

2015年7月28日

WebRTCプラットフォーム「SkyWay」にて 世界初のマルチブラウザ対応音声認識 API などを無償公開

NTT コミュニケーションズ(略称：NTT Com)は、2015年7月28日より、WebRTC^{*1}技術を簡単に活用できるプラットフォーム「SkyWay^{*2}」にて、世界初のマルチブラウザ対応^{*3}の音声認識 API を無償公開します。また、WebRTC 技術によるアプリ開発をより容易とする、多人数接続・画面共有ライブラリ、WoT^{*4}サンプルコード、WebRTC セキュリティ文書も併せて無償公開します。

1.背景

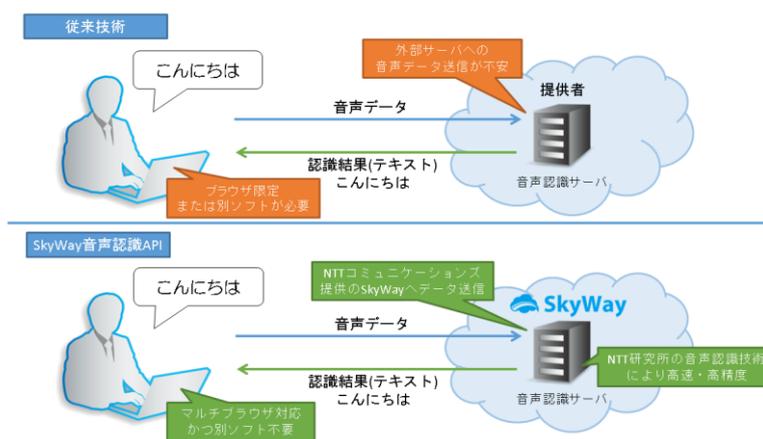
「SkyWay」は2013年12月に国内初の WebRTC プラットフォームとして公開を開始しました。これまで、開発者の方に、Web アプリケーションやスマートフォンアプリ用の開発環境を提供しており、それにより、開発登録数は2,200^{*5}を超え、様々なシーンで活用されています。

多くの開発者の方にご利用いただくなかで、音声認識 API について、マルチブラウザで利用したいというご要望を数多くいただきました。これは、従来、音声認識機能に対応できるブラウザが限られている、ブラウザの種類によってはソフトウェアを別途インストールする必要があるなどの課題があったためです。また、一部のブラウザでは、音声データが第三者の音声認識用サーバーへ外部送信されるため、その二次利用について不安の声もいただきました。

2.マルチブラウザ対応の音声認識 API について

今回、SkyWay にて無償公開するマルチブラウザ対応の音声認識 API は、JavaScript で構成されているため、利用者側のブラウザで新規ソフトウェアのインストールは不要です。また、音声認識サーバーは SkyWay 上で提供されるため、個別に音声認識エンジンを構築する必要はありません。加えて、SkyWay 上で提供する音声認識 API は、NTT 研究所の最先端の技術を活用しており、入力された自由発話音声を高速度・高精度で認識します。

【SkyWayの音声認識APIイメージ図】



(<https://github.com/nttcom/SkyWay-SpeechRec> より利用可能)

3.各種ライブラリ・WoT サンプルコード・WebRTCセキュリティ文書について

(1) 多人数接続ライブラリ (<https://github.com/nttcom/SkyWay-MultiParty> より利用可能)

SkyWay では WebRTC 技術を容易に扱えるようライブラリを提供していますが、多人数が参加するグループチャットなどを開発する場合、複雑な制御をプログラミングする必要性がありました。本ライブラリの利用により、多人数参加のグループチャットサービスを、数行程度のコードで実現できるようになります。

(2) 画面共有ライブラリ (<https://github.com/nttcom/SkyWay-ScreenShare> より利用可能)

WebRTC を利用できるブラウザは、プライバシー上の問題から、初期状態では画面共有機能が具備されておらず、個別開発が必要でした。本ライブラリの利用により、数行程度のコードで、開発者が自由に画面共有用の拡張機能(Chrome であれば Extension、Firefox であれば Add-ons)を開発できるようになります。

(3) WoT サンプルコード (<https://github.com/nttcom/SkyWay-DrivingVehicle> より利用可能)

WebRTC を活用してロボット操作を可能にするサンプルコードを、オープンソースとして公開いたします。本サンプルコードを参照することにより、ロボット操作を行う Web of Things アプリの開発が容易になります。

(4) WebRTC セキュリティ文書 (https://webrtc-security.github.io/report_ja より利用可能)

WebRTC のセキュリティに関わる全体像・詳細・関連技術などを伝える WebRTC セキュリティ文書を、オープンソースと同様の形式で公開します。本文書の活用により、WebRTC を利用したアプリケーションの利用・開発にあたり、セキュリティ上、注意すべき点を確認できます。

4. 今後の展望

今回公開する音声認識 API、各種ライブラリや WebRTC セキュリティ文書を活用していただくことで、音声認識機能を持つ新たなアプリケーションや、より安全なアプリケーションの開発などが効率的に実現できます。NTT Com は、これからもアプリケーション開発者の皆様と連携しながら、社会を変革するイノベティブかつセキュアなサービスの創出を目指していきます。

- *1 : Web Real-Time Communications の略。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）を実現するオープン技術。広義の HTML5 の一つとして、W3C と IETF で標準化が進められている。
- *2 : NTT Com が無償提供する、WebRTC 技術を活用するために開発したプラットフォーム。ソフトウェア開発支援サービス GitHub 上でライブラリ・ソースコードを公開するなど、積極的にオープンソースプロジェクトに参加している。
- *3 : 「Chrome」「Firefox」「Opera」に対応。次期マイクロソフト社製ブラウザ「Edge」にも対応予定。
- *4 : Web of Things の略。IoT(Internet of Things)アプリケーションの作成を簡単にするアプリケーションレイヤを提供する。
- *5 : 2015 年 7 月 15 日現在の API キーの発行数。