

Getting Started - Timed Text Speech

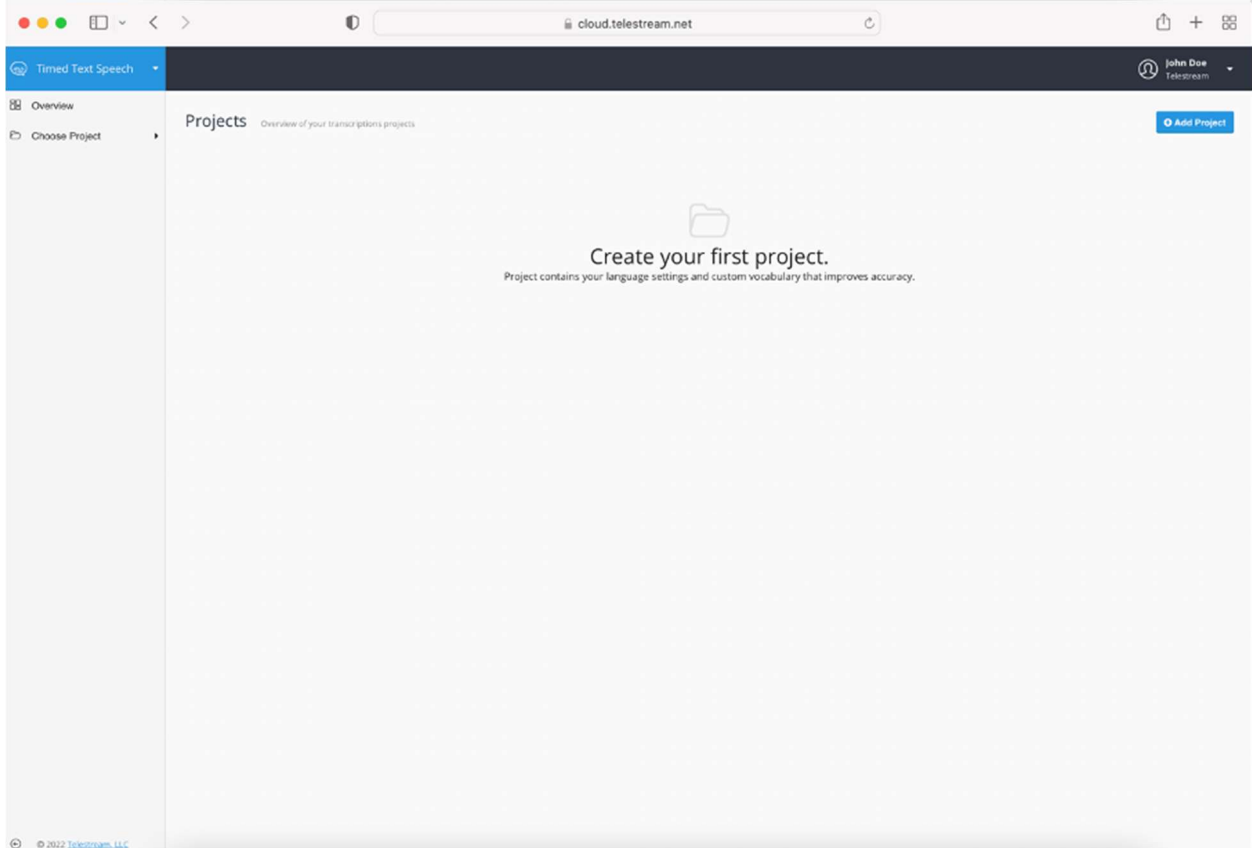
