Ready for Prime Time?

The State of AI-Based Air Talent

SUMMIT TECHNOLOGY GROUP

Paul Stewart, Summit Technology Group

About Me

President and Founder of Summit Technology Group. My background is largely in broadcast automation and related software. I am a self-taught developer but these days, I rely on my talented development team to build better products than I could ever build on my own.

About Summit Technology Group

Broadcast software development company with products that utilize various cloud architectures, artificial intelligence, and machine learning to provide new and innovative workflows to the broadcast industry.

Background of AI Voices

- **What is an AI Voice?**

  A voice that is synthesized using machine learning to deliver or speak a group of words, typically in an on-demand fashion.

- **What AI Voices are Available?**

  Cloud providers such as AWS, Microsoft, and IBM provide cloud-based voices while dozens of other providers have SAPI5 voices that install locally to a computer. Available in a variety of languages and dialects.

- **Where is this Technology Used?**

  - Call centers and IVRs
  - Assistive technology for cognitive disabilities
  - Dynamic paging such as mass transit
  - Increasingly being used in broadcast

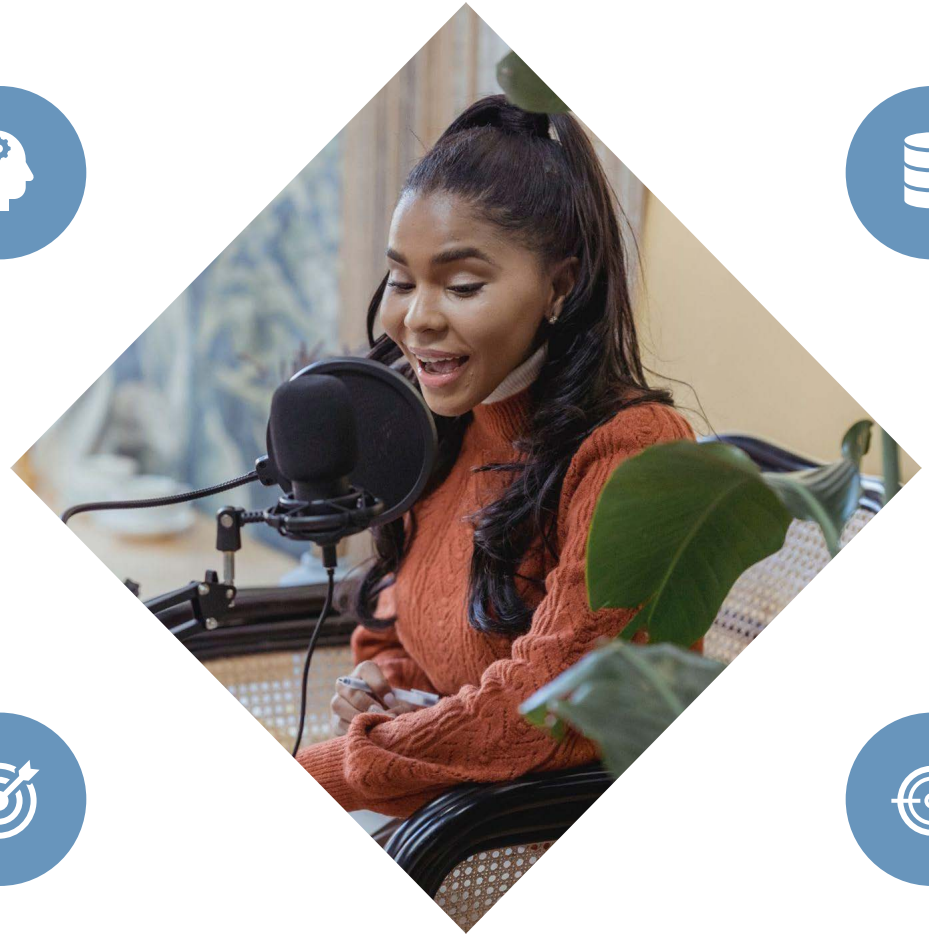
The Catalyst for Change

Broadcast has seen an increase in the adoption of AI-based technologies, including synthesized voices for a variety of reasons.

- Safety measure due to COVID-19 and local lockdown orders
- Reduced staffing due to current shifts in labor market
- Competition with on-demand media and podcasts
- Increased cost of operation in a broadcast facility

Standard vs. Neural

While standard text to speech engines utilize concatenative synthesis, meaning words are synthesized independently and concatenated, neural synthesis uses a continuous spectrogram and vocodes the result for a more natural sound.



