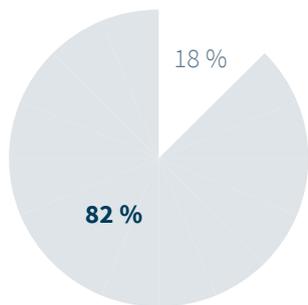# Human-Assisted AI Voice
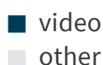## Solution Brief

Why Milengo?

Milengo is an ISO-certified language service provider with 30 years of experience in the industry. We are your one-stop shop for multilingual and multimedia production, from translation and transcription to professional voice-over services. We work with a wide range of global players in the e-learning industry to produce projects in over 70 different languages.

Milengo's brand-new Human-Assisted AI Voice service combines the latest AI-based voice technology with optimized human quality assurance.

- A future-proof alternative to traditional voice-over
- The perfect solution for multilingual e-learning initiatives
- Natural-sounding voices almost indistinguishable from those of a real human
- Human quality assessment and correction step to ensure your proprietary terminology is pronounced correctly
- Easy-to-edit audio recordings for further tweaks down the line
- Save up to 70% compared to the cost of traditional voice-over

18 %

82 %

Projections for global data traffic
by 2022
■ video
■ other

Listen for yourself and
hear the quality of our
Human-Assisted AI
Voices with your own
ears! **Click here for an
audio sample**

According to current forecasts, videos will account for around 82% of all global data traffic by 2022[1]. Audio-visual communication is the new normal. This presents a huge challenge for many businesses – not least because of the high costs and complexity associated with professional audio production. That's where Milengo's Human-Assisted AI Voice solution comes in! This cost-optimized service opens the door to new opportunities and will set you on the path to making multimedia content one of the cornerstones of your communications strategy.

AI-optimized voices + human quality assurance

Wherever you promote your company, it is essential to ensure that your corporate language is pronounced correctly – that goes for both basics like your company name, product names, and any other important terminology used in your industry. But with conventional text-to-speech plugins, this is far from guaranteed. The engines these tools use are susceptible to mispronunciation, especially when it comes to rare technical terms and neologisms. Where our service differs is that we enlist trained specialists to check your audio output and systematically correct the emphasis and flow of the spoken text. We ensure that both the audio and visual elements of your multimedia content are in line with professional standards, so your in-house production teams can focus on their real core tasks.

Much-needed relief for your multimedia budget

Traditional voiceover is a huge drain on both resources and personnel. Professional voice artists need to be cast and retained, sound equipment needs to be hired, and then the whole thing needs to be mastered and put through several rounds of corrections before it can finally be published. The fact that the production costs for text-to-speech are so much lower is one of the main reasons why it is so appealing. With Milengo's Human-Assisted AI Voice service, you can slash your expenditure by up to two thirds, thus enabling you to add high-quality voice output to more of your multimedia content and make more of your content overall accessible to everyone.

Simpler planning cycles

Milengo's Human-Assisted AI Voice service offers more than just lower costs – it can also help you to streamline your planning cycles, reduce the workload on your in-house teams and generally make your life less complicated. Professional voice artists need to be booked weeks or even months in advance, and often won't be available for the entirety of a long-term project. Even if you employ in-house voice talent to save costs, it is still a highly time-consuming process. Thanks to the reduced human involvement and technical workload it offers, TTS technology allows you to complete the same projects in just a fraction of the time. And of course, the voices of a text-to-speech engine are always available, so you don't have to worry about your original voice artist dropping out halfway through the project and ruining the consistency of your content.

[1] https://biteable.com/blog/
video-marketing-statistics/

The perfect choice for projects with frequent content updates

If you make regular changes to your audio-visual content, traditional voiceover probably isn't the right solution for you. Even minor corrections to your script can prove costly, as they mean you'll need to book a new session with your voice artist and carry out more expensive recording sessions in a professional studio environment. For many companies, it is impossible to reconcile the time and workload required with the benefits of re-recording. This leads to content quickly becoming outdated or even obsolete, especially in the field of e-learning. This is one of the key advantages text-to-speech has to offer: it allows you to edit the texts you have recorded as many times as you want, finally making the dream of regularly updated content with audio support a reality.

How does text-to-speech work?

Text-to-speech (TTS) is a software-based process that converts written text into spoken language. Modern text-to-speech technology is based on artificial intelligence, and every aspect of the process is tailored to the client's individual needs. As a Milengo customer, you can choose from over 220 voices in more than 40 different languages and set other parameters to suit your requirements, such as the pitch and speed of the voice, how it pronounces abbreviations, and how to tackle dates and times.

What applications can text-to-speech be used for?

This technology can be used for any kind of audio-visual content that is informative or descriptive. If you are producing expensive marketing videos or drafting sensitive employee communications, of course, it is still best to use a professional voice artist. But text-to-speech can be a valuable aid in plenty of other scenarios – from training videos and software demos to in-house safety briefings. Studies show that TTS-based audio experiences are particularly effective in increasing user engagement in e-learning courses and learning management systems.

Does text-to-speech meet the professional standards for corporate use?

Absolutely. The days of harsh, "robotic" voices with low-quality text delivery are long gone. Thanks to neural networks, the synthetic speech generated by the new models produces phonemes, syllables and words with a natural tone, stressed in a similar way to those of a human speaker. Users find these voices much more friendly and expressive, making that distracting, halting monotone a thing of the past.

Try out our Human-Assisted AI Voice service today as part of our free pilot project! Email us at sales@milengo.com to arrange a consultation.