

Audio data is binary data. You can read the binary data directly from a gRPC response; however, JSON is used when responding to a REST request. Because JSON is a text format that does not directly support binary data, Text-to-Speech returns a response string encoded in Base64 (<https://en.wikipedia.org/wiki/Base64>). You must convert the base64-encoded text data from the response to binary before you can play it on a device.

JSON responses from the Text-to-Speech include base64-encoded audio content in the `audioContent` field. For example:

To decode base64 into an audio file:

