

Crew training videos localized into new languages without reshoots or studio time

Client

An inflight digital services provider delivering connectivity, entertainment, and retail for over 50 airlines.

Company Size

140+ employees

Location

HQ in Barcelona, Spain

Featured Partners



The client operates crew training videos in a single source language. Global expansion meant every new region needed the same training in a new language, and traditional re-recording is cost-intensive, slow, and operationally complex. Marlabs proactively designed and delivered a working AI-powered multilingual dubbing solution that produces professional-quality voice-over videos without recreating the original content.

The solution runs a seven-stage automated pipeline that segments each video into 15-second scenes, transcribes, translates, synthesizes a natural AI voice, time-aligns the audio, dubs it over the original track, and assembles the final video. Re-syncing at every segment boundary keeps long-form videos aligned end to end. The client can add languages on demand, and the capability is production-ready and scalable.



Cognitive AI



Generative AI



AI Automation



Application Development

The Challenge: Crew training across global locations existed in only one language



Objective: Enable rapid multilingual crew training across new regions without recreating the original video content.



Existing Issues: Re-recording every video in every language was cost-intensive, slow, and operationally complex, and manual localization drifted out of sync on long-form videos.



Solution Needed: Marlabs built an AI-powered dubbing pipeline that segments, transcribes, translates, voices, time-aligns, and assembles each video automatically.



Outcome: The client gained a production-ready capability that adds new languages on demand with minimal incremental effort.



The client's crew training library existed only in its original source language. Expanding into new regions meant every crew needed the same standardized training, but re-recording each video per language demanded studio time, voice talent, and scheduling the business could not absorb.

The Solution: A seven-stage AI pipeline that dubs any video into any target language

Marlabs delivered a working solution built on Azure AI services. It splits each video into 15-second scenes, then transcribes, translates, voices, time-aligns, dubs, and assembles them into a finished video. Re-syncing at every boundary keeps long-form training aligned end to end.

Phase 1: Discovery & Pipeline Design

Marlabs assessed the client's crew training library and designed a segment-based architecture that reuses the original footage as-is.

Workstreams:

- Source video assessment
- Pipeline architecture design
- Language scope definition

Phase 2: Segmentation & Transcription

The team used PySceneDetect to split each video into 15-second scenes and Azure AI Speech to transcribe every segment.

Workstreams:

- Scene detection
- Segment-level transcription
- Timing reference capture

Phase 3: Translation & Voice Synthesis

Our team translated each transcript with Azure AI Translator and generated a natural AI voice track with Azure AI Speech.

Workstreams:

- Transcript translation
- Neural voice generation
- Target language selection

Phase 4: Time-Alignment, Dubbing, & Assembly

We fit the new audio to each segment, dubbed it over the lowered original track, and stitched the final video together with FFmpeg.

Workstreams:

- Segment time-alignment
- Audio ducking & overlay
- Final video assembly

Services and Technologies Used:

Services:

- Cognitive AI
- Generative AI
- AI Automation
- AI Strategy
- Application Development
- Platform Development

Technologies:

- Azure AI Speech
- Azure AI Translator
- FFmpeg
- PySceneDetect
- Streamlit

The Results: Impact on the client organization

The client received a working, production-ready solution rather than a concept. A single-language training library can now reach new regions as a pipeline run instead of a production, and the same capability extends to product, compliance, and onboarding videos across the business.



Faster Regional Rollout: The solution turned each new-language rollout into a pipeline run rather than a studio production.



New Languages on Demand: The solution supported any source and target language pair so that adding a language took minimal incremental effort.



Lower Localization Cost: The team removed reshoots, voice talent, and studio time from the localization process by reusing the original footage as-is.



No Drift on Long-Form Video: The pipeline re-synced timing at every 15-second boundary, which kept long training videos aligned from start to finish.



Consistent Training Everywhere: The client kept one standardized training message across every region and airline crew since each language came from the same source.



A Reusable AI Capability: Marlabs delivered a production-ready capability that scales across languages, regions, and other video libraries in the business.