Tích hợp lồng tiếng âm thanh dựa trên AI vào phát trực tiếp với AWS Media Services và CAMB.AI

**📖 Bài viết gốc:** https://aws.amazon.com/vi/blogs/media/integrating-ai-based-audio-dubbing-into-live-streaming-with-aws-media-services-and-camb-ai/  
**👤 Tác giả:** Chris Zhang  
**📅 Ngày xuất bản:** 09 tháng 7 năm 2025  
**🌐 Nguồn:** AWS Blog  
**👨‍💻 Người dịch:** Nguyễn Quang Bình  
**📅 Ngày dịch:** 08/07/2025  
**⏱️ Thời gian đọc:** 15 phút

──────────────────────────────────────────────────

## 📋 Tóm tắt

Khi nhu cầu về các luồng video trực tiếp tăng trưởng nhanh chóng, các nhà phát sóng và nhà sản xuất đối mặt với những thách thức phức tạp và chi phí đáng kể để làm cho nội dung có sẵn cho khán giả toàn cầu, đa ngôn ngữ. Lồng tiếng trực tiếp track âm thanh sang một ngôn ngữ nói khác theo truyền thống...

**🎯 Đối tượng đọc:** Developers, Solution Architects, DevOps Engineers  
**📊 Độ khó:** Intermediate  
**🏷️ Tags:** AWS, Cloud Computing, Architecture

──────────────────────────────────────────────────

## 📚 Mục lục

• Giới thiệu  
• Kiến trúc giải pháp  
• Triển khai  
• Kết luận  
• Glossary - Thuật ngữ  
• Tài liệu tham khảo

──────────────────────────────────────────────────

Khi nhu cầu về các luồng video trực tiếp tăng trưởng nhanh chóng, các nhà phát sóng và nhà sản xuất đối mặt với những thách thức phức tạp và chi phí đáng kể để làm cho nội dung có sẵn cho khán giả toàn cầu, đa ngôn ngữ. Lồng tiếng trực tiếp track âm thanh sang một ngôn ngữ nói khác theo truyền thống là một quy trình rất thủ công, điều này thêm chi phí và độ phức tạp đáng kể cho các nhà sản xuất. Nhờ những tiến bộ mới nhất với lồng tiếng AI, và khả năng mở rộng và tính linh hoạt được cung cấp bởi đám mây Amazon Web Services (AWS), các giải pháp mới có thể giải quyết những điểm đau di sản này. Lồng tiếng âm thanh AI cho phép người dùng phân phối luồng trực tiếp bằng các ngôn ngữ khác nhau, được xử lý tự động và nhúng trong luồng vận chuyển video gốc.

Lồng tiếng truyền thống đắt đỏ do nhu cầu về diễn viên lồng tiếng chuyên nghiệp, kỹ sư âm thanh và quy trình sản xuất phức tạp. Nền tảng được hỗ trợ bởi AI của CAMB.AI tự động hóa quy trình lồng tiếng, giảm đáng kể chi phí bằng cách loại bỏ nhu cầu về nguồn nhân lực rộng rãi và thiết lập studio. Hơn nữa, quản lý độ trễ luồng trực tiếp và đồng bộ hóa giữa nhiều âm thanh lồng tiếng là thách thức.