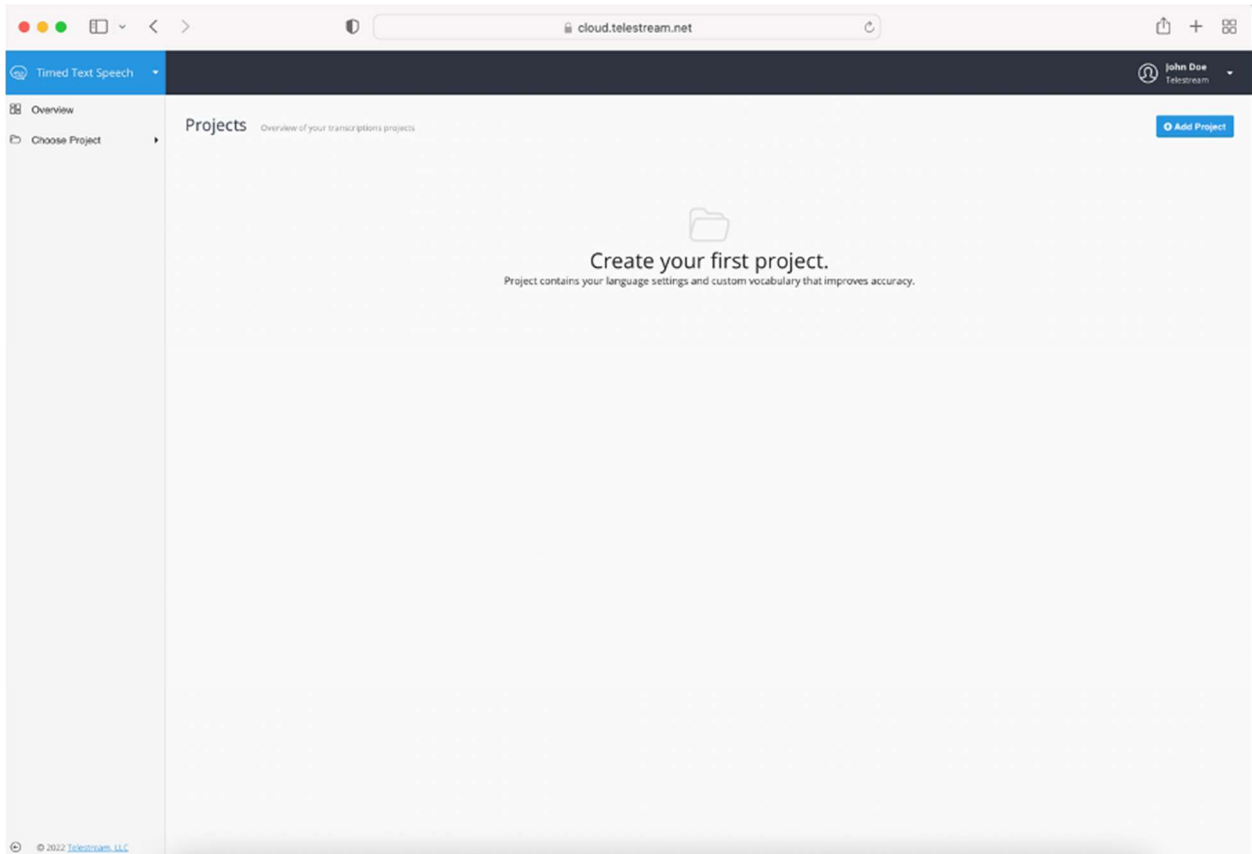
Creating a Project

