they make their selection, they need to press the “Upload” button. The user then needs to select the “Generate audio” button which will send their uploaded file or input LaTeX to be converted. This may take a few seconds. The user will then be presented with a mini audio player which will automatically begin playing the generated voicing out of their uploaded file.

3 Running the software

3.1 Installing the software from source

As of now the product is not ready for deployment so the only way to run the software is to download and run it from source code. Once one has downloaded the source from the project GitHub repository (<https://github.com/bilalsiraj/sumOfSounds>), they can run the program by following the instructions below.

To start the client and development server:

1. Download yarn and npm for your system
Visit this link for the download file and installation instructions:
<https://classic.yarnpkg.com/en/docs/install/>
<https://www.npmjs.com/get-npm>
2. Using your system's shell, navigate into the project root directory named "formula-reader"
3. Run the command: yarn
4. Navigate into the folder named "client"
5. Run the command: yarn
6. Navigate back into the project root directory
7. Run the command: yarn dev

This should launch the client and development server which will open up a new tab on your internet browser if it was already open. If not, launch your internet browser and enter the following URL into your address bar: <http://localhost:3000/app>

You should now see the main page of the web application on your screen.

3.2 Direct access through the web

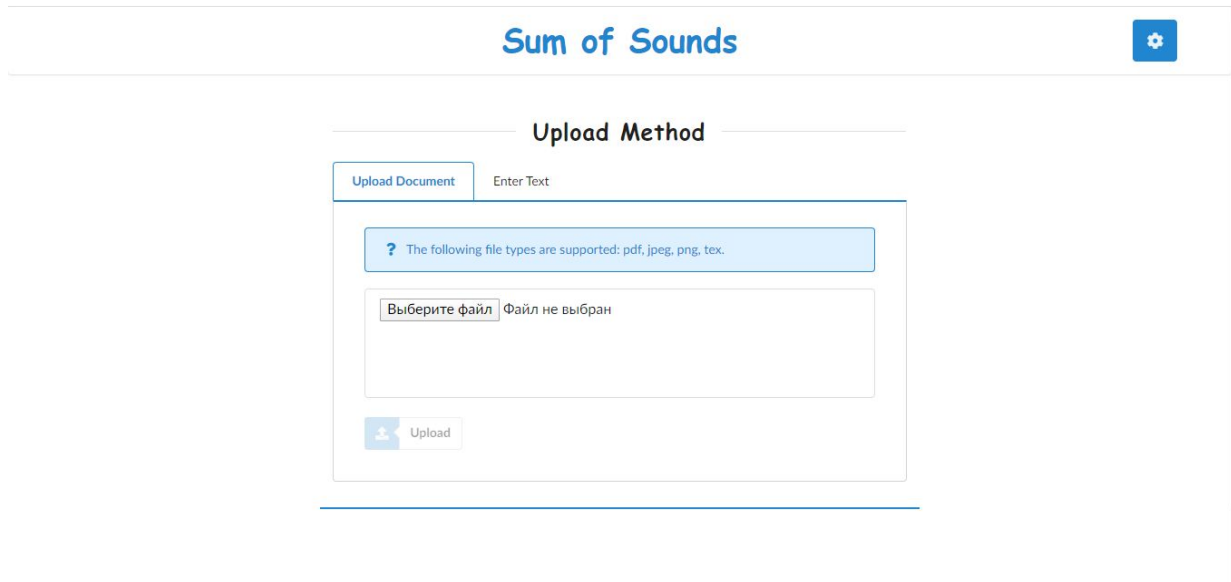
We are currently preparing the software for deployment online. You may visit the following link <https://sum-sounds.herokuapp.com/app> to use the service hassle free. When it is ready to use, you will see a message on the main page, until then it may work on and off so assume it is in development.

4 Using the Sum of Sounds Text-to-Speech converter

4.1 Uploading a file or source LaTeX and listening to the generated audio

The following are basic instructions to navigate the user interface of the web application and use its functionality.

The image below shows the Sum of Sounds main page.



To select a file to upload, simply click the “Browse” button. This will open a file browser through which you can select the file you want to upload. The application accepts files of type PDF, png, jpg, and tex.

Upload Method

Upload Document
Enter Text

? The following file types are supported: pdf, jpeg, png, tex.

Выберите файл mathequation.png

Upload

Play

Generate Audio

Once a file is selected, click the “Upload” button.

Sum of Sounds ⚙

Upload Document
Enter Text

? Copy and paste the LaTeX string to the text bar.