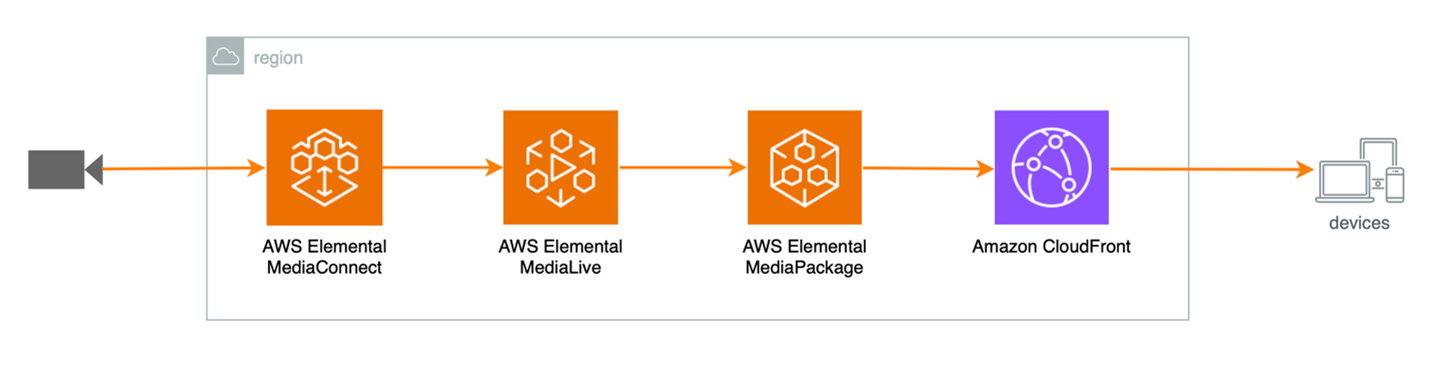
Trong bài viết này, chúng tôi giới thiệu bạn với AWS Partner CAMB.AI, người cung cấp một giải pháp lồng tiếng AI có thể dễ dàng cắm vào luồng kiến trúc AWS Media Services. Chúng tôi giải thích cách lồng tiếng AI có thể tích hợp vào một kiến trúc phát trực tiếp có khả năng phục hồi và dự phòng, để cung cấp thể thao trực tiếp, tin tức và sự kiện cho khán giả trên toàn thế giới. Hơn nữa, chúng tôi đưa ra một bản xem trước của các bước cần thiết để bắt đầu.

**Về CAMB.AI**

CAMB.AI, một AWS Partner, là một công ty tổng hợp giọng nói và dịch thuật AI tiên tiến. Nền tảng Dub Stream của CAMB.AI cung cấp lồng tiếng thời gian thực được hỗ trợ bởi AI sang hơn 140 ngôn ngữ, tất cả trong khi bảo tồn tông giọng của người nói và sắc thái cảm xúc. Các track âm thanh lồng tiếng thay thế được thêm vào cho phát trực tiếp cho phép các tổ chức cung cấp cho khán giả toàn cầu nhiều tùy chọn hơn để truy cập các luồng của họ. CAMB.AI rất phù hợp với người dùng AWS vì nó có thể được tích hợp liền mạch với AWS Elemental Media Services. Người dùng có thể linh hoạt nhúng một hoặc nhiều track ngôn ngữ lồng tiếng mỗi luồng. CAMB.AI tích hợp vào quy trình làm việc phát trực tiếp thông qua dịch vụ AWS Elemental MediaConnect, hỗ trợ phân phối độ trễ thấp, an toàn của các luồng vận chuyển.