Everything in Timed Text Speech service happens within a Project. This is where you define:

- Which language which will be used for transcription (base transcription model)
- Custom Vocabulary (optional) which helps to improve results by adding names, acronyms, places, specialized terms, etc.

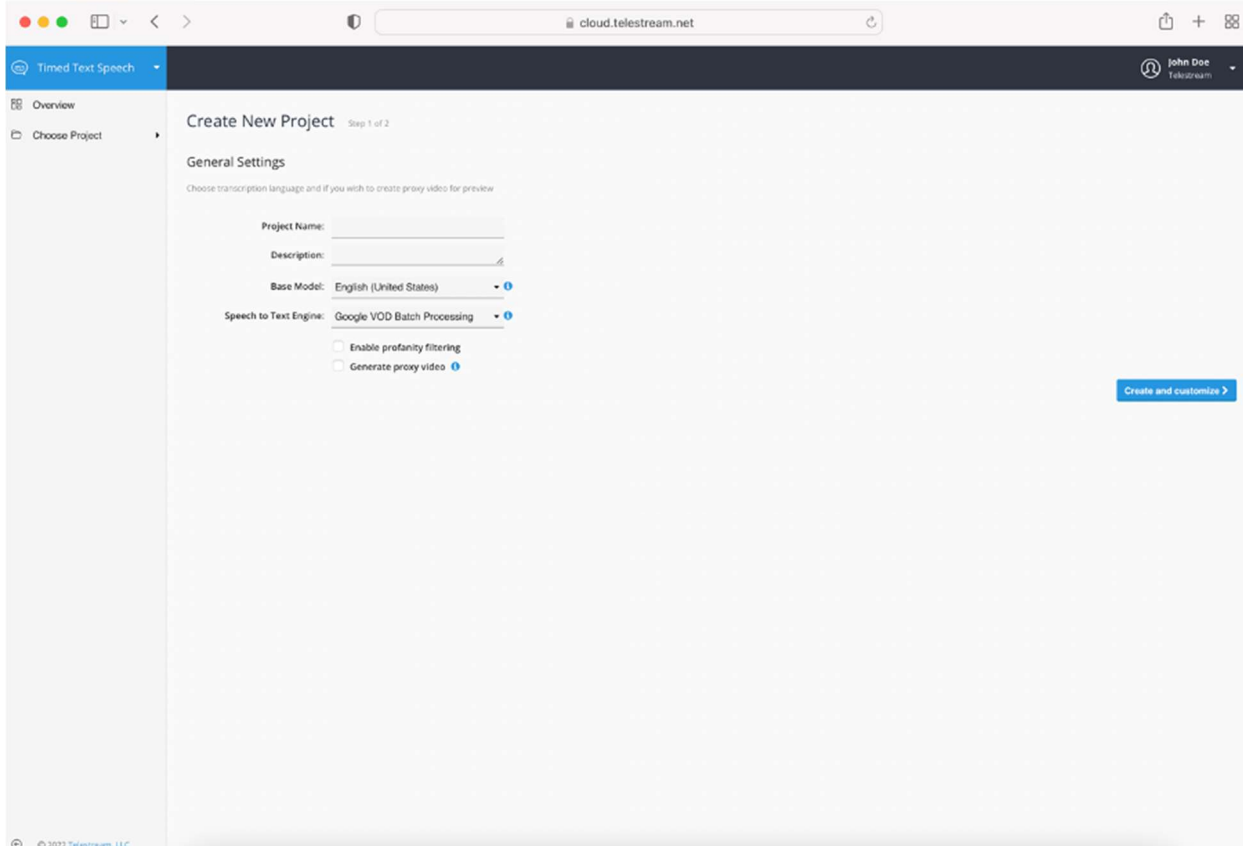
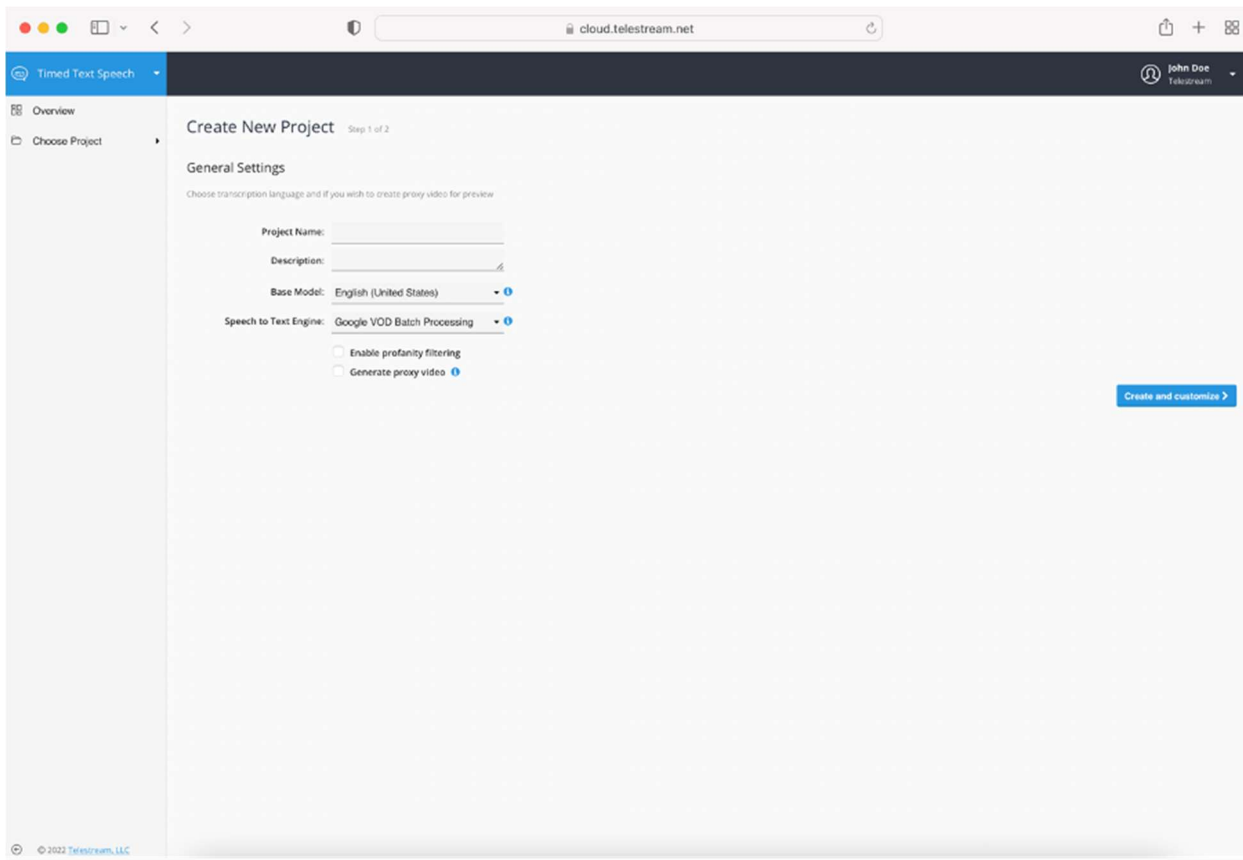
Transcription jobs are submitted to a Project for processing and once finished they are ready for review and editing if needed.