Speaking Style and Prosody

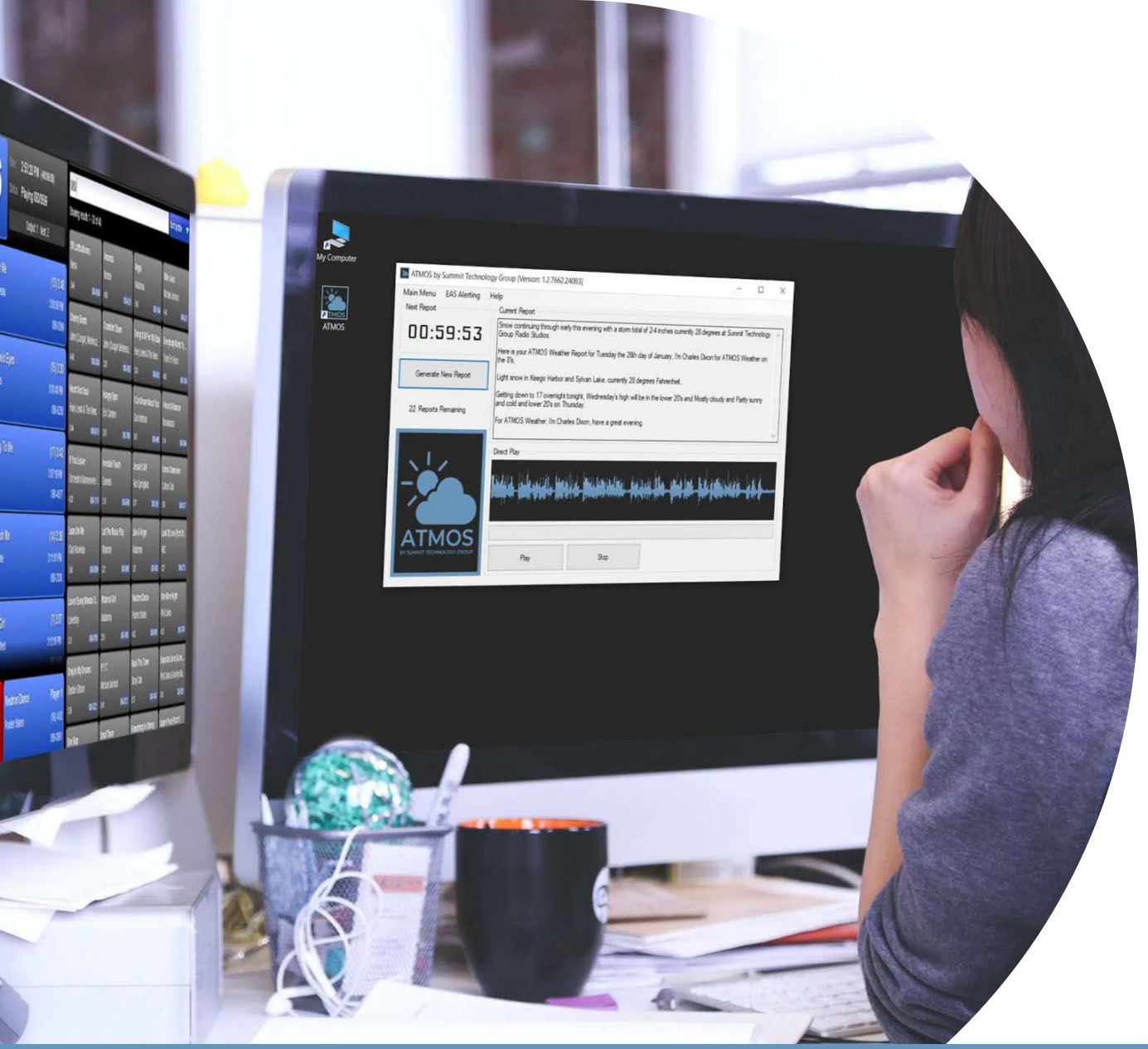
Various speaking styles, including newscast and conversational, are available. These can be further customized with prosody settings include speed, intensity, pauses, breaths, pitch, and timbre.