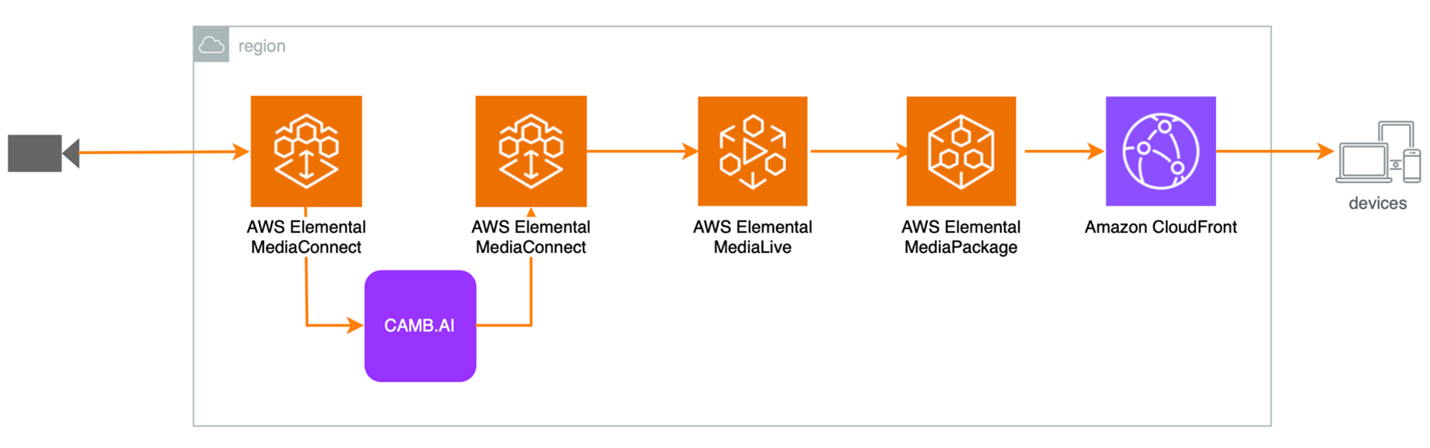
**Tổng quan giải pháp: kiến trúc phát trực tiếp với lồng tiếng âm thanh**

Hình sau đây cho thấy một quy trình làm việc phát trực tiếp sử dụng AWS Media Services, như MediaConnect, AWS Elemental MediaLive, AWS Elemental MediaPackage, và Amazon CloudFront. MediaConnect cung cấp vận chuyển an toàn video từ các nguồn tại chỗ lên đám mây. MediaLive chuyển mã video thành nhiều luồng thích ứng bitrate. MediaPackage phục vụ như máy chủ gốc và packager, và CloudFront là mạng phân phối nội dung (CDN) phân phối luồng trực tiếp đến người xem. Việc sửa đổi một luồng trực tiếp sau giai đoạn chuyển mã là phức tạp, do đó chúng tôi khuyến nghị tích hợp các ngôn ngữ lồng tiếng sớm trong luồng stream, trước khi được chuyển mã bởi một dịch vụ như MediaLive.



Hình 1: Kiến trúc phát trực tiếp trên AWS.

Hệ thống Dub của CAMB.AI chấp nhận các luồng giao thức Secure Reliable Transport (SRT) cho ingest và egress, và SRT cũng được hỗ trợ nguyên bản bởi MediaConnect. Một luồng được khuyến nghị có thể xem trong hình sau, mô tả một bộ mã hóa video đóng góp để stream đến một endpoint MediaConnect, được CAMB.AI nhận, và sau đó chuyển trở lại MediaConnect trước khi chảy đến MediaLive. Sau khi người dùng cấu hình MediaLive để ingest các track âm thanh lồng tiếng được thêm vào từ luồng SRT đến, và chuyển mã chúng sang định dạng đầu ra thích hợp để phân phối, điều này cho phép tích hợp liền mạch với một kiến trúc phát trực tiếp có khả năng mở rộng cao và có sẵn tiêu chuẩn ngành sử dụng MediaLive, MediaPackage, và CloudFront. Các trình phát video HTML5 thiết bị khách hàng hiện đại cũng thường có hỗ trợ nguyên bản các kênh âm thanh ngôn ngữ thay thế và làm cho chúng có thể chọn được cho người dùng cuối thông qua GUI trình phát.



