

WDR implementiert Text to Speech System



Der WDR hat ein Text to Speech (TTS) System implementiert, um die Barrierefreiheit der Senderangebote noch weiter zu verbessern. Dieses System wandelt bei wichtigen Ereignissen oder im Katastrophenfall Ad-hoc eingeblendete Textnachrichten oder Ticker Meldungen in Sprache um, die zusätzlich mit dem Sendesignal als Audio-Signale ausgespielt werden. Visuell beeinträchtigte ZuschauerInnen können die Textmeldungen dann auch als Audiosignale empfangen – besonders wichtig im Katastrophenfall. LOGIC media solutions hat das System implementiert und die nahtlose Integration in die Produktionsabläufe des WDR sichergestellt.

Zur Sprachsynthese wird das Allinga Voice TTS-Modul genutzt, eine gemeinsame Entwicklung der Fraunhofer Institute IAIS und IIS. Es besteht aus einem redundanten Docker-Deployment, welches die Allinga Voice TTS und einen von LOGIC programmierten Client beinhaltet, wobei das Allinga Voice TTS Modul beim WDR Köln on-prem angesiedelt ist. Wenn im Sendezentrum ein Signal zur Texterstellung ausgelöst wird und eine Textnachricht am Newsfeed-Server im ARD CN anliegt, fragt der Client diese am Server ab und die Allinga Voice TTS synthetisiert daraus Sprache in Form von Audiosamples. Die Audioserver sind aus Redundanzgründen doppelt vorhanden. Das produzierte Audiosignal wird dann via AES-3 in der Sendezentrale verfügbar gemacht und bei Bedarf ausgespielt. Dabei senkt der dazu genutzte Embedder den eigentlichen Sendeton ab und spielt das Audiosignal der Audio-Server auf den Sendeton auf.

„Wir freuen uns, dass wir mit unserer Allinga Voice TTS Technologie dazu beitragen können, die Soforthilfe und den Informationsfluss in Katastrophenfällen zu unterstützen. Unsere Allinga Sprachassistentenlösungen sind vielfältig einsetzbar und tragen maßgeblich zur Barrierefreiheit bei. In diesem Kontext bieten wir über das TTS-Modul hinaus noch weitere Technologien, wie die Spracherkennung, die gesprochene Sprache in Text umwandelt. So sind auch Einsatzszenarien in anderen Kontexten denkbar – etwa die Transkription von öffentlichen Reden“, sagt Oliver Hellmuth, Abteilungsleiter am Fraunhofer IIS.

Bei der Implementierung des Systems legte LOGIC besonderen Wert auf einen weitestgehend automatisierten Ablauf sowie auf die Schnelligkeit des Zuschaltens, da das System besonders bei Katastrophenfällen zum Einsatz kommen wird. Der Fokus lag hierbei auf der direkten Ansprache der Lösung ohne zusätzliche Arbeitsschritte oder Geräte. Das TTS-System wird über den Broadcast-Controller (KSC) gesteuert und das Starten und Stoppen erfolgt in der Praxis vom Newsroom aus durch den Anwender (Redaktion) ohne einen SAW-Ingenieur zu benötigen.

www.logic.tv