Gender, Language, Dialect

A voice has multiple characteristics including gender, age, language, and dialect. Surprisingly, many voices are available for nuanced differences like EN-GB and EN-Scottish

Custom Pronunciation

Many AI voice engines allow users to customize pronunciation by using specific tags within the script or by uploading a custom lexicon.



Short Form Reporting

Weather, news, sports, market, and agriculture reports can all be automatically synthesized

Commercial Advertising

Commercial messages and underwriter announcements can be delivered with AI

Liners and Voiceovers

In certain cases where the voice matches the station image, AI voices can be used as a substitute for human operators to deliver liners and other voice work

Example Weather Report

Personalize the Script

Use the script to greet your listener with time-accurate dayparts, dates, and times to establish legitimacy. Use station imaging and branding to retain listener attention and boost brand recognition.

Randomize the Delivery

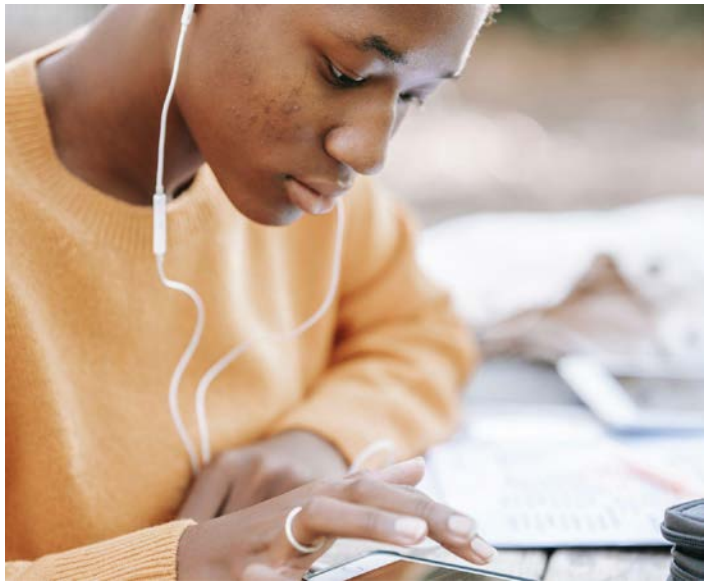
Use generic but distinct phrases to randomize the delivery to keep the content fresh. Try randomizing greetings and pleasantries and experiment with meaningful data such as news and weather.

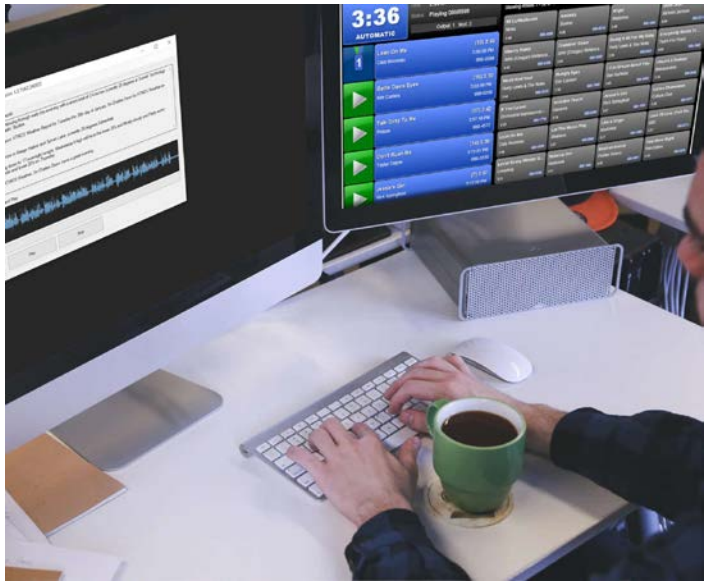
Generalize the Data

Most thermometers can tell us the temperature to the 1/100th of a degree. Your listener wants to know whether to pack an umbrella or a pair of sunglasses.

Use Programmatic Cues

Programmatic cues can help personalize your content based on the listener and what message is being delivered.





Voice Cloning

Voice cloning products are already widely used as a voice replacement for those with ALS or similar conditions, and broadcast-quality voice cloning is quickly becoming a reality.

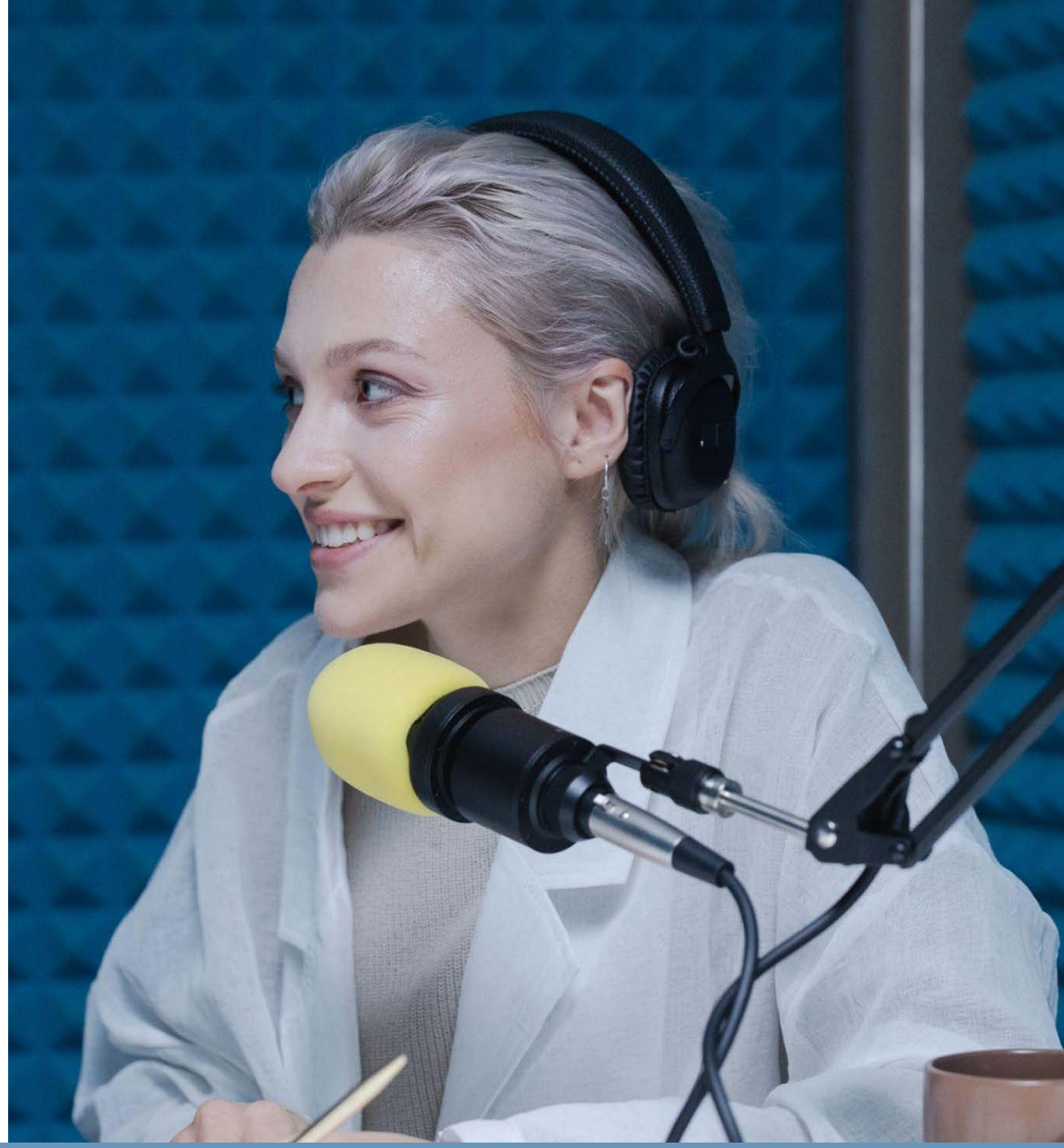
Conversational Commerce

On-demand and OTT outlets can leverage this technology to customize messaging based on location, consumer behavior, and previous buying decisions (including remarketing).

Broadcast Integration

As with all technology adopted by broadcasters, the need for a seamless integration with broadcast automation and newsroom systems is paramount.

Where AI Voices Fall Short



AI is not a replacement for human talent

While AI-based voice engines can be impeccably similar to that of a human, and while the tools that rely on these engines are powerful enough to think and synthesize stopsets on demand, they are still a long way from replacing humans completely. Human talent connects with your listeners over common local knowledge and pop culture as well as evoking emotions AI voices are contextually not good at handling...yet.