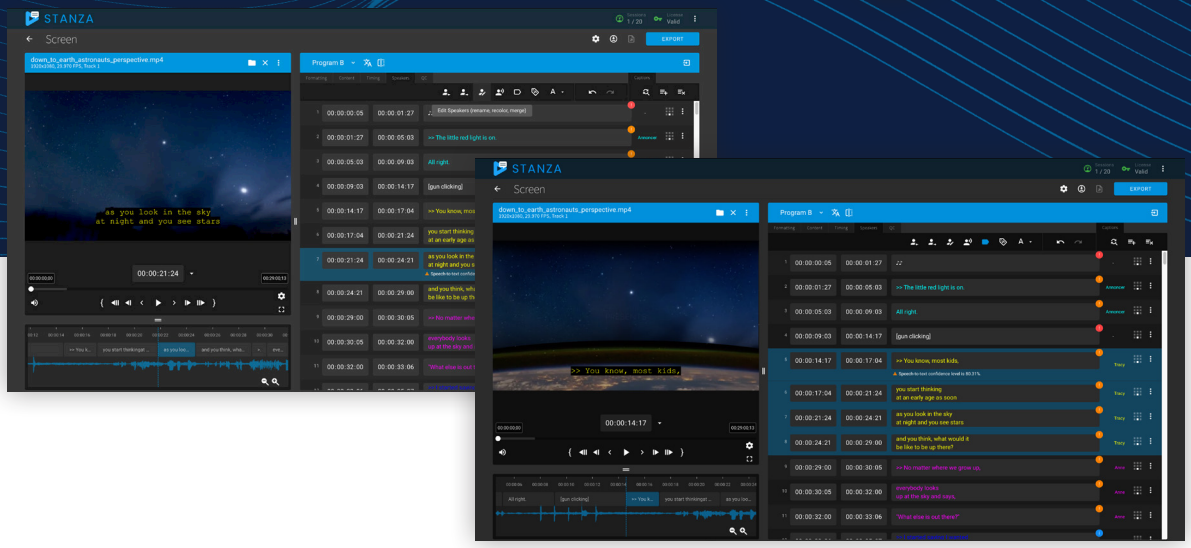
Stanza

Broadcast Quality Captioning & Subtitling



Executive Summary: Intelligent AI-Driven Speech-to-Text Transcription, Translation, & Speaker Identification

Stanza is the premier broadcast-quality captioning tool with state-of-the-art AI-powered speech-to-text, translation, and speaker identification that run locally on your Stanza server. With its browser-based UI and compatibility with all major operating systems, editors can access media files from anywhere. Stanza is powered by Telestream's GLIM engine for fast, uncomplicated playback of high-res files and integrates with Telestream Vantage for end-to-end workflow captioning.

Stanza meets all FCC regulations for closed captioning and is available in an operationally cost-friendly subscription model.

What's New in Stanza 4.0

- Speaker identification (diarization) automatically labels who is speaking, adds speaker change markers such as ">>" or "[NAME]", and saves reusable speaker profiles for recurring speakers.
- Profanity detection and substitution during speech-to-text and translation, with importable and exportable word lists.
- Faster, more accurate speech-to-text powered by Telestream Vantage AI Speech models, with better handling of fast conversation, overlapping speech, and music transitions.
- Improved AI translation, plus new V2 models for higher accuracy in less common languages.
- Smarter AI line breaking for speech-to-text, translation, Reformat Rows, and Square Up.
- Mouse editing on the waveform: drag caption edges to trim, roll, or ripple, or press C to split a caption.
- Simplified Single Sign-On (SSO) with Keycloak, with automatic SSL certificate generation and renewal.
- API tokens and more concurrent AI jobs for automated workflows.

Advantages of Stanza

- Offers user access through an intuitive browser interface across different operating systems - Windows, MacOS, or Linux — without requiring high-performance computers.
- Automatically transcribes speech to text using artificial intelligence for highly accurate captioning, and identifies and labels each speaker.*
- Supports many languages and provides a range of punctuation, smart AI text breaking, and formatting customization capabilities.
- Translates captions into many languages using AI processed locally on your Stanza server.
- Detects profanity and automatically substitutes words from customizable lists.
- Optimizes video playback for quality and bitrate, automatically scaling as needed based on network conditions and server load.
- Meets all FCC regulations for both TV and Internet closed captioning with the industry's leading support for CEA-708 caption file formats, QC and troubleshooting of caption files, plus automated adjustments and conversions.
- Supports integrated end-to-end captioning for Telestream Vantage, plus a REST API with user API tokens and multiple concurrent AI jobs for automation.
- Plays virtually all common video/audio formats without requiring transcoding proxies.
- Fast import on video files with a separate closed captioning track.
- Built on Telestream's industry-leading playback technology enabling proxy-less workflows and seamless captioning, removing the need to move large media files across a local area network.
- Provides time-saving customizable keyboard shortcuts and mouse editing on the waveform that facilitate user tailored and optimized captioning workflows.
- Supports secure Single Sign-On (SSO) through Keycloak / OpenID Connect.
- Offers a subscription model, avoiding the need for a large expense associated with permanent licenses or special equipment.

Technical Specifications

Import Formats

- Embedded 608/708 CC
- Embedded Teletext (OP47 / OP42)
- SCC
- MCC
- TTML (all profiles: IMSC, SMPTE-TT, iTT, EBU-TT, DFXP, and more)
- TTML in MXF (IMF / DCP)
- WebVTT
- SRT
- EBU-STL
- Spruce STL
- Cheetah CAP
- Lambda CAP (Japanese)
- PAC
- 890
- Plain Text with timecodes
- Plain Text (untimed / unformatted)
- Avid DS
- Ultech ULT
- TDS

Supported Video/Audio Formats

- MXF (OP1a, OP Atom)
- IMF
- MOV (including ProRes)
- MP4
- MPEG-2 Transport / Program stream (.ts / .mpg)
- WAV

*As with any automatic transcription system, we recommend checking the results and correcting any mistakes before exporting or delivering the captions.

Export Formats

- EBU-STL
- MCC
- SCC
- Spruce STL
- SRT
- TTML (all profiles)
- WebVTT
- Plain Text
- DCP (SMPTE / Interop)
- Subtitle Overlay
- Subtitle Burn-in
- Embed 608/708 CC
- Embed Teletext OP47 (up to 20 languages)
- Embed DVB Teletext (multiple languages)
- Avid DS
- Avid AAF

Supported Direct Play from Cloud / Network Storage

- S3 (AWS / Generic)
- Azure
- Google Cloud
- WebDAV
- Windows Shares (SAN / SMB / UNC paths)
- Physical media (local drives)

System Requirements (Stanza 4.0)

- Windows 11 or Windows Server 2022 / 2025
- macOS 15 (Sequoia) or later on Apple Silicon (Intel Macs not supported)
- NVIDIA RTX GPU (Pascal–Blackwell) with 8 GB VRAM for GPU-accelerated AI; 12 GB+ recommended for speaker identification and translation
- Multi-user servers and load-balanced deployments supported
- Clients: Google Chrome (version 150 or later recommended)

November 2026

Learn more about Stanza:
telestream.com/stanza

Want to schedule a demo?
Contact us at telestream.com/contact-us

Specifications may be subject to change. All pictures are for illustration purposes only. Telestream® and Stanza™ are registered trademarks or trademarks of Telestream 2, LLC. All other trademarks mentioned or used are the property of their respective owners. 2026 Telestream 2, LLC. All rights reserved.