


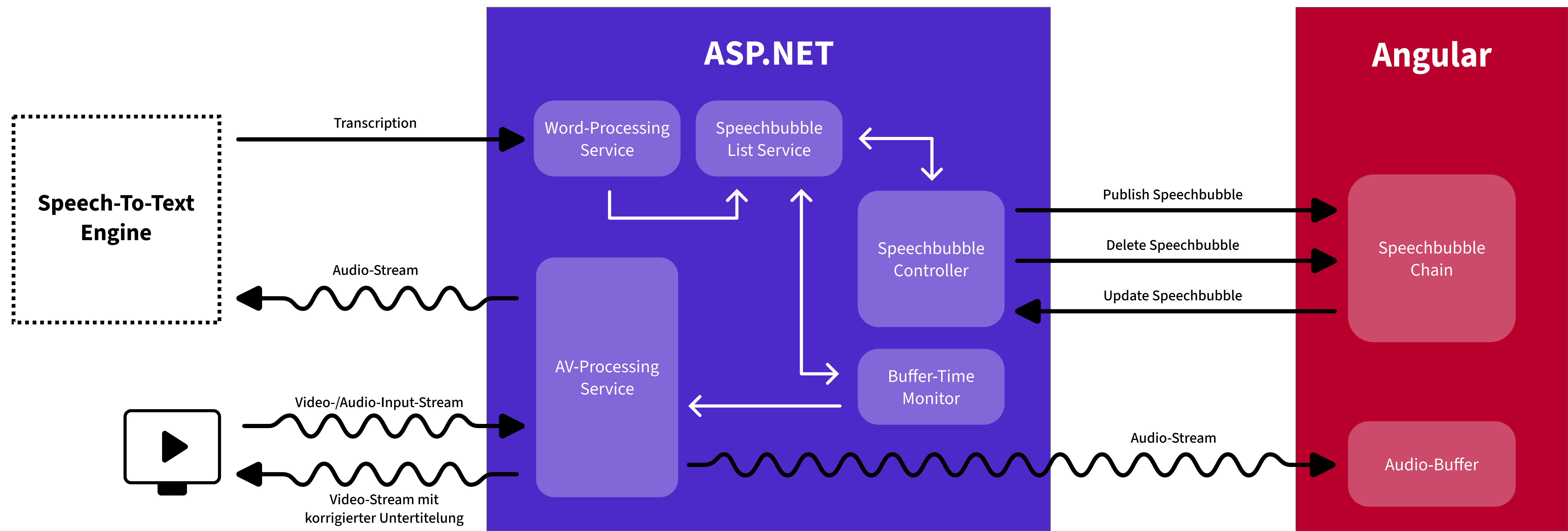
## Live Stream Editor zur Korrektur von Untertiteln

Mit der von unseren Entwicklern erdachten Software, möchten wir ein Werkzeug schaffen, welches das Korrigieren von Live-Untertitelung, so einfach wie möglich macht. Dabei liest unsere Software einen beliebigen Videostream ein, erstellt eine Untertitelung mit einer Speech-Engine, bietet die generierten Untertitel dem Anwender zu Korrektur an und liefert zuletzt den fertig untertitelten Videostream an den Endkunden aus. Und das auf intuitive Weise, in Übereinstimmung mit der DSGVO und unter Berücksichtigung der Barrierefreiheit.

