

Viele Radiosender haben es erkannt, das enorm hohe Potenzial einer modernen Text-to-Speech-Lösung. Der rbb hat – wie auch der WDR und die BBC – in diesem spannenden Pilotprojekt gelernt, wie leicht und einfach man die Microsoft Cognitive Services auch in die eigenen Produktionsabläufe integrieren kann und so für den Hörer enorme Mehrwerte schafft, ohne, dass es viel kosten muss!



Ausgangslage

Viele Hörerinnen und Hörer in der Region Berlin-Brandenburg nutzen das rbb-Inforadio zur Information, und auch die Smartphone App des rbb Inforadio oder die Website „inforadio.de“ als Informations-Portal. Die beliebtesten Dienste sind dabei natürlich regional: Verkehrsnachrichten und der Wetterbericht. Gerade der letztere wird natürlich rund um die Uhr erstellt, aber über Nacht ist oft gar keine Sprecherin oder kein Sprecher mehr im Studio, um die aktuellen Informationen einzusprechen, damit sie per Audio über die App oder das Web abgerufen werden können.

Lösung

Microsoft und Ceteris haben für den rbb eine neuronale Stimme auf Basis der Microsoft Speech Services ausgewählt, die in überraschender Qualität die Verkehrs- und Wettermeldungen "vorliest". Die Herausforderung für Ceteris war dabei die Anbindung an die Systeme des rbb: die neuen Texte müssen der Redaktionsdatenbank entnommen werden, die fertigen Audio-Dateien werden im Content Management System zurück erwartet. Die Kommunikation erfolgt dabei aktuell meist über SFTP! Besonders wichtig war dem rbb, dass diese Prozesse über ein Web-Portal jederzeit überwacht und beeinflusst werden können, etwa, wenn man blitzschnell eine manuelle Verkehrsmeldung einfügt, oder wenn man den Überblick behalten muss, welche Version welches Meldungstextes denn jetzt schon fertig aufgenommen vorliegt.

Verbesserungen

Mit der heute vorliegenden Lösung sind alle sehr zufrieden, obwohl sie erst einmal kaum mehr als einen Piloten darstellt. Die nächsten Schritte sind daher klar: der rbb wird in Zukunft nicht nur die von Microsoft zur Verfügung gestellten Stimmen "Katja" und "Conrad" verwenden, sondern sich auch eine eigene "Custom Neural Voice" erstellen lassen. Außerdem wird Ceteris die riesige Aussprachen-Datenbank der ARD, die vor unter anderem auch komplizierte Orts- und Straßennamen enthält, zur weiteren Verbesserung der Ausgaben in die Erstellung der Audio-Dateien miteinbeziehen.

