AsiaSat | AMT

LLMS IN ACTION:
ENHANCING TIME-SYNCED SUBTITLES FOR FILE-BASED & LIVE CONTENT

A BROADCASTER'S GATEWAY TO GLOBAL REACH

In today's broadcast and OTT landscape, subtitles have evolved beyond mere compliance requirement; they are essential for connecting with international audiences. By providing accessibility, subtitles open up new markets. However, audiences can quickly notice when subtitles lag, feel awkward, or fail to match the natural rhythm of speech. Such issues can undermine viewer trust and, ultimately, brand credibility.

While traditional neural machine translation (NMT) has facilitated subtitle localization, it has inherent limitations. As NMT processes each segment independently, it cannot retain the full context of conversations. This results in translations that may feel clumsy or disjointed, particularly in fast-paced live environments where accuracy and timing are critical. Large Language Models (LLMs), on the other hand, provide a more holistic approach, allowing for smoother and more coherent translations that resonate better with viewers.

THE TIME SYNCHRONIZATION CHALLENGE

A typical subtitle localization workflow begins with transcribing spoken content and segmenting it using time codes. Each segment is translated separately before being reassembled into a complete subtitle track. This process introduces two major challenges. First, without a broader transcript, translations may lack coherence and lose the intended meaning. Second, subtitles may not align smoothly with the speaker's rhythm, and as a result, disrupting the viewing experience.

For live localization, the challenges intensify. Since traditional NMT systems cannot carry translation history forward, and treat each segment in isolation, thus hindering fluent and consistent subtitles throughout the program. For broadcasters, these limitations lead to increased manual intervention, higher costs, and the risk of losing audiences due to poor quality.

HOW LLMS TRANSFORM SUBTITLE LOCALIZATION

LLMs address these challenges by processing the entire transcript along with timestamps, rather than in isolated fragments. This holistic approach enables them to generate coherent and natural translations across segments while ensuring precise synchronization with time codes.

In live broadcasts, LLMs can dynamically reference the translation history, allowing for real-time subtitles refinement and a smooth viewing experience. For viewers, the difference is evident: subtitles flow naturally and stay perfectly in sync, all while the process remains fast and scalable for broadcasters.

WHY THIS MATTERS FOR BROADCASTERS

The implications for broadcasters and OTT platforms are significant. LLMs improve translation quality by producing human-like subtitles that are tightly synchronized with audio. They also make large-scale subtitling more affordable, as smaller models such as AWS Nova and Qwen-MT have shown that they can deliver excellent results at a lower computational cost. Moreover, LLMs can support terminology intervention, guiding the use of predefined glossaries for specialized or technical terms, which is essential in fields like sports, legal programming, or technical content.

Equally important, LLM-based workflows are adaptable to both file-based operations, such as OTT library localization, and live events, including sports and news. This versatility allows broadcasters to scale their localization efforts without sacrificing quality or inflating operational budgets.

CHALLENGES TO KEEP IN MIND

Like any emerging technology, LLMs come with challenges. They may occasionally produce inaccurate or fabricated text, a phenomenon often described as "hallucination". This risk can be mitigated through careful input design and monitoring, but it remains a consideration. Latency is crucial, especially in live broadcasting where every second counts, and systems must deliver translations at near-instant speed. Finally, the computational resources required

for large models can be significant, necessitating a balance between quality and efficiency. Many broadcasters are now exploring hybrid approaches that combine the fluency of LLMs with the cost-efficiency of NMT, creating reliable and practical workflows.

FUTURE HORIZONS FOR BROADCAST LOCALIZATION

The possibilities extend beyond subtitles alone. LLM-driven workflows can also facilitate dubbing by generating natural, time-aligned scripts for voiceovers. With open-source transcription tools such as Whisper, paired with smaller LLMs, the entire localization pipelines can be operated on-premises. This reduces dependence on cloud infrastructure, enhances data security for sensitive content, and lowers operating costs.

Looking ahead, multimodal LLMs, that integrate audio, video, and text, promise an even richer contextual understanding of content. This evolution could unlock subtitling and dubbing experiences that feel seamless and intuitive, driving deeper engagement with global audiences.

SUMMARY

For Asian broadcasters and OTT platforms, LLMs represent more than a technical upgrade. They address the long-standing timing and context issues of traditional NMT, enabling the delivery of high-quality, synchronized subtitles at scale. The outcome is a smoother viewing experience, reduced operational costs, and greater confidence in content readiness for global distribution.
LLMs are not just aiding audiences in understanding the words on screen; they are enhancing subtitles credibility, engagement, and future readiness for a global stage. ∎

by Andy Chang
Chief Technology Officer, AMT

ASIASAT MEDIA TECHNOLOGY LIMITED (AMT), *a subsidiary of AsiaSat, specializes in innovative media technology solutions designed specifically to enhance media and broadcast capabilities throughout the Asia-Pacific region. AMT's advanced services include high-quality live-streaming, AI-driven localization, and comprehensive broadcasting technologies, addressing the evolving dynamics of the video ecosystem in close partnership with broadcasters across the region.*