

### 3 FutureVoice の音質向上・UI を刷新

## 顧客のニーズに応じたバリエーション豊かな自然な“音声”を提供し、顧客価値を向上

NTT テクノクロス株式会社（以下、NTT-TX）は NTT の研究所の最先端技術を活用して、自然な音声を作成可能にする音声合成ソリューション“FutureVoice”を提供している。お客様のニーズを汲み取る形でさまざまな音声バリエーション・利用形態への対応や、より自然な音声の実現に取り組み続け、新しい価値を提供し続けている。

#### 人が話すような自然な音声を手軽に利用可能に

音声合成は、対話型ロボットや電子書籍の読み上げ、館内放送などで利用されているが、一般には「機械的な音声」というイメージが強いのではないだろうか。NTT-TX が提供している音声合成ソフトウェア“FutureVoice Crayon（以下、FutureVoice）”を使うと、人が話すような自然な音声を業界最高レベルで実現できる。標準で様々な用途での活用を想定した性別・世代別の 50 パターン以上の音声を用意されているほか、日本語以外の言語や、嬉しい・嫌悪・悲しいといった感情を伴う口調も表現可能だ。

収録した人の声を音声合成として提供するサービスメニューも用意されている。NTT の研究所が深層学習（DNN：Deep Neural Network）を活用して開発した最先端の音声合成技術により、任意の人の少量の音声データから、音声合成を可能にし、「自然な音声」と「低コスト」という、従来は相反していたメリットが両立する。それにより、様々なニーズに応じた音声の利用を実現している。

提供形態として、サーバーシステ

ム、組み込み開発向けの SDK、オンライン連携が可能なインタフェースと Web ブラウザにテキストを入力するだけで音声を作成できるユーザインタフェース（以下、UI）を有するクラウドサービスなど、さまざまな製品ラインナップがある。

#### 人気声優の声も手軽に利用可能 声優ビジネスにも大きなメリット

なかでも特徴的なものが、声優の声を利用できる“FutureVoice Actors”だ。誰でも知っているような有名声優をはじめ、20 名以上の声優の声が用意されている。本サービスを実現・提供するにあたっては、NTT グルー



NTT テクノクロス株式会社  
カスタマーエクスペリエンス事業部  
マネージャー 鳥居 崇氏

プ会社に加え、声優事務所、音声利用権の調整を行う企業と協力している（図 1）。

人気声優は非常に忙しく、全ての依頼に応えることは不可能だ。しか



図 1 FutureVoice Actors ユースケース・サービススキーム

し FutureVoice Actors なら音声合成向けに、一回音声を収録するだけで、さまざまな場面で声を利用してもらえる。少量の音声データから音声合成できるため収録の負担も小さい。また NTT の研究所独自の「クロスリンガル音声合成技術」により、人の声質を維持し、外国語の音声合成を実現し、バイリンガルではない声優の声で多言語の吹き替えを行うといったビジネス展開も今後考えられる。声の利用者はもちろん、声優ビジネスの事業者にも大きなメリットがあるサービスとなっている。

実現にあたり技術的な問題以上に、新しいビジネスモデルならではの難しさがあったという。

「声優さんの声を利用するには複雑な権利調整が必要ですので、手間を最小化して声優さんの音声を利用できることも FutureVoice Actors の大きな特長です。ただし今までにないビジネスであったため、IT と声優という業界の違いによるすり合わせには苦労がありました」(鳥居氏)。

### 簡単に読み・アクセントを調整できる UI と、より自然な音声を実現

FutureVoice では、指定したテキストを読み上げる形で音声を作成される。テキスト中に「辛い思い出」と「カレーが辛い」のように異なる読みがある単語についても、文脈に応じて正しい読みが選択される。ただし意図した読みにならないことも有り得るため、テキストエディターのような UI で読みを変更できるようになっている。同様にアクセントが意図したものと異なる場合も、同じ画面でアクセントの位置を編集



図2 ICT デザイン室と共同で開発した新しい UI

し、調整することが可能だ。

しかし文章中から目的の箇所を探し出し、アクセントの位置を編集する作業は、直感的でなく、人によっては使いづらかった。

「IT に詳しくない現場でも使いやすいものにするため、“ここを動かす ICT デザイン室” (本特集“2 ForeSight Voice Mining 《FSVM》の UI 刷新とグローバル展開”参照) と一緒にいろいろな業界のお客様にヒアリングしました。『どこが使いづらいか』、『どうすると使いやすいか』、といったことを聞き、その内容に基づいて開発した新しい UI を、2021 年 7 月のバージョンアップで提供開始しました」(鳥居氏)。

新しい UI のサンプルを図 2 に示す。長い文章ではなくセリフ単位で編集するような UI を採用したため、「この部分だけ嬉しい調子で」といった設定をしやすい。また顧客からニーズの高かった部分的な音の上げ下げや長さを調整する機能も付加した。先進的な UI であるため、UI に関する特許を出願している。

また 7 月のバージョンアップでは、更に自然な音声を実現している。や

はり DNN を活用し NTT の研究所が開発した最先端技術によるものだ。

### お客様のニーズに応え新しい顧客価値を提供し続ける

今後もお客様のニーズを把握し、NTT-TX の特徴である「思考と試行による探索」による Explore 型のビジネス開発を続けるとして、鳥居氏は次のように述べている。

「『世の中にバーチャルアナウンサーが登場しはじめたのでそれに必要な声を揃えよう』といったように、これまでお客様との対話からニーズを拾い上げ、ラインナップや機能の追加により新しい価値を提供してきました。まだ声優さん本人による読み上げでなければ難しいという部分も多いので、より自然な音声合成についての研究開発も続きます。このほか、たとえば『音声がどのように使われたか追跡したい』というニーズがあります。少なくとも将来的に本物と見分けの付かない自然な音声を実現した際には必要になることです。こうしたことも含め、さまざまなことを検討しています。」