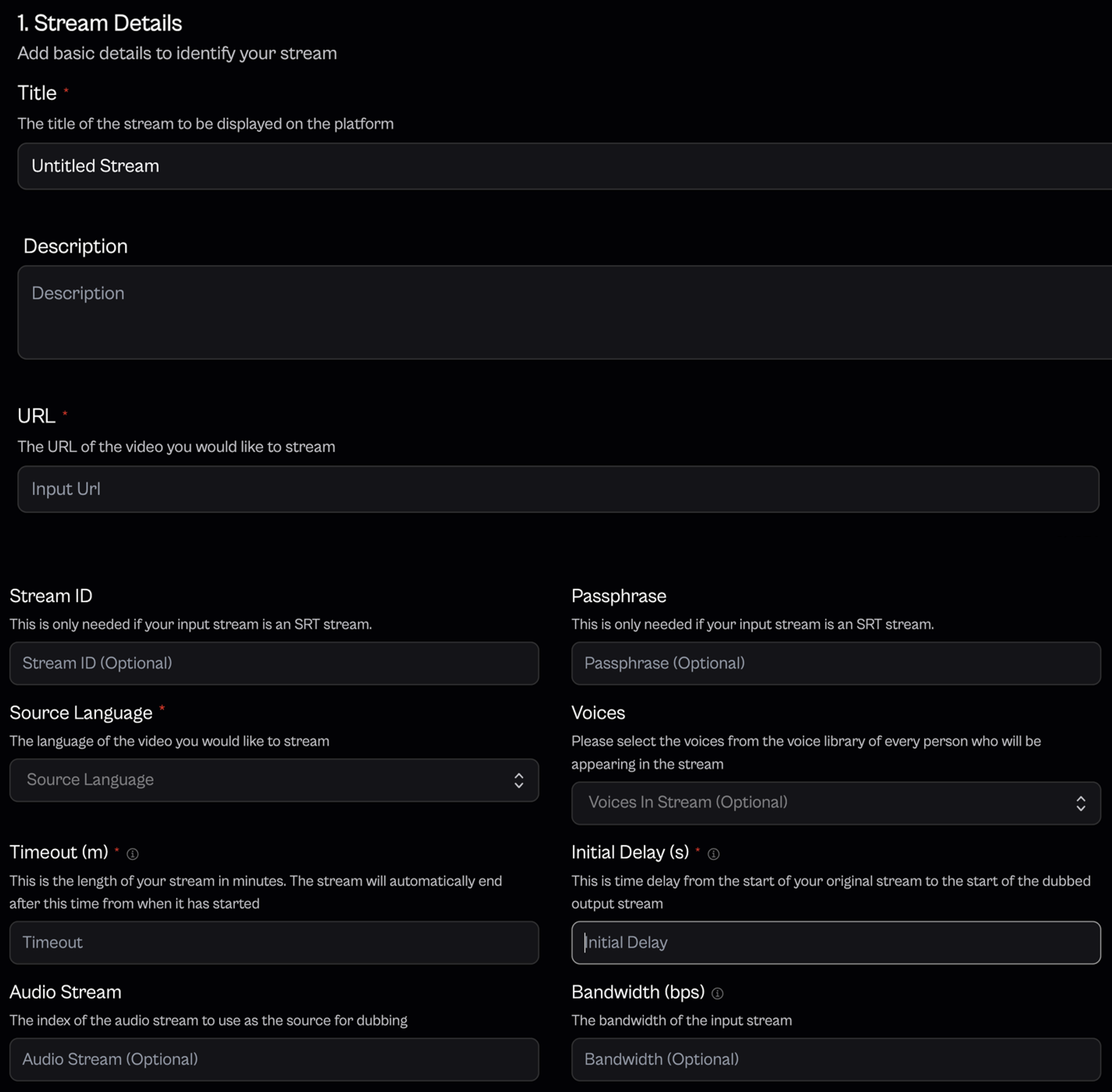
Hình 2: Giải pháp CAMB.AI được tích hợp vào quy trình làm việc phát trực tiếp.

**Bắt đầu với CAMB.AI**

Khi bạn có một tài khoản với thông tin đăng nhập truy cập từ CAMB.AI, bạn có thể thiết lập tích hợp lồng tiếng AI.

Hình sau đây cho thấy giao diện người dùng để cấu hình đầu vào luồng cho Dub System. Trường URL là nơi bạn có thể cấu hình giao thức SRT với trường Stream ID và Passphrase tùy chọn. Ví dụ, srt://<ip address>:<port number>. Tham khảo tài liệu dịch vụ để biết chi tiết cụ thể về cấu hình một SRT listener hoặc caller trong MediaConnect.

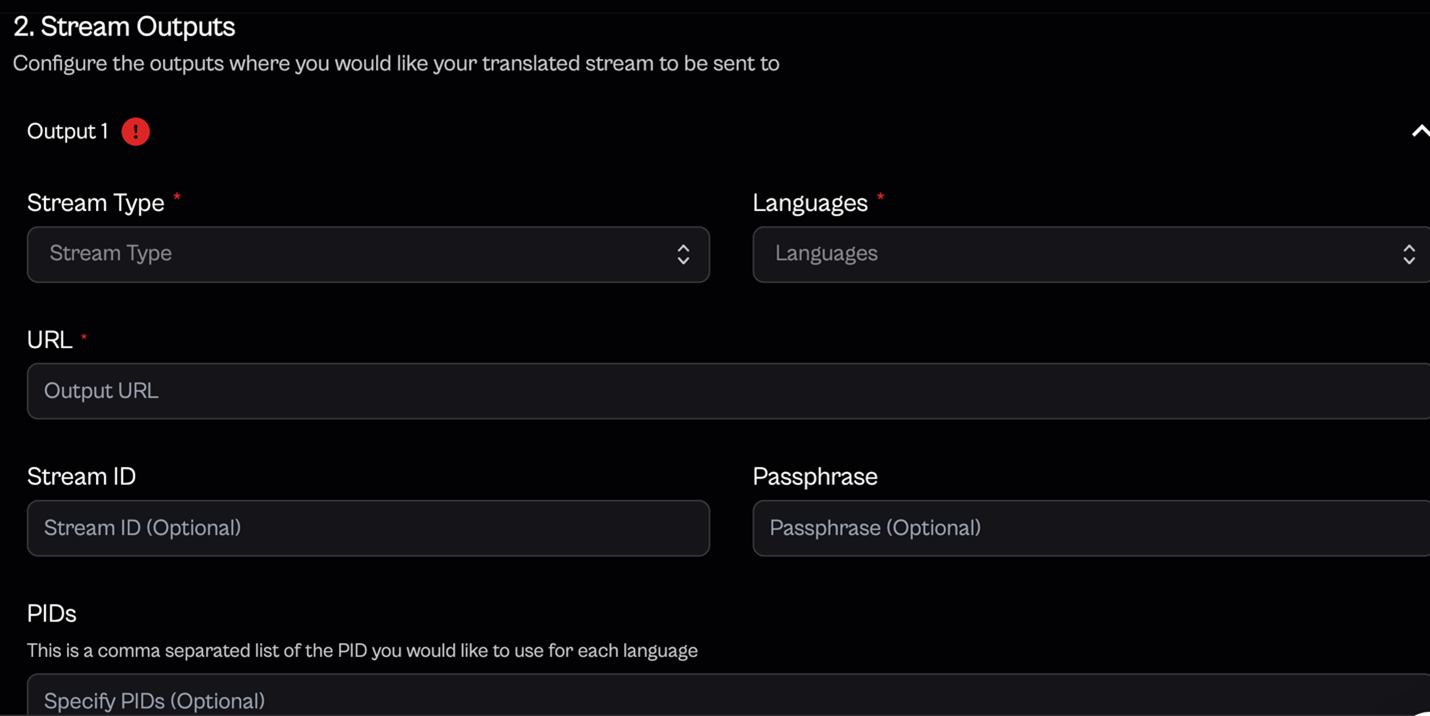
Dropdown Source Language cho phép bạn chọn ngôn ngữ của âm thanh gốc. Dub System cũng cho phép người dùng chọn một giọng nói cụ thể thông qua menu dropdown Voices.



Hình 3: Đầu vào luồng CAMB.AI Dub System.

Khi đầu vào luồng được cấu hình, bước tiếp theo là cấu hình đầu ra âm thanh lồng tiếng AI. CAMB.AI không sửa đổi video và âm thanh gốc của bạn, nhưng giải pháp trộn âm thanh được tạo bởi AI như các kênh âm thanh được thêm vào trong luồng SRT.

Hình sau đây cho thấy các tùy chọn để cấu hình một luồng đầu ra trong Dub System. Bạn chọn giao thức thích hợp trong trường Stream Type và chọn các ngôn ngữ mục tiêu trong trường Languages để lồng tiếng hoặc dịch thuật.



Hình 4: Đầu ra luồng CAMB.AI Dub System.

Sau khi cấu hình các tham số đầu vào đầu ra, khởi tạo luồng của bạn bằng cách chọn Start Stream. Tại thời điểm này, luồng của bạn được tăng cường với âm thanh lồng tiếng được tạo bởi AI và được tích hợp liền mạch vào quy trình làm việc phát trực tiếp. Tham khảo hướng dẫn người dùng MediaLive Input settings—Audio selectors để cấu hình MediaLive sử dụng đúng track.

**Kết luận**

Bài viết này mô tả cách kết hợp công nghệ lồng tiếng AI, như CAMB.AI, vào quy trình làm việc phát trực tiếp. Tích hợp này cho phép người dùng cung cấp phát trực tiếp được bản địa hóa lồng tiếng AI cho khán giả toàn cầu. Để tìm hiểu thêm về tạo một sự kiện trực tiếp sử dụng AWS, tham khảo giải pháp Live Streaming on AWS của chúng tôi, giúp kiến trúc. Nếu bạn quan tâm đến việc bắt đầu với AWS Media Services, hãy truy cập trang sản phẩm. Đối với giải pháp lồng tiếng âm thanh của CAMB.AI, hãy truy cập CAMB.AI.

──────────────────────────────────────────────────

## 📖 Glossary - Thuật ngữ

|  |  |  |
| --- | --- | --- |
| English | Tiếng Việt | Định nghĩa |
| Auto Scaling | Tự động mở rộng quy mô | Khả năng tự động tăng/giảm resources dựa trên demand |
| Load Balancer | Bộ cân bằng tải | Phân phối traffic đến multiple servers |
| Microservices | Kiến trúc microservices | Architectural pattern chia application thành small services |

## 🔗 Tài liệu tham khảo

**Tài liệu gốc**

• AWS Blog: https://aws.amazon.com/vi/blogs/media/integrating-ai-based-audio-dubbing-into-live-streaming-with-aws-media-services-and-camb-ai/

• Related Articles: https://aws.amazon.com/vi/blogs/media/delivering-low-latency-captions-and-voice-translation-for-live-sports-news-and-ott-platforms-with-syncwords-aws/

──────────────────────────────────────────────────

## 💬 Ghi chú của người dịch

**Challenges trong quá trình dịch**

• Technical Terms: Giữ nguyên các thuật ngữ AWS để đảm bảo tính chính xác

• Cultural Context: Điều chỉnh ví dụ cho phù hợp với bối cảnh Việt Nam

• Complex Concepts: Giải thích rõ ràng các khái niệm phức tạp

**Insights gained**

• Technical Learning: Hiểu sâu hơn về kiến trúc AWS và best practices

• Language Skills: Cải thiện kỹ năng dịch thuật kỹ thuật

• Industry Knowledge: Nắm bắt xu hướng công nghệ mới

──────────────────────────────────────────────────

## 🤝 Đóng góp và Feedback

Bài dịch này được thực hiện trong khuôn khổ FCJ Internship Program.

**📧 Liên hệ:** ngqbinh456@gmail.com.  
**💬 Feedback:** Mọi góp ý để cải thiện chất lượng dịch thuật xin gửi về email trên  
**🔄 Updates:** Bài dịch sẽ được cập nhật dựa trên feedback từ cộng đồng

──────────────────────────────────────────────────

© 2024 - Bản dịch thuộc về Nguyễn Quang Bình. Vui lòng credit khi sử dụng.