

2019年1月23日

音声・ビデオ通話のデータをクラウドサービスと簡単に連携できる 「Media Pipeline Factory」を提供開始 ～レアジョブとの英会話習熟度の可視化実験に活用～

NTTコミュニケーションズ（以下 NTT Com）は、「Enterprise Cloud WebRTC Platform SkyWay（以下 SkyWay）」の音声・映像データを音声認識 AI や機械翻訳 AI、ライブ配信などのさまざまなクラウドサービスとリアルタイムで連携できる SkyWay 簡単機能拡張キット「Media Pipeline Factory（以下 本キット）」を開発しました。2019年1月23日より、無料トライアルの提供を開始します。

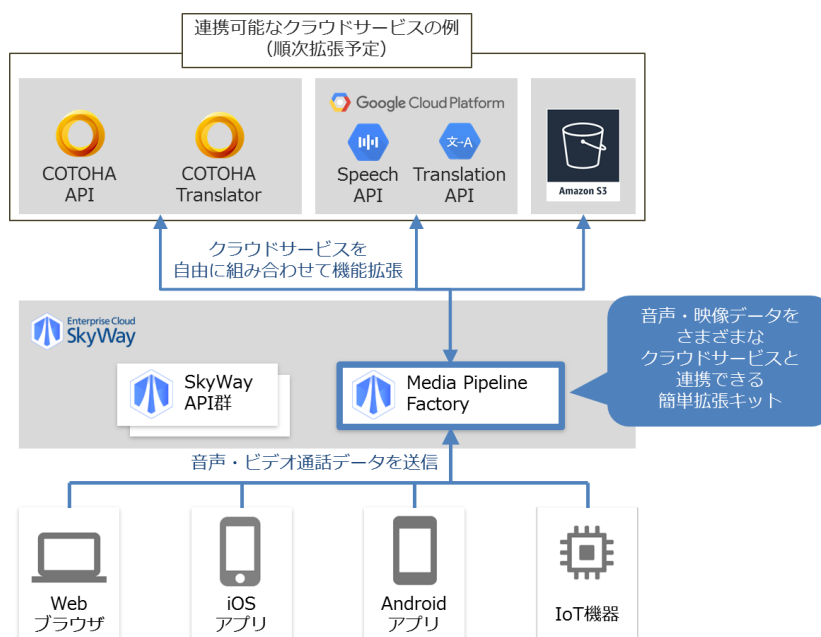
1. 背景

近年、特に音声認識、機械翻訳、画像認識などの分野において、クラウド上で AI 機能を提供するサービスが登場しています。しかし、WebRTC^{*1}により音声・映像データを、それらのクラウドサービスに送信して AI 機能と連携させるためには、個別のシステム開発が必要になるなど、多大なコストが生じる課題がありました。

2. 概要

今回、NTT Com は SkyWay^{*2}の音声・映像データを、対話型自然言語解析 AI エンジン「COTOHA Virtual Assistant」（以下 COTOHA VA）^{*3}をはじめとする、さまざまな AI やクラウドサービスと自由に組み合わせることが可能なキットの提供を開始します。WebRTC の機能拡張を容易にし、音声・映像データの活用領域を拡大させることで、お客さまのデジタルトランスフォーメーション（DX）を加速させます。

SkyWay 簡単機能拡張キット「Media Pipeline Factory」のイメージ



NTTコミュニケーションズ株式会社 広報室

NTT Communications Corporation Public Relations Office

〒100-8019 東京都千代田区大手町 2-3-1 大手町プレイスウエストタワー

OTEMACHI PLACE WEST TOWER 2-3-1 Otemachi, Chiyoda-ku, Tokyo 100-8019, Japan

Tel (03)6700-4010 International +81 3 6700 4010

3. 本キットの特長

本キットは、以下の3つの特長を備えます。

(1) 多用途な活用が可能

本キットにより、SkyWay で送信した音声・映像データと多様なクラウドサービスを組み合わせ、さまざまな業界や利用シーンで DX を実現できる機能を提供します。

<代表的なユースケース>

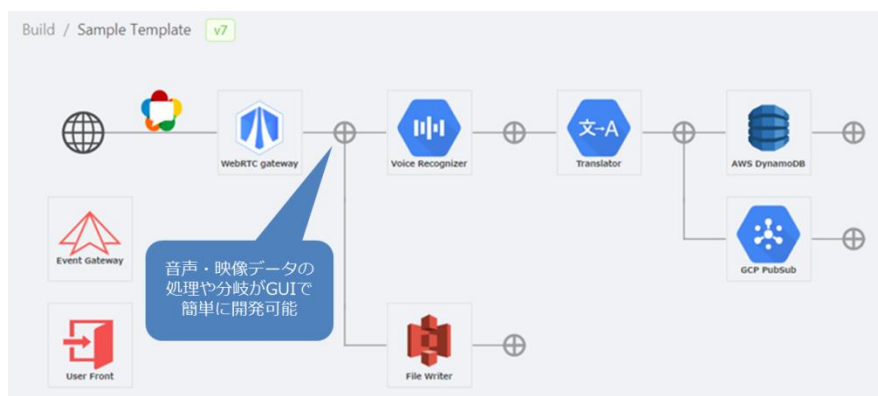
- ・ AI サービスとの組み合わせによる、音声認識、機械翻訳、画像認識機能の提供
- ・ ストレージサービスとの組み合わせによる、録音録画機能の提供
- ・ SIP サーバーとの組み合わせによる、VoIP との相互通話機能の提供
- ・ スタジアム Wi-Fi や CDN との組み合わせによる、低遅延ライブ配信機能の提供

(2) 素早い開発が可能

グラフィカルユーザーインターフェース（GUI）上でコンポーネントを並び替えることにより、シンプルな処理の開発であれば、プログラミングが不要となります。

また、コンポーネントに分割されたマイクロサービスアーキテクチャを採用しており、複雑な処理を伴う開発の際も、高いアジリティと柔軟性を実現します。

<Media Pipeline Factory の操作画面イメージ>



さらに、各コンポーネントのオーケストレーション基盤も合わせて提供し、ログやサービスの状態を一元管理することが可能となります。これにより、デバッグが容易になり、開発効率の向上を実現します。

(3) スケーラブルな対応が可能

サーバレス・アーキテクチャを採用しており、一度に大量のアクセスが発生した場合や、アクセスの急増に対しても、安定した機能提供を実現します。

4. 無料トライアルについて

2019年1月23日より、録音、音声認識、機械翻訳の各機能をプログラミングすることなく利用可能な6種類のビルトインコンポーネントを、無料トライアルとして提供します[別紙]。

お申し込み方法などの詳細につきましては Web サイト(<https://webrtc.ecl.ntt.com/m-pipe/>)をご参照ください。

5. 今後について

今後も継続してビルトインコンポーネントの開発を行い、録画、画像認識、低遅延ライブ配信などの機能を提供していく予定です。また、機能拡張とともに、さまざまなお客さまとの共同実験を順次開始し、ユー

スペースを拡大してまいります。

6. 導入事例

株式会社レアジョブ（以下 レアジョブ）と NTT Com は本キットを活用して、オンライン英会話レッスンにおける英会話習熟度に関する実証実験（以下 本実験）を実施します。レアジョブでは、受講者の英会話習熟度の可視化に取り組んでいますが、自社システムで大量のオンライン英会話レッスンの音声データを保持し、分析を繰り返していくには限界があるなど、可視化を進めていくには課題がいくつかありました。

本実験では、オンライン英会話の音声データを、本キットによって連携されたストレージサービスに蓄積し、レアジョブの発話解析 AI システムによって分析することで、レッスン中に使用した語彙数の分析による受講者の習熟度の見える化などを実現します。

今後、受講者のフィードバックなどを得ながら、サービス品質を強化し、学習意欲向上をサポートすることを目指してさらなる検証を行います。

<発話解析 AI システムによる分析結果（開発中のイメージとなります）>



※1: WebRTC: Web Real-Time Communication の略。 HTML5 仕様の一部として、W3C, IETF によって標準化されたリアルタイムコミュニケーション機能を実現するための技術仕様。パソコンやスマートフォンのブラウザやネイティブアプリで、サーバーを介さない P2P のリアルタイム通信(音声・データ)を実現します。

※2: SkyWay は、パソコン、スマートフォン、タブレット、IoT 機器などの端末間で、WebRTC を使った音声・ビデオ通話に必要なサーバー群を NTT Com が運用し、ソフトウェア開発環境をお客さまに提供するサービスです。現在、オンラインによる英会話、診療、通訳、遠隔作業支援や、コンタクトセンター、低遅延ライブ配信などのサービス提供者に広く利用されています。

※3: NTT Com が提供する、日本語での対話が可能な AI エンジンです。NTT グループが 40 年以上にわたって蓄積・精錬した日本語辞書と AI 関連技術「corevo[®] (コレボ)」と NTT Com が独自開発した自然言語解析技術により、的確な日本語の理解による自然な対話を実現しています。また、対話だけでなく、業務システムと連携して自動で処理を行うことも可能です。



「corevo[®]」は、日本電信電話株式会社の登録商標です。 <http://www.ntt.co.jp/corevo/>



* 記載されている会社名、製品名やサービス名は各社の登録商標または商標です。

■無料トライアルの提供機能一覧

機能	詳細
Build	GUI上でマイクロサービスコンポーネントを分岐・連結することにより、音声・映像データの処理の手順を作成できます。
ビルトインコンポーネント	<p>利用頻度の高い処理は、プログラミングすることなく環境変数の設定のみで利用できるオープンソースのビルトインコンポーネントとして提供します。当初提供するのは以下の6種類です。</p> <ul style="list-style-type: none"> Filewriter : 音声を wav 形式に変換し、チャンクに区切ってオブジェクトストレージに保存 Recognizer : 「Google Speech API」を利用して音声認識を実行 Translator : Recognizer の結果を入力し「Google Translate API」により機械翻訳を実行 Logger : 動作確認用のログを出力 Dbwriter : Recognizer および Translator の結果をデータベースに保存 Publisher : Recognizer および Translator の結果を Google Cloud Pub/Sub に publish
カスタムコンポーネント	プログラミングによりコンポーネントを自由に作成できます。SDK をオープンソースソフトウェアとして提供します。当初は Node.js に対応します。
Deploy	Build で作成したシナリオをデプロイできます。テスト環境や本番環境など、それぞれのステージごとに異なる環境変数を設定できます。
Insights	各コンポーネントのログ、メトリクス、ヘルスチェックを確認できます。