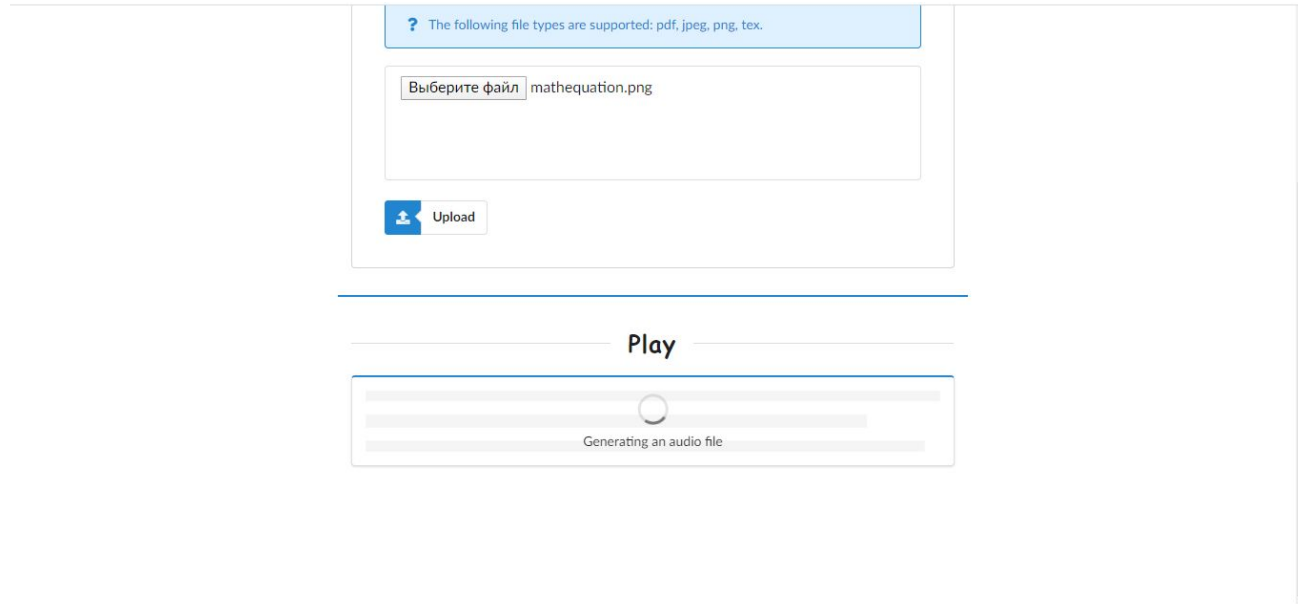
LaTeX string

$\frac{a}{b}$
↻

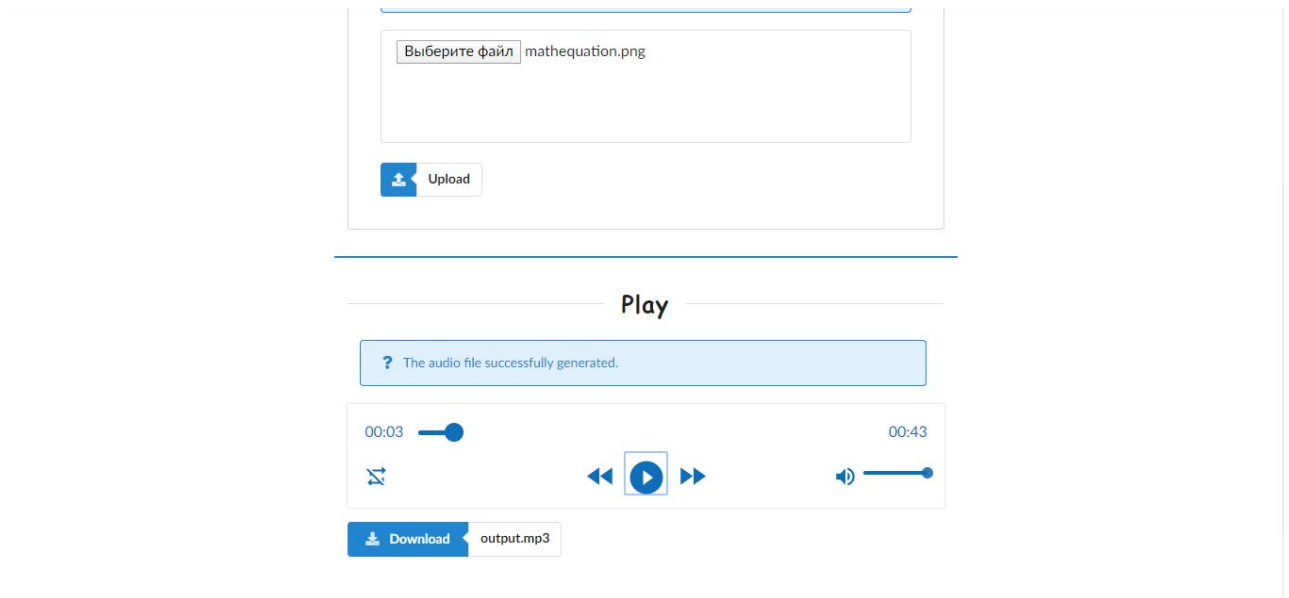
Submit

Alternatively the user may choose to enter LaTeX directly. In which case they must click on the “Enter Text” tab which will present them with a text entry box in which they can enter their LaTeX. After that, they must simply press the “Submit” button.

In both cases, what follows is the same.



Then click on the “Generate Audio” button under the “Play” section. The program will begin generating the audio file. This may take some time.



Once the audio is generated, a small media player will open up. This will allow skipping through parts of the audio, pausing, playing and adjusting volume. In the case that the user wants to save the audio file for later use, they may download it by clicking on the “Download” button. The download will begin automatically.

Upload Method

Upload Document

Enter Text

?

The following file types are supported: pdf, jpeg, png, tex.

Выберите файл

IMG_20200527_184252.jpg

Upload

Play

The application could not read the file. Try it again.

In the case that an unreadable file is uploaded and an attempt to generate the audio is made. The user will receive an error message.

4.2 Changing the settings

Users have the option of changing the generated reading voice to their liking, along with the color theme for the application.

Sum of Sounds

Settings

Audio speed: 0.50

Pitch:

Speaker

en-US-Standard-B

Theme

Black

Cancel

Save

To enter the settings menu, the user needs to click the gear icon which is present on the top right hand corner of the interface at all times. This gives them access to change the speed of the reading voice by moving the slider under “Audio Speed”. By default it is set at natural

speaking pace, 1. But users can change it to a range of values starting at a quarter of the natural speaking pace, 0.25, going all the way up to four times the natural speaking pace , 4.

The user is also presented with the option of choosing from a set of predefined speakers from the “Speaker” dropdown menu, which can be male or female, and have different accents. The reading language is currently limited to English.

By clicking on the “Theme” dropdown menu, users can also choose to change the color scheme of the application to suit their liking.