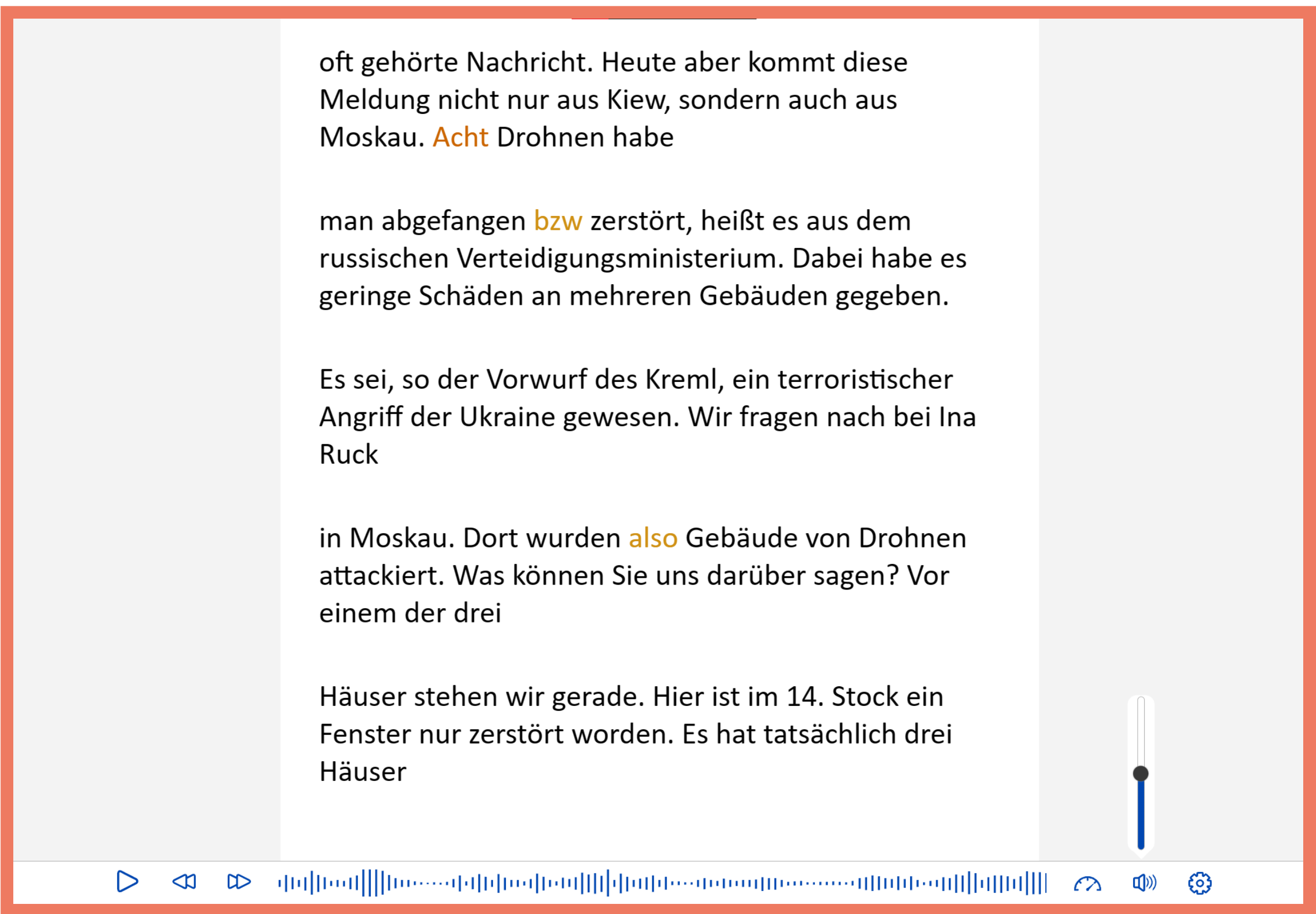


Try it yourself today!

**Zielstellung**  
Live-Untertitelung existiert schon heute, wie die automatische Untertitelung von YouTube zeigt. Diese ist jedoch noch fehleranfällig und eignet sich daher nicht für Live-Streams bei denen ein hoher Wert auf Korrektheit gelegt wird, wie es unter anderem bei Landtagsdebatten der Fall ist. Dabei besitzen diese Technologien das Potenzial, Berufsgruppen wie Stenografen zu entlasten und lässt Menschen die auf Untertitel angewiesen sind, von besserer Untertitelung profitieren. Unser Ziel ist es, bereits heute existierende Software für automatische Untertitelung, um Echtzeit Korrektur zu ergänzen. Dabei sollen die Bedürfnisse von Stenografen mitgedacht und die DSGVO- Konformität eingehalten werden. Somit kann unsere Software überall da eingesetzt werden, wo es auf korrekte Untertitel und Datenschutz ankommt.



**Funktionalitäten**  
Unser MVP erlaubt es dem Nutzer, hereinkommende Transkriptionen, sogenannte Sprechblasen, zu sehen und zu editieren. Jede Sprechblasen beinhaltet gesprochene Wörter, welche in einer Liste zusammengefasst werden. Dies stellt die Übersichtlichkeit der Software sicher. Zudem werden Wörter bei dem sich die Speech-To-Text Engine unsicher ist, farbig markiert. Möglich ist dies dadurch das Wörter mit einem Konfienzwert versehen werden. Zusätzlich zur reinen visuellen Repräsentation des Gesagten, ist laufend das Audio des Video-Streams zu hören, sodass der Anwender sich vergewissern kann, die Sprechblasen richtig zu Korrigieren. Fallen Sprechblasen und der Audio-Stream aus dem Zeitfenster, so werden diese wieder mit dem Video zusammengeführt und sind damit bereit für den Export.



**Technologien**  
Unsere Software setzt auf ASP.NET (Version 7) im Backend und Angular (Version 16) im Frontend. Zur Erstellung der Untertitel wird eine Speech-Engine verwendet. Das Streaming der Untertitel und der Audio zum Frontend, erfolgt durch SignalR, eine .NET Bibliothek die das Hinzufügen von Echtzeit-Webfunktionen erlaubt. Um mit verschiedenen Stream-Formaten intern umgehen zu können, wird FFmpeg im Backend und die Web Audio API im Frontend verwendet. Das Deployment besteht aus zwei geteiltes Docker-Image für Frontend und Backend. Mittels einer Docker-Compose werden diese Images zusammengeführt. Das Ausführen der Software wird dann via Podman auf den Servern der HTWK durchgeführt.

Wir danken allen Beteiligten: