Asianet News
India | 2020

Google
News Initiative

Innovation
Funding

Asianet News develops an AI translation engine to automatically convert videos into different Indian vernacular languages

Asianet News Media and Entertainment Pvt. Ltd. (Asianet News) is a leading digital news provider in Southern India that publishes across six Indic languages: Malayalam, Kannada, Hindi, Bangla, Tamil, and Telugu.

OPPORTUNITY

Asianet News wanted to use technology to reduce the resource and time requirements associated with producing news content in different Indian vernacular languages.

APPROACH

Asianet News developed a new tech-based approach to quickly and easily produce news content in many Indic languages. Previously, content was created in each separate region and translated manually before it was published on the channel.

To overcome this intensive process, the team developed a machine learning (ML) program that translates content into multiple languages — the central video production team creates content in one language, then feeds that through the ML program, where it's automatically translated into other Indic languages. While these scripts still need to be manually corrected, this centralized process has helped Asianet News reduce duplication, minimize costs, and increase overall production capacity for both translated and untranslated videos.

ASIANET NEWS DEVELOPED A ML PROGRAM THAT STREAMLINED CONTENT TRANSLATION

1 Initially a video is produced in English (or any other source language)

2 The script of this video is provided as feed to translation box

3 Translated output is corrected manually

4 Voice over with the target language script is done in the same video

5 Final video outputs in Malayalam, Kannada, Hindi and Bangla are posted on the respective websites

Building up the technology team: Asianet News set up an ML team to develop AI-based Automatic Speech Recognition (ASR) and translation technologies. The team explored technology service providers, and then onboarded an ASR and cloud provider that offered the best services for different Indian languages.

Building the central video team: Next, Asianet News established and trained a central video production team, the Central Synergy Team. Under the new model, this team is responsible for producing the primary video content, which then gets translated into the various languages. The Team became the nodal point for the development of videos across platforms, and a

centralized process has reduced duplication and increased overall video production capacity, even in the case of non-translated videos.

Developing the technology: The next step was to build the technology that would allow Asianet News to quickly and easily translate content produced in one Indic language into many others, allowing them to reach a broader audience with each story. To start, Asianet News built its translation engine on the Google cloud translation API. It then needed to train up the translation engine for traditionally underserved Indic languages, so the Team provided learning data based on existing translations, upon which manual corrections had been completed. Next, they developed an interface that had the required APIs for voice recognition, translation, manual correction of script errors, editing, and documenting setup so that the editors and production team could reproduce the videos.

Training the team: Editors and staff in the video production team were given in-depth training in how to use the video translation tool. Asianet News created standardized workflows and processes for video content production and translation. Other training was provided to staff as necessary across production, post production, video meta-data, SEO, managing copyright on different social media platforms, animation, and drone and 360 video production.

Streamlining model usage: To make life easier for the video editors, Asianet News created a one-window dashboard to handle the input and output stream. At the backend, the model captures the manual corrections at each stage.

Launching the new content: To publicize and raise awareness of the translated videos, Asianet News ran promotional content on all its websites and social media channels, in seven languages.
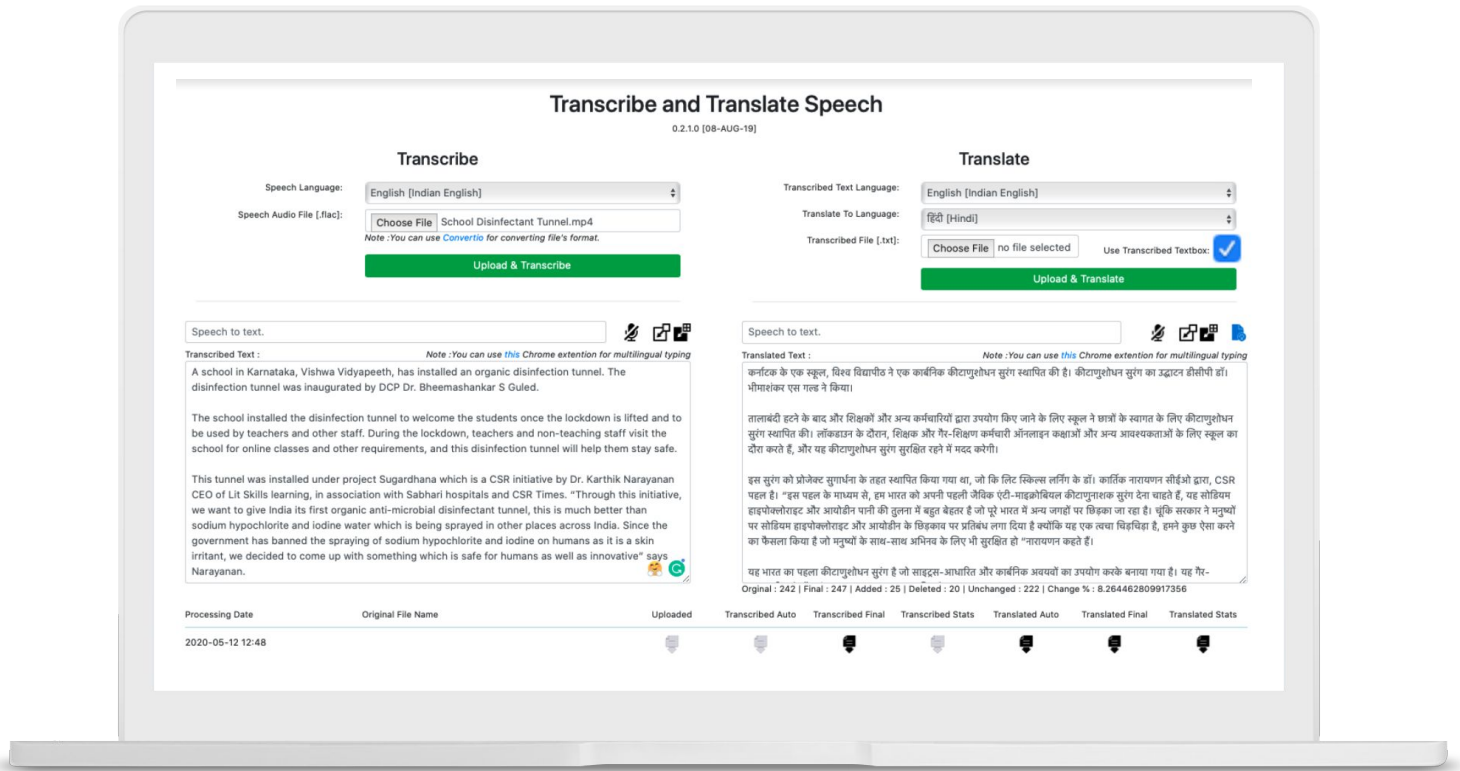
Increasing reliability: Asianet News are iterating on the translation tool to tweak and improve things like voice recognition, accuracy, and efficiency — it is ultimately aiming for 95% accuracy, but isn't there yet. The model is most efficient while processing data from English to Hindi with results varying from 80-90% efficiency. This is mainly due to the higher learning database of parallel sentences available for this segment. Next is English to Bangla with around 74% efficiency, followed by Tamil and Malayalam with 70% efficiency. For Kannada and Telugu it's around 60%. By its nature, the more translation the tool undertakes, the more reliable it will become.

The process after: Asianet News' ML translation tool in action

RESULTS

Asianet News successfully developed a reliable AI-based model for translating video from one language to another, and were able to create more news videos at scale for six Indian vernacular languages.

- The translation model has **65-75% efficiency for automatic voice recognition.**
- The model has **70-80% efficiency for translation.**
- This resulted in a **2X increase in overall video production** capacity, allowing Asianet News to increase video upload volume.

"

The project has been successful in transforming the mindset of our organization — we're now more of a tech-led news media agency focused on video-first content.

"

Abhinav Khare,
CEO of Asianet News Network

LEARNINGS

Look for simple ways to increase efficiency: Asianet News looks for all kinds of ways to enhance the efficiency and accuracy of its product. For example, the team maintains an archive of video files (in low resolution) with scripts and proper tagging to use as training data for future projects. The Central Synergy Desk also reduces duplicate content.

Teach everyone about the benefits of technology: For optimum output from such a big technology investment, it is very important to keep everyone in the loop about developments. Asianet News ran regular training sessions with editors and producers so they could get the most out of the translation product. With a greater understanding of technological solutions, the editorial and production teams were excited about bringing the benefits of technology into their daily workflows.

Don't rush the transition from old to new tools: The ML translation tool is constantly being improved. Currently, the output still needs human correction and voiceover to become more appealing to the users. To maintain the video performance index of its channel and to ensure audience engagement, Asianet News has an optimum mix of newly developed video formats and existing ones.

Focus on national news: If you're producing news in multiple languages, align your editorial teams so they're focused on national news and not niche topics that are only relevant for one small region. This helps standaradize content at the production stage so that minimal processing is required once the content is translated.