Setting up a Project is a pretty straightforward process, we've made it as easy as possible.



When creating a project - apart from naming it - you need to specify the Base Model. Available Speech to Text Engines will be filtered out based on the Base Model selected.

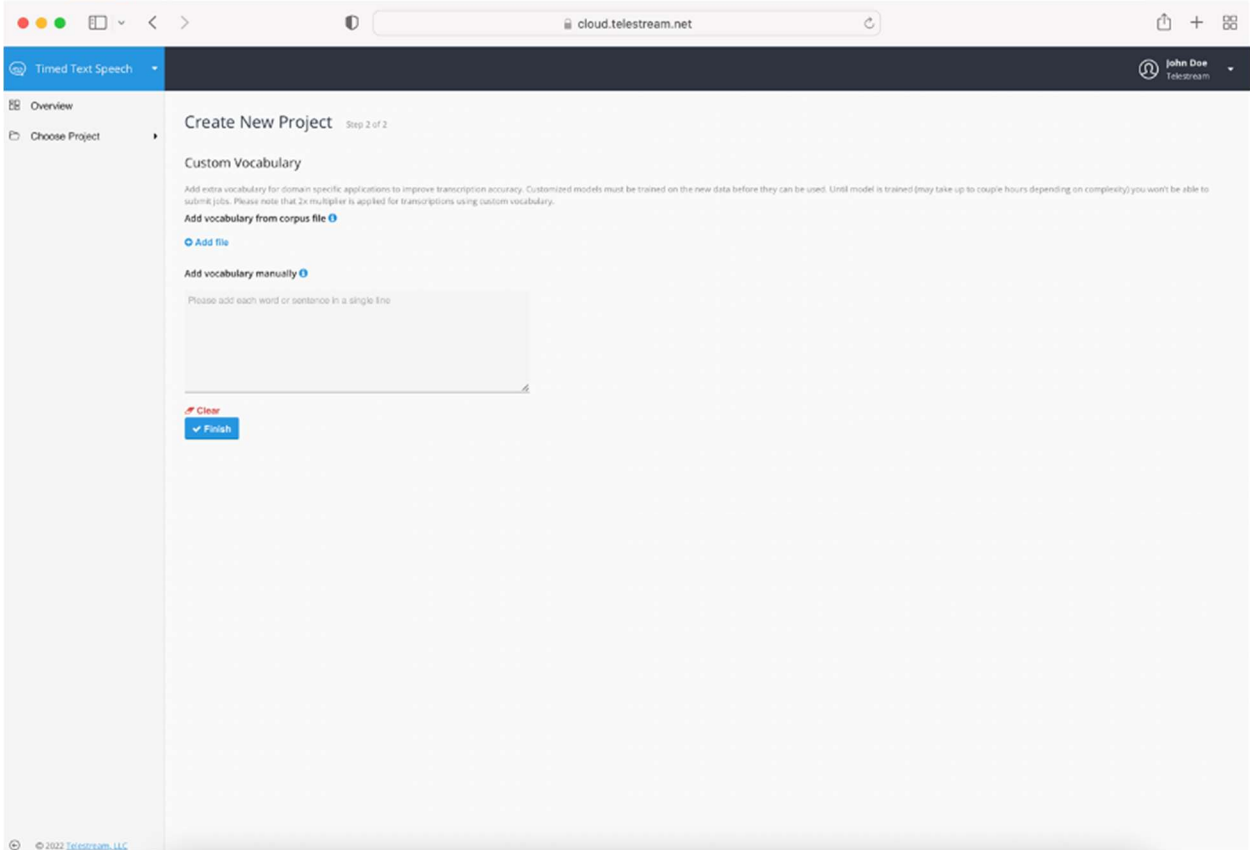
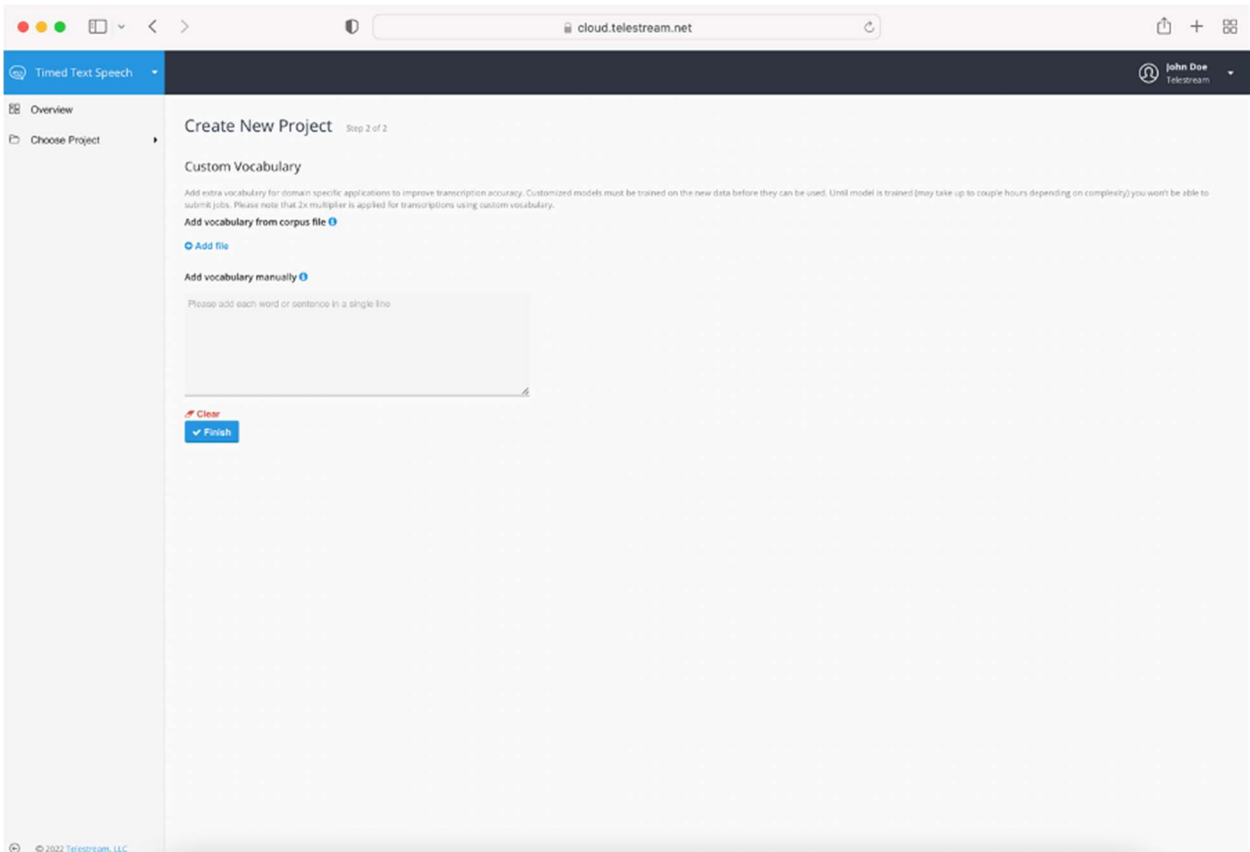
- **Google VOD Batch Processing**
This model will schedule the processing of the file(s) once all audio is completely uploaded. The output will be generated only when the entire file is processed.
- **Google Near Real Time Processing**
This model will process the file(s) continuously as data is delivered. The transcription results appears near real-time as the audio is being processed.
- **IBM Watson**
This model is a data analytics processor that uses natural language processing (NLP) and analyzes human speech. This option offers only near real-time processing, where results appear as the audio is being processed.



