

2013年9月10日

NTT アイティ株式会社

スマホやタブレット等で簡単・高精度な音声認識が利用できる クラウド型音声認識サービス「SpeechRec Cloud」を提供開始

～音声認識を活用したアプリケーション開発をより簡単に実現～

NTT アイティ株式会社（本社：横浜市中区、代表取締役社長：長谷雅彦）は、スマホやタブレット等で簡単・高精度な音声認識が利用できる「SpeechRec Cloud」を、9月24日（火）より提供開始いたします。

「SpeechRec Cloud」は、従来のオンプレミス型の「SpeechRec」を、スマホやタブレット等からより簡単にご利用いただくため、新たにクラウド型サービスとして開発したものです。従来のオンプレミス型と比べ、短期間かつ低コストでご利用いただくことが可能となります。

「SpeechRec Cloud」は、音声認識エンジンとして、WFST 音声認識技術※1）を駆使して NTT 研究所が新たに開発した高速・高精度の音声認識エンジン「VoiceRex」を搭載しています。また、背景雑音の抑圧や音声データのエンコードなど、音声認識を利用する上で必要となる機能を搭載したクライアントライブラリもご用意しています。クライアントライブラリのご利用により、高性能な音声認識を利用するアプリケーション開発を、より簡単に行うことが可能となります。

本サービスは、「Human Sensing 2013」（10月23日-25日、パシフィコ横浜）、「コールセンター/CRM デモ&コンファレンス 2013 in 東京」（11月14日-15日、池袋サンシャインシティ・コンベンションセンター）に出展いたします。

※1 WFST 音声認識技術：重み付き有限状態トランスデューサ（WFST: Weighted Finite-State Transducer）音声認識技術は、従来の個別モデル（音響モデル、単語発音モデル、言語モデル）を統合・最適化した変換モデルによる音声認識技術で、超大語彙の環境で高速かつ高精度音声認識が可能となります。

■販売開始の経緯

これまで、音声認識の活用は、大規模なコールセンタでの通話録音音声の認識分野など、オンプレミス型での利用が中心となっており、「SpeechRec」も、コールセンタや IVR などの分野に導入されてきました。

近年、スマホやタブレット等の普及により、音声認識を活用したさまざまなアプリケーションが登場し、音声認識は身近なインタフェースとして認知されるようになってきました。

今回サービスを開始する「SpeechRec Cloud」は、もっと簡単に、かつ低コストで音声認識を利用したいというお客様からのご要望にお応えし、「SpeechRec」をクラウドサービスとして提供するものです。

■「SpeechRec Cloud」の概要

「SpeechRec Cloud」の概要を図1に示します。

「SpeechRec Server」にインターネットを介して音声データを送ることにより、リアルタイムで音声認識結果のテキストデータを受け取ることができます。音声データの送信やテキストデータの受信などの処理は、API を介して簡単に行うことができます。さらに、音声認識の効果を導入検討の段階で評価することができるお試し環境もご用意しています。

■ 「SpeechRec Cloud」 の特徴

(1) 最先端の WFST 音声認識エンジンを搭載

「SpeechRec Cloud」は、最先端の WFST 音声認識技術を駆使して NTT 研究所が新たに開発した超高速・高精度の音声認識エンジン「VoiceRex」を搭載しています。これにより、数百万語に及ぶ超大語彙に対しても高速かつ高精度の音声認識が可能となります。

(2) 音声認識の利用形態に合わせた 2 種類のサービス種別を提供

リアルタイム性や利用頻度といった利用形態に合わせて、「同時接続数保証型」と「ベストエフォート型」の 2 種類のサービス種別をご提供します。

(3) 音声認識に必要な基本的な処理をまとめたクライアントライブラリを提供

音声認識を利用するためには、周囲雑音を抑圧する処理、端末のマイクの制御、音声データ送信のためのエンコード処理など複雑な処理が必要となります。これらの基本的な機能を簡単に利用できるクライアントライブラリをご提供します。

