

# **AWS POLLY & S3**

By Richard Hoehn – CSCI 7400  
MTSU – April 2023





# Introduction to AWS Polly

- Amazon Polly is a cloud service to convert Text into lifelike Speech
- It supports Neural Text-To-Speech (NTTS) voices, that delivers human like speech and pronunciation!
- Common uses cases are:
  - Games, eLearning, Accessibility (visually impaired)
- **Benefits of AWS Polly**
  - Cost effective – Only pay what you use (SaaS) (~50hours @ \$22.00)
  - Low Latency
  - High Quality synthesis for natural speech – Solid pronunciation
  - Supports many voices, including French Canadian (demo will use VoiceId = “Ruth”)



# Why AWS Polly

- **Text to Speed in the Cloud:**

The ability to have an API (cloud SaaS) service provide text to speech at scale enables users of the platform to build application to reach more users.

- **Access to a Broader and more Diverse Audience:**

Text to speech opens doors for people with disabilities, second-language learners, and older adults struggling with increasingly complicated user interfaces.

- **Better Outcomes for Corporate Learning Programs:**

Everyone learns differently. Some prefer to read content. Others retain more when they hear it. Many do best with bimodal presentation, in which materials are delivered by text and speech simultaneously.



# Why AWS Polly (continued)

- **News Sites:**

News are using Polly (USA Today, WAPO, CBS) to give their readers of the web the ability to hear stories.

- **Telephony:**

Sites like FICO, Best Western Hotels, Twillio, Vonage us AWS Polly to help users with automated customer service calls.

- **Learning:**

Sites like CommonLit, DuoLingo, Roybi Robots use text to speech for their users to understand and learn skills in STEM, Games, Songs, and other creative activities.



# Introduction to AWS S3

Amazon S3 or Amazon Simple Storage Service is a service offered by Amazon Web Services that provides object storage through a web service interface.

- **Availability**

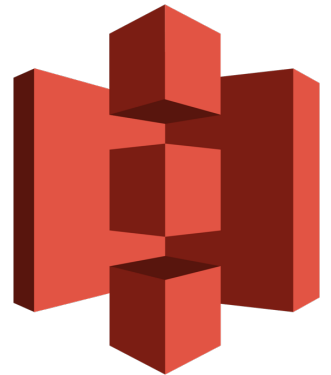
Designed for 99.999999999 percent durability, AWS S3 also provides easy management features to organize data for websites, mobile applications, backup and restore, and many other applications.

- **Flexibility**

S3 is ideal for a wide range of uses like data storage, data backup, software delivery, data archiving, disaster recovery, website hosting, mobile applications, IoT devices, and much more.

- **Cost**

S3 is very low cost and lets you store data in a range of “storage classes.” These classes are based on the frequency and immediacy you require in accessing files.





# Introduction to Ready 2 Learn

- Pre-K – 3<sup>rd</sup> Grade Supplemental Educational App
- Used by students (kids) in Canada and the USA
- Aims to improve Reading, Writing, and Math
  - R2L is a Teaching Platform – not a Testing Platform
- Launch date to be early August 2023 (US & CA)
- Website: <https://www.ready2learn.app>





# R2L Use Case for AWS Polly

## Why Richard chose AWS Polly to Review

- While kids/students are learning from the Ready 2 Learn App, they will at some point get “stuck” in their learning.
- Most app users (kids K - 3<sup>rd</sup> grade) cannot read very well or at all!
- Using Text-To-Speech the Ready 2 Learn App can help kids get “**un-stuck**” on their own by contextual descriptions of the problem.
- This theoretical is just an idea, but the demo can show the uses of SaaS Text-To-Speech options with an AWS Polly demo.
- Round trip test are as low as 200ms for AWS Polly to convert text to 6-8seconds of audio mp3 sound.
- Demo is a simple NodeJS app that uses the AWS SDK to interface with AWS Polly.



# Demo

## Process Flow (see console output on right):

1. Application requests user's name.  
Example user enters: **MTSU CSCI Department**
2. Creates a simple text saying "Hi..." text to the user and reciting the name of the day and how many days are left until the weekend. (*illustrates change*)
3. Sends (API) text to AWS Polly for synthesis that returns a "Readable Stream"
4. Saves mp3 file to disk
5. Uploads mp3 to S3
6. Retrieves size of object (uploaded mp3) and displays on console
7. Plays the AWS Polly mp3 file on user's PC 😊

```
Hi,  
What's your name: MTSU CSCI Department  
  
Sending to AWS Polly:  
>>> Hi MTSU CSCI Department, I hope you are doing well on this beautiful  
Wednesday? Only 2 days until the weekend! <<<  
AWS Round Trip: 263.506ms  
  
Saving (2023-03-22T20:00:31.543Z.mp3) to local disk now...  
Local save complete  
Local Save: 250.559ms  
  
Uploading (2023-03-22T20:00:31.543Z.mp3) to S3 now...  
Uploading to S3 complete  
AWS Round Trip: 471.409ms  
  
Get size of uploaded file from S3 now..  
File (2023-03-22T20:00:31.543Z.mp3) retrieved is 43.92kb  
AWS Round Trip: 40.919ms  
  
Playing returned "mp3" now...  
Sound Play Time: 7.767s  
  
Done with Demo – Have a good day!
```





# Demo (cont.)

## Demo in Terminal



# Conclusion

## Thoughts on AWS Polly in conjunction with the CSCI 7400 Assignment 02

- Polly is very fast – To convert text to speech of 5-7seconds takes **~250ms!**
- Pay as you go – I envision this being a very good way to let developers test speech synthesis and build apps that can help a variety of users.
- Easy to use – Before this assignment I (Richard) had not prior experience with AWS Polly. Getting it setup and running was really easy and took a few days to understand.
- Will seriously think about using AWS Polly on Ready 2 Learn in the future!



# Thank you

Details and source code can be found on GIT  
@

<https://github.com/richardhoehn/mtsu.csci.7400.assignment02>



# References / Citations

- <https://aws.amazon.com/polly/customers/>
- <https://www.readspeaker.com/blog/benefits-of-text-to-speech/>
- <https://nodejs.org/dist/latest-v18.x/docs/api/>
- <https://aws.amazon.com/sdk-for-javascript/>
- <https://www.ready2learn.app>
- <https://www.simplilearn.com/tutorials/aws-tutorial/aws-s3>