

HTTP/TTS Protocol

[Home](#) >

Description

Our VoiceXML browser can connect to a TTS engine using HTTP. The HTTP protocol is used to transform the prompt text to an audio file. The audio file can be store in a cache directory in order to optimize the TTS ressources using. The first access request to generate the audio file, and save it into the cache. The next times, if you use the same text content, the VoiceXML browser will directly use the file in cache, as a prerecorded audio.

This protocol is simple :

- From the VoiceXML browser, you configure to use HTTP, a (POST recommended) request containing mainly the text content and additional parameters (like language, voice...).
- The server treats your request.
- The VoiceXML browser receives an audio file (cpataible with Asterisk) : it store its in cache and play it.
- If you try to use the same content, the VoiceXML will check and use the cache content instead of requesting the TTS engine.

VoiceXML Browser configuration

The main TTS configuration is set in /etc/vxmld.conf, in the section "TTS server configuration", parameters prefix "client.prompt.resource.0."

- **method** : When you set the 'method' with POST or GET the HTTP/TTS protocol is used to process <prompt> text contents. If you set the value ASTERISK, the VoiceXML browser will send the content to the Asterisk module (as a text/UTF8 or XML/UTF8 content).
- **urivideo** : same as uri but when you sent the xml:language="video".
- **uri** : You need to set the 'uri' for the TTS (or TextToVideo) service (our scripts install the services in <http://ip/tts/provider/tts.php>).
- **format** : Configure the audio file 'format' used, all the scripts not support all the format. Have a look on the install documentation to check and set the correct format.
- **formatvideo** : same as format but when you sent the xml:language="video".
- **maxage** : The parameter 'maxage' force to refresh the cache after sometime. The value 0, disable the caching, the HTTP request will be use for each prompt. The value -1 define infinite age. If the file exist in the cache, it will be always used from the cache directly.
- **cutPrompt** : The option 'cutPrompt', allows to slice the prompt in order to mutualize the maximal contents (cuts in '!', ',', ':' ...).
- **ssml** : The option 'ssml' for to send the text as a SSML/XML well formatted content (with <?xml> and <ssml> roots tags).

Configuration example :

```
#####
# TTS server configuration #
#####

#client.prompt.resource.0.uri                VXIString
http://localhost/tts/flite/tts.php
#client.prompt.resource.0.uriVideo           VXIString
http://localhost/tts/video/ttv.php
client.prompt.resource.0.method              VXIString    POST
client.prompt.resource.0.cacheDir            VXIString    /tmp/cacheContent
client.prompt.resource.0.format              VXIString    wav
client.prompt.resource.0.formatVideo         VXIString    3gp
client.prompt.resource.0.maxage              VXIInteger   -1
client.prompt.resource.0.checkBreak          VXIInteger   0
client.prompt.resource.0.cutPrompt           VXIInteger   0
#client.prompt.resource.0.api                VXIString    voxygen
#client.prompt.resource.0.user               VXIString    demo
#client.prompt.resource.0.password           VXIString    demo
client.prompt.resource.0.ssmI                VXIInteger   0
```

HTTP Parameters

- **text** : the text to synthesize : from the <prompt> content (UTF8 format).
- **language** : the language used (en-GB, fr-FR...) : from the xml:lang attribut.
- **format** : the audio format to return (wav, gsm, mp4... formats supported by Asterisk) : from the configuration.
- **voice** : the voice (Carla, Marcos... depends on the TTS provider) : from the xml:lang attribut.
- **size*** : the size of the image : from the property promptsize.
- **background*** : the image reference or color used for the background : from the property promptbackground.
- **color*** : the color for the text : from the property promptcolor.
- **font*** : the size of the font : from the property promptfont.
- **offset*** : the offset X shift to the text : from the property promptoffset
- **position*** : the position Y shift to the text : from the property promptposition
- **hmac** : MD5 key generated for Voxygen Cloud integration.

* : Only for TextToVideo function. When you set xml:language="video".

From:
<https://wiki.voximal.com/> - Voximal documentation

Permanent link:
https://wiki.voximal.com/doku.php?id=legacy:text-to-speech_guide:tts_http_protocol:start&rev=1425416165

Last update: 2017/07/28 22:29