Optionally you may add custom vocabulary that is specific to a certain domain - law, sports, medicine, technology, and so on. It can be done in two ways, either by uploading corpus files or by adding words manually. Speech Text Engine which corresponds with the specific Base Model determines the behavior for custom vocabulary.

The table below illustrates the possible options which Timed Text Speech offers when applying the Custom Vocabulary feature:

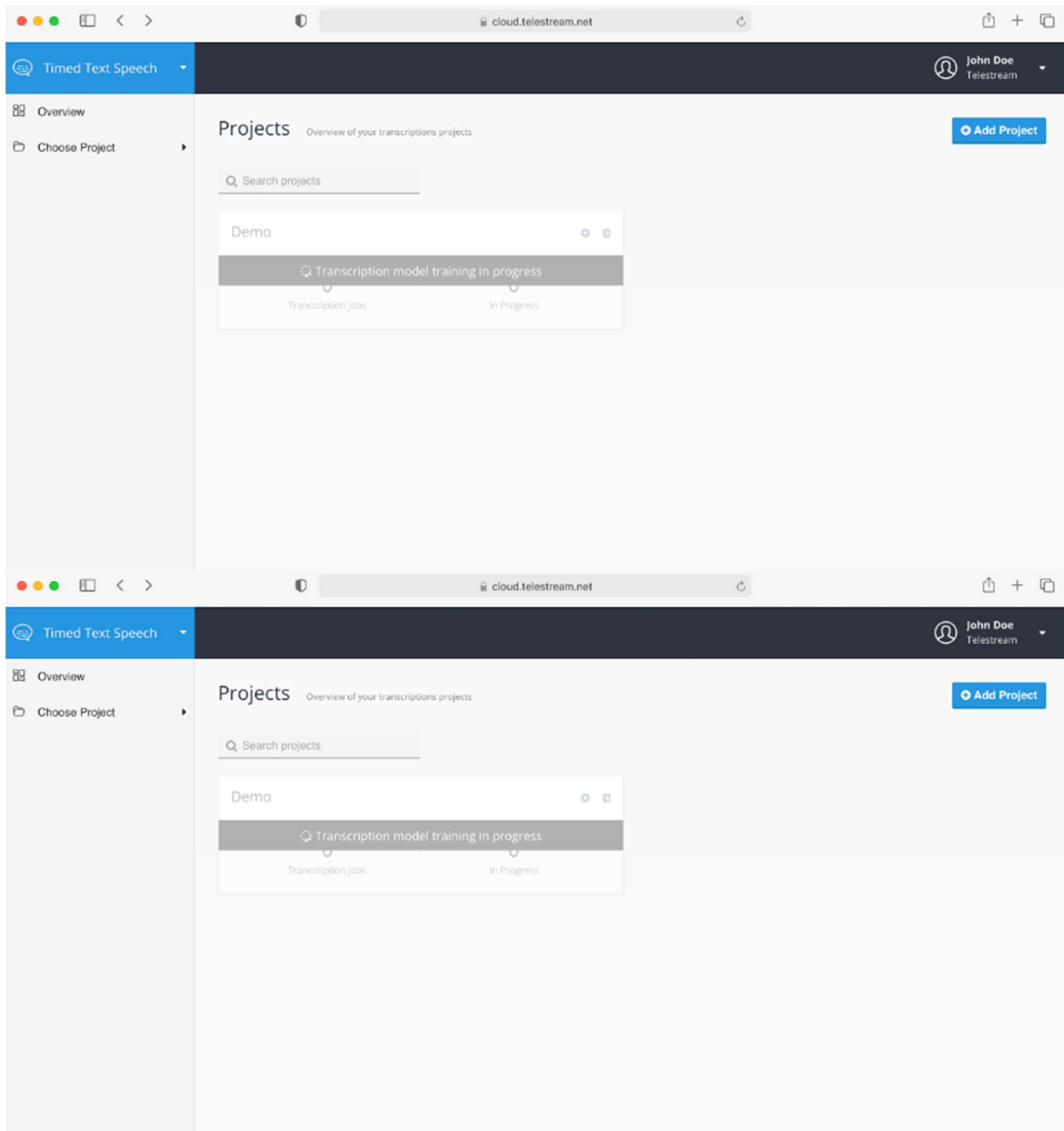
Speech Text Engine	Custom Vocabulary
IBM Watson: Arabic (Modern Standard), Chinese (Mandarin),	Not supported
IBM Watson and Google	Can be specified manually and via corpus file
Google VOD Batch Processing	Can be specified manually
Google Near Real Time Processing	Can be specified manually



A corpus file is a plain UTF-8 text file that ideally contains sample sentences with domain-specific words used in the context they usually appear in, to improve recognition accuracy. If you already have or plan to create corpus files please make sure each sentence is in a separate line and uses consistent capitalization.

You can also add words manually - again using separate lines for each word or sentence is essential. Custom vocabulary can be added or updated later in the Projects Settings.

If you added custom vocabulary then the speech recognition engine needs to be trained on using it properly. This may take some time and until then you won't be able to submit jobs to the Project.



To make sure transcriptions represent high fidelity audio 16,000 Hz sample rate as default is used.