■ 「SpeechRec Cloud」 サービス種別

(1) 同時接続数保証型

- 「SpeechRec Server」のプロセス数（同時に認識処理を実行できる数）単位でご契約いただけます。
- リアルタイムの認識が必要で、常時利用が見込める用途に適しています。
- 音声認識エンジンで使用する辞書は、ご要望により専用の辞書として固有名詞等を登録することも可能です。（オプション）
- 利用例：スマホやタブレットのテキスト入力に音声認識を利用する場合、スマホのリモコンアプリを音声認識で利用する場合など。

(2) ベストエフォート型

- 「SpeechRec Server」のプロセスを複数のユーザで共有するサービスです。ユーザ数単位でご契約いただけます。
- リアルタイム性はそれほど重要でなく、1 ユーザ単位で低コストに利用したい用途に適しています。
- 音声認識エンジンで使用する辞書は共用のものを利用します。
- 利用例：会話を録音したファイルを音声認識でテキスト化したい場合など。

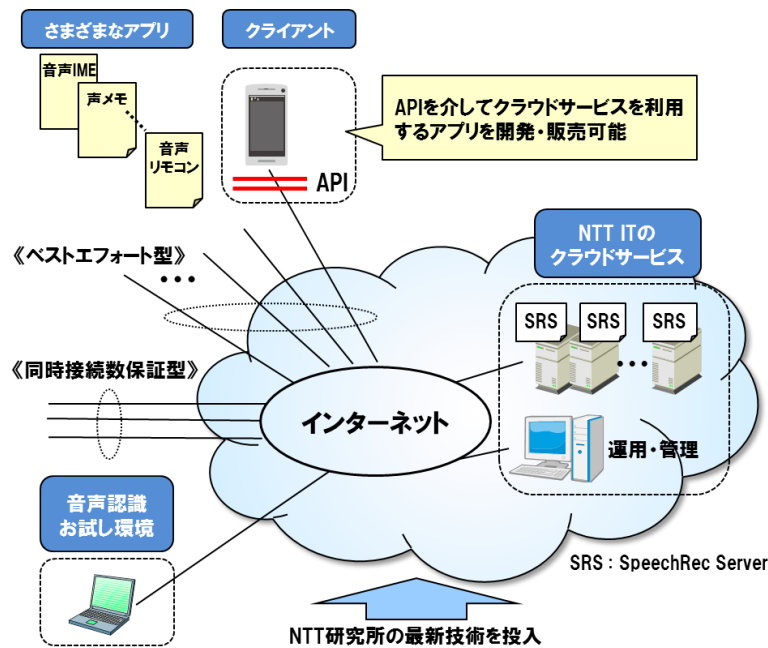


図 1 SpeechRec Cloud の概要

■クライアントライブラリの概要

図 2 にクライアントライブラリの概要を示します。Android 版のクライアントライブラリは、Activity 形式でご提供します。(API#1 を利用) 独自の音声処理機能を搭載したい場合は、SpeechRec の基本的なライブラリである「SpeechRec Client」を直接利用することも可能です。(API#2 を利用) クラウドライブラリは、今後、iOS 版や Windows 版もご提供予定です。

