

# NTTの研究開発技術を融合しAIフェイク解決に挑む、VOICENCE (ヴォイセンス) カンパニー始動

生成AI (人工知能) による音声合成の普及とともに、声の無断利用やフェイク音声は社会的課題となっています。日本では法制度上、声そのものを直接保護する明文規定がなく、このギャップが新たなリスクを生む中、NTT西日本は音声AI事業「VOICENCE (ヴォイセンス)」を立ち上げました。本稿では、NTT人間情報研究所の音声処理技術とNTT西日本が独自開発した真正性証明「トラスト技術」を組み合わせた新市場創造への挑戦とその社会的意義について解説します。

## AI 音声が直面する新たな社会課題

生成AI (人工知能) の普及に伴い、「声」の無断利用やフェイク音声拡散という社会問題が深刻となっています。こうしたリスクは、著名な声優や俳優、タレントなどの実演家の声が、自身の意思とは無関係にデジタル上で拡散される“フェイク音声”問題としてマスメディアでも取り上げられています。

一方、日本の現行法制では声そのものを直接保護する明文規定が存在せず、肖像権やパブリシティ権の一部でしか対応できません。実演家の活動領域はグローバルに広がる一方で、無断生成や不正利用によるブランド毀損・機会損失が拡大している中、この制度的なギャップを放置したままでは、クリエイターや実演家が自らの声をデジタル時代において安全に活用することが困難となり、生成AIによって生まれる新たな価値とリスクのバランスが欠如してしまいます。

この課題認識を背景として、NTT西日本は2025年10月27日、音声AI事業「VOICENCE (ヴォイセンス)」\*1の開始を発表しま

した<sup>(1)</sup>。公式プラットフォームを介して“公認AI”を生み出し、声の権利保護とAI音声の健全なビジネス活用を実現し、AI音声の活用促進をめざします。

## VOICENCE とは——音声IPの保護・活用基盤

VOICENCEの核心は、「音声IP (Intellectual Property) としての声の保護と活用」にあります。

実演家本人や権利者が提供した音源データを基にAI音声モデルを生成し、そのモデル自体と生成されたAI音声コンテンツを「音声IP」\*2として統合的に管理できるプラットフォーム基盤を構築します (図1)。これにより、声のライセンス情報、生成モデル、

- \*1 VOICENCE (ヴォイセンス) : NTT西日本が開始した音声AI事業。名称は「Voice (声)」「License (権利)」「Essence (本質)」を組み合わせた造語。
- \*2 音声IP : 声を知的所有としてとらえ、話者本人の同意に基づき権利化・管理・活用する概念。音源データ、AI音声モデルおよび生成した音声コンテンツを含みます。



図1 VOICENCEプラットフォーム

コンテンツ利用権が一元的に整理され、実演家本人により公認されたAI音声である真正性と利用履歴が追跡可能となります。

VOICENCEという名称は、Voice（声）、License（権利）、Essence（本質）の3つの意味を含めた造語であり、声の価値を守りつつその本質的な可能性を拡張する事業であることを示しています<sup>(2)</sup>。

### 技術的基盤——合成音声と音声印象制御技術

VOICENCEの合成音声は、NTT人間情報研究所で研究・開発されてきた技術を基盤としており、少量の音声データから話者の声質や口調の特徴を抽出し、本人らしい音声を高精度に生成する技術が用いられています。

本事業では、NTT人間情報研究所の研究成果として、Zero/Few-shot<sup>\*3</sup>およびクロスリンガル音声合成技術<sup>\*4</sup>を活用し、日本語話者モデルから英語、中国語、韓国語、フランス語、スペイン語といった複数言語の音声を、声質を保ったまま生成することが可能となっています<sup>(3)</sup>。

さらに、同研究所が開発した音声印象制御技術<sup>\*5</sup>を実装している点も特徴です(図2)。声質を維持したまま、抑揚や話し方のニュアンス、口調の強弱といった表現特性を制御することで、同一話者モデルから用途に応じた複数の発話表現を生成できます。

