



最優秀賞

東芝デジタルソリューションズ株式会社 「コエステーション」™

独自の音声合成技術で人の声から“声の素”を生成 多様なサービス展開を実現可能に

DATA

活用領域・解決する課題	・人の声をもとにした音声合成サービスの提供 ・安全かつ自由な「声」の収集・蓄積
テクノロジー・デバイスキーワード	音声合成、AI、スマホアプリ



東芝デジタルソリューションズ
RECAIUS事業推進部
コエステーション事業推進 プロジェクトリーダー
金子祐紀氏

東芝グループでIoTやAI、ICTなどデジタル分野のビジネスを担う東芝デジタルソリューションズは、人間の声の“分身”を作り、その声で読み上げができるユニークな仕組みを生み出した。

独自の音声合成技術を活用して、声を登録、収集・蓄積し多様なサービス展開を可能にするプラットフォーム「コエステーション」である。一般ユーザー向けのスマホアプリを提供するとともに、ビジネスユーザー向けのツールも用意した。

「コエステーション」のベースに使わ

れているのは、東芝が研究所で40年以上にわたって進化させてきた音声合成技術だ。

RECAIUS事業推進部コエステーション事業推進プロジェクトリーダーの金子祐紀氏は、その特徴を次のように説明する。

「文章を読み上げた人の声からきめ細かな特徴を抽出・学習して声の素となるデータ『コエ』（以下同様）を生成し、音声合成エンジンによって人工的に作り出したその人の声でテキストデータを読み上げさせることができます。最低で10文、多くても数百程度の

文を読み上げて登録すれば、チューニングなどの必要なく、ほぼ全自動で『コエ』が出来上がります。さらに、カーナビなど組み込み機器での利用を想定し、データ量を抑え処理が軽くなるように開発されているため、ネット上で多くの声データを収集する仕組みとしても適しています」。

音声インターフェースの普及に着目 β版で確かなニーズを確認

「コエステーション」を発想したきっかけは、スマートスピーカーだった。「スピーカーからの声家族や友人、好みのタレントなどの声だったら面白い。AIやIoTの普及に伴って音声インターフェースの活用シーンが増えれば、ビジネスとしての可能性も大きく広がります。そこで、いろいろな人の声を集めて多種多様な音声合成サービスを提供できるプラットフォームを開発することにしました」と、金子氏は話す。

市場の反応を探るべく2017年7月にβ版をリリースすると、すぐに業種

コエステーション「エディター」の画面イメージ



も用途も多岐にわたる問い合わせが寄せられた。各種メディアでも相次いで取り上げられた。

まずは個人向けにスマホアプリ 自分の「コエ」を手軽に生成

確かな手応えを感じて開発を加速し、2018年4月にまずは一般ユーザー向けに「コエステーション」の名称でiPhone対応アプリの提供を開始した。

同アプリは、先述の通り、規定のセリフを読み上げることで、『コエ』を最短15~20分程度で生成し、入力したテキストをしゃべらせることができる。

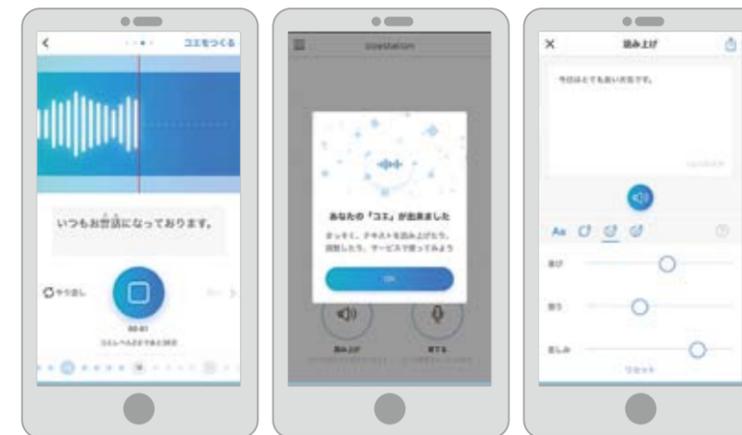
『コエ』生成のためのセリフは、10文のレベル1から200文のレベル5まで5段階が設定されており、上位レベルほど生の声に近い『コエ』を生成できる。また、『コエ』は年齢や喜怒哀楽など簡易なパラメータでチューニングを行える。

さらにSNS (LINE、Facebook、Twitter) への音声投稿機能、『コエ』を使ったゲームやスタンプサービスなど、利用促進につながる付加価値も提供している。

同時に、声の権限を守るためのセキュリティにも万全を期した。

まず、『コエ』生成に規定のセリフを指定順に読み上げる方法を採用したこと。「技術的には任意の文章でも声の特徴を学習できますが、テレビの

スマホアプリ「コエステーション」の画面例



音声などから有名人の『コエ』を勝手に作ってしまうといった不正利用を防止するため、声登録の条件を設定しました(金子氏)という。

また、『コエ』の利用に関する本人認証や他人からの利用申請、それに対する許可・拒否を選択できる仕組みを実装し、本人の意向に反した利用ができないような工夫も考えている。

さらに、すべての音声合成処理のログ保管や、生声か合成音かを識別するための信号の埋め込みなどの手立ても施している。

スマホアプリのプロモーションはほとんど行わなかったが、Webニュースや利用者のSNS投稿がきっかけで話題となり、認知度が急激に上がった。サービス開始から1か月足らずで利用者は1万人を突破し、すでに3万人以上の『コエ』が集まっているという。

ビジネス向けツールで事業本格化 多様な業種とのコラボに意欲

2018年11月には、ビジネスユーザー向けの2つの有償ツール——ブラウザ上で手軽にテキストから音声合成を作成・編集し音声ファイルを出力できる音声コンテンツ制作用の「エディター」と、テキストを音声合成に変

換する「Web API」をリリースした。いずれも、タレントや声優、有名人などを含めたさまざまな人の『コエ』を利用可能で、一般ユーザー向けアプリよりもはるかにきめ細かな合成音声のチューニングが行える。

金子氏は、「さまざまな企業とコラボレーションして『コエ』を使ったサービスを普及拡大していくことが、プラットフォームである『コエステーション』の本来の役割」と話し、多種多様な業種・業態での採用事例を積み上げていく考えだ。

孫の『コエ』で高齢者の自動安否確認をしたり、好きなタレントの『コエ』が選べるカーナビやスマートスピーカーができた、その地域の出身タレントが観光案内をしたりと、いろいろな用途が考えられる。

金子氏は、「ゲームキャラクターがプレイヤーの名前を呼んでくれる、チャットボットで何万人からの質問に同時に返答するなど、ニーズはあるものの生声では対応不可能あるいは非効率なサービスも、合成音声なら実現できるものが多々あり、アイデア次第で用途は大きく広がります。新しい音声文化を作り上げていきたい」と、意気込みを語っている。

図 「コエステーション」の概要