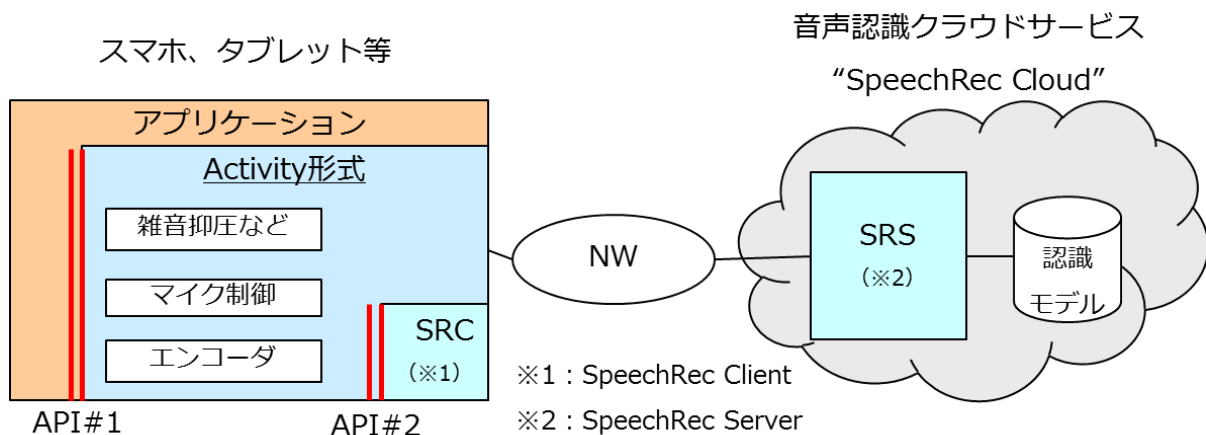


図 2 クラウドライブラリの概要

■アプリケーション例

- スマホで、メモしたいことや、かかってきた電話を録音し、音声認識した結果をテキスト情報としても残しておくことにより、備忘録や用件の確認などに活用できます。(図 3)
- ネットワーク対応型のサイネージ画面で、コンテンツの選択などの制御を音声認識で行うことができます。(図 4)

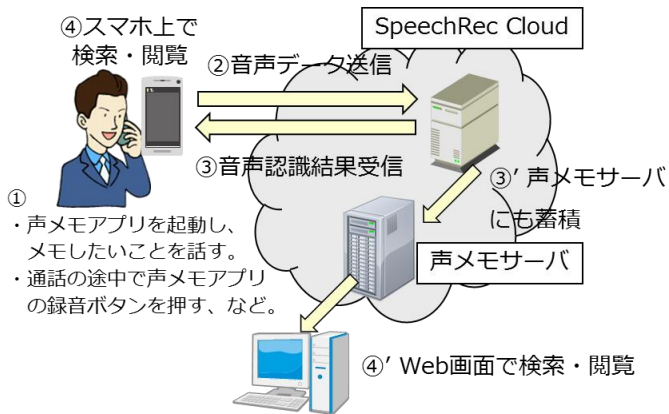


図3 アプリケーション例（声メモ）

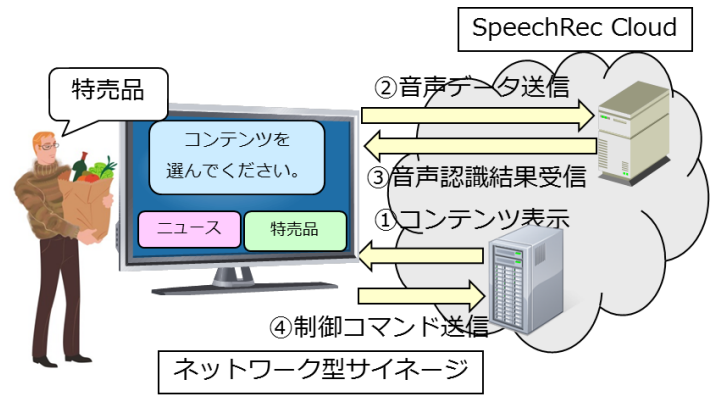


図4 アプリケーション例（サイネージ連携）

■提供形態

NTT アイティが運用する SpeechRec サーバを共同で利用するクラウド型のサービスです。

■価格

表1 クラウドサービスの価格（税抜価格。消費税等を別途加算いたします）

サービス内容		価格	備考
同時接続数 保証型	初期登録料	150,000 円	
	月額利用料 （ご利用規模、内容に 応じて個別見積）	～100,000 円/月	・1 接続あたりの月額利用料
ベスト エフォート型	初期登録料	150,000 円	
	月額基本料	30,000 円/月	・1 契約あたり月額基本料
	1 ユーザあたり 月額利用料 （100 ユーザまで）	500 円/月	・1 ユーザあたりの月額利用料 ・101 ユーザ以上は個別見積
クライアントライブラリ		個別見積	
開発サポート費（オプション）		1,200,000 円	・3 ヶ月間の開発サポート費用

お試し環境をご用意していますので、利用条件など個別にお問い合わせください。

■販売開始日と販売目標

販売開始：2013年9月24日（火）

販売目標：年間 同時接続数：100 接続、ベストエフォート型ユーザ数：5,000 ユーザ

「SpeechRec」はNTT アイティの登録商標です。
その他の商品名などは、各社の商標または登録商標です。