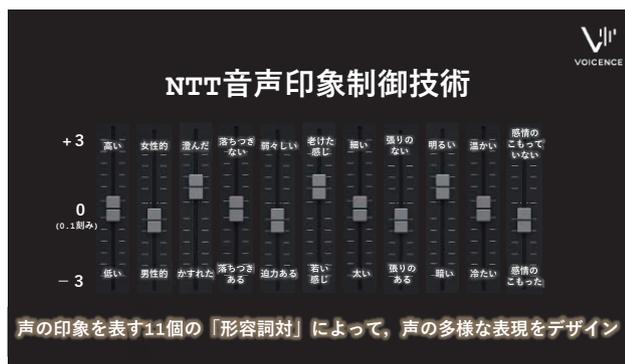


図2 NTT音声印象制御技術

これらの技術により、VOICENCEは単なる音声生成にとどまらず、話者の個性を保持しつつ多言語・多用途に適応できる「コンテンツプロデュース事業」を展開しています。

### トラスト技術——NTT西日本開発の真正性証明技術

VOICENCEの最大の技術的な特色は、NTT西日本が独自開発した「トラスト技術」<sup>\*6</sup>と呼ばれる真正性証明機能<sup>\*7</sup>の実装です。単に音声を生成するだけでなく、それが「誰の声であるか」「どのモデルから生成されたか」「どのような契約条件で使われるか」を、ブロックチェーンとVC (Verifiable Credentials: 検証可能資格情報)<sup>\*8</sup>によって記録・管理する仕組みを持ちます(図3)。

この仕組みは次の3層構造で提供されます。

- ① 改ざん耐性のある分散台帳(ブロックチェーン)による真正性情報の保管
- ② 階層型VCによるコンテンツとモデル、音声データとの関係性の記録
- ③ 利用条件・契約情報(用途証明)の付与

これにより、無断生成された類似音声データとVOICENCE正規生成音声を識別可能にし、利用者企業・実演家双方が安心して活用できる信頼基盤を提供します。

また、本技術は2025年度NEDO(国立研究開発法人新エネルギー・産業技術総合開発機構)懸賞金活用型プログラム「GENI-AC—PRIZE」領域03のトライアル審査で採択されるなど、外部評価も得ています。

### 権利保護から流通までを一貫して担い“声の経済圏”を創出

NTTは電話からはじまり、「声」を大切にしてきました。そのDNAと積み重ねてきた国産技術を活かし、“声の経済圏”を生み出すことで、ジャパンコンテンツ産業の中にある「声」という資産が新たな価値を生み出すことを推進し、グローバル展開を含む

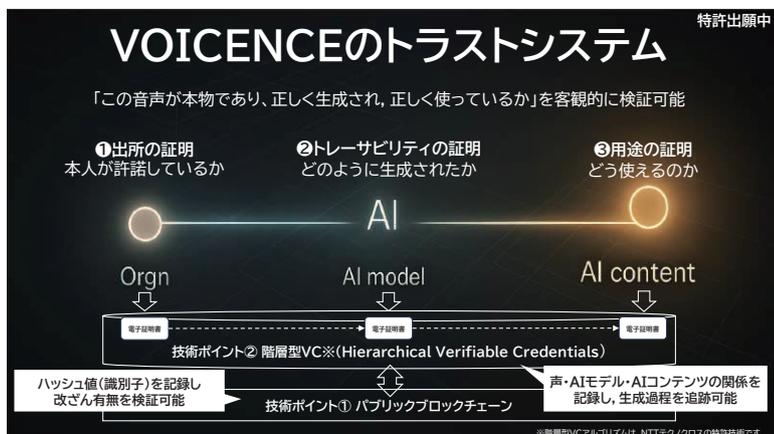


図3 VOICENCEのトラストシステム

- \*3 Zero/Few-shot: 少量または数秒程度の音声データから話者の声の特徴を学習し、新たな発話をその話者らしい音声として生成する技術。
- \*4 クロスリンガル音声合成: 単一の話者モデルを用いて、声質を保持したまま複数言語の音声を生成する技術。多言語対応の音声対話や国際展開用途に用いられます。
- \*5 音声印象制御技術: 話者の声質を保持したまま、抑揚、話速、口調、感情表現などのパラ言語的特徴を制御する音声生成技術。用途に応じた発話表現の生成が可能。
- \*6 トラスト技術: 生成データの真正性や利用条件を暗号的に証明・管理するための技術群。ブロックチェーンやVCなどを使用。
- \*7 真正性証明: 生成音声が正規の話者モデルおよび契約条件に基づいて作成されたものであることを技術的に証明する仕組み。改ざん防止および追跡可能性を担保。
- \*8 VC: 発行者・保持者・検証者の三者モデルに基づき、データの真正性や属性を暗号的に検証可能とするデジタル証明方式。



図4 連携するNTTグループ各社との集合写真

さらなる成長へ貢献したいと考えています。音声IPの真正性・契約情報・生成技術が統合されることで、企業はブランドボイスやナレーション、広告音声などを安心して活用でき、実演家は自らの声を安全に収益化できる市場環境が整います。

いち早く市場を創造していくためにVOICEENCEカンパニーという社内独立組織を立ち上げ、機動的な事業推進体制を整えます。これにより、ライツマネジメント機能と音声コンテンツ企画・制作機能を備えたコンテンツプロデュース事業を行い、実演家と企業やブランドの双方に対して、権利保護から流通まで一貫したサービスを提供します。

## VOICEENCEがめざす姿

VOICEENCEがめざすのは、「声の真正性が担保されたAI音声コンテンツ」のデファクトスタンダードを確立することです。トラスト技術によって付与される証明データや利用ログを、ステークホルダーの要請に応じて提供することで、誰もが正当に声の権利を行使できる社会基盤の整備を進めていきます。さらに、音声コンテンツの保護に関する事例を継続的に蓄積し、社会的な抑止力を育むとともに、市場における「声の権利」への認識定着を図ります。

併せて、声の権利研究をリードするNTT社会情報研究所と連携し、同研究所が創出した研究成果である「声の権利を保護し適切に扱うための解説資料」を活用することで、法的・社会的観点からの権利保護の普及を推進します。また、知的財産権を専門とする法律家、行政・業界団体、アカデミアなどの有識者と連携しながら、ルール整備と運用基盤の構築を段階的に進めていきます。

VOICEENCEカンパニーは、こうしたルールに準拠したサービスを提供することで、AI音声の信頼基盤を社会に実装します。これにより、実演家、企業、クリエイターが、より自由かつ安全にAI音声ビジネスへ参入できる環境を整え、「声の経済圏」の創出をめざします(図4)。

VOICEENCEカンパニーはNTTグループ各社と協力してビジネ

イベント名：The Unknown City ～時を超える広島の記憶とスパイミッション～  
開催期間：2026年1月23日(金)～3月29日(日)  
開催場所：広島市中央公園エリア付近  
参加費：無料  
参加方法：以下記載のQRコードを読み取って、ARG型謎解きデジタルラリーの導入体験や、最新情報の取得を行います。



<https://www.ntt-west.co.jp/news/2601/260116a.html>



図5 NTTアーバンソリューションズ主催イベントでのAI音声を活用したコンテンツ事例

図5 NTTアーバンソリューションズ主催イベントでのAI音声を活用したコンテンツ事例

ス活用に注力していきます。

- ・NTT人間情報研究所：合成音声
- ・NTT社会情報研究所：声の権利保護
- ・NTTテクノクロス：AI音声モデル制作、AI音声コンテンツ制作、トラストシステムへの技術提供
- ・NTT ExCパートナー：AI音声コンテンツ制作
- ・NTTアーバンソリューションズ：AI音声を活用した施策オーナー

### ■コンテンツ事例

AI音声を活用したコンテンツ事例として、NTTアーバンソリューションズ主催のイベントを紹介します(図5)。概要は下記のとおりです。

- ・イベント名：The Unknown City ～時を超える広島の記憶とスパイミッション～
- ・開催期間：2026年1月23日(金)～3月29日(日)
- ・開催場所：広島市中央公園エリア付近
- ・参加費：無料
- ・参加方法：公式ティザーサイト ([https://www.hiroshima-theunknowncity.com/trial?utm\\_source=pr](https://www.hiroshima-theunknowncity.com/trial?utm_source=pr)) より、ARG型謎解きデジタルラリーの導入体験や、最新情報の取得を行います。

### ■参考文献

- (1) <https://www.ntt-west.co.jp/news/2510/251027b.html>
- (2) <https://voicence.jp/>
- (3) [https://www.rd.ntt/research/JN202504\\_33343.html](https://www.rd.ntt/research/JN202504_33343.html)

### ◆問い合わせ先

NTT西日本  
VOICEENCEカンパニー