

# 生成AIと声が変わるとき

## — AI音声の社会実装とPublic Engagementの実践 —



特集  
ハイテク推進  
セミナー

株式会社 CoeFont

Vice President 山田 泰裕 氏

AI Voice and Social Implementation in the Era of Generative AI

Key Words: Voice Ethics, Public Engagement, Social License

### 1. はじめに

生成AIの急速な普及により、私たちの日常生活は大きく変わりつつある。文章生成や画像生成に続き、近年とくに注目を集めている分野が「AI音声」である。声は単なるデータではなく、人の感情や記憶、さらには人格そのものと結びつく媒体であるため、その社会実装は技術的課題に留まらず、文化・倫理・制度を巻き込んだ総合的テーマとして議論される必要がある。AIによる音声の生成・変換が一般化するほど、社会は「声とは何か」という根源的な問いに向き合うことになる。本稿では、AI音声プラットフォーム「CoeFont」の事例を用いて、生成AI時代におけるPublic Engagementの実践を論じ、先端技術を社会へと浸透させるための条件を考察する。

### 2. 技術ではなく“声”から始まったアプローチ

株式会社CoeFontは2020年に創業した。深層学習にもとづく音声合成・変換技術を背景にしながらも、その出発点には創業者である大学生の「自分の声に自信が持てない」という素朴な感情があった。声を

“文字のフォントのように誰でも選べる”世界をつくるという着想は、一般的な技術起点とは異なる、人間の感情に由来する視点である。この発想は単なるUI/UXの改善に留まらず、声を個人の社会参加や表現の基盤と捉え、人と声の関係そのものを問い直すアプローチへと発展した。

だが、当初のAI音声には大きな課題があった。社会一般はAIへの不信感を抱えており、「AIが声を奪う」「偽物の声が氾濫する」といった漠然とした恐怖と不安が先行していた。技術そのものの精度向上だけでは、この溝を埋めることはできなかった。CoeFontが直面したのは、技術開発と社会受容の間に横たわる深い断絶であり、これを乗り越えるには、人間の尊厳や文化的文脈に配慮したPublic Engagementが不可欠であった。

### 3. 倫理から始める社会実装

CoeFontの社会実装が大きく前進した契機は、一通の手紙であった。咽頭がんで声帯を摘出する患者から、「手術の前にどうしても自分の声を残したい」という切実な依頼が届いたのである。これは単なる技術的相談ではなく、「声とは何か」「声を失うとはどういう経験なのか」という根源的な問いを突きつける出来事であった。

この要望に応えるかどうかの意思決定は、即座に行われた。その背景には社会的理解も十分ではない状況を打破するには社会的なインフラとなる必要があると確信していたからである。まずは声帯摘出者に、続いてALS患者に対するAI音声の無償提供を決断した。これは事業戦略というより、倫理的な立脚点を明確に社会へ示す行為であった。この取り組みを通じて、AI音声は「仕事を奪う技術」ではなく、「人の尊厳を守る技術」として新たな文脈で理解されるようになった。また、医療機関、福祉関係者、行政などとの連携が進み、AI音声が生活の質(QOL)



講師 山田 泰裕 氏

向上に貢献する具体的な姿が社会に共有されていた。

この過程は、公的な許可や法制度ではない“社会的許容”(social licence to operate)を獲得するプロセスであり、Public Engagementの基礎を形成するものであった。技術への信頼は、倫理的な行動の積み重ねを通じて初めて得られるという事実を示す事例でもある。

#### 4. 社会的理解を広げるための文化的アプローチ

AI技術の普及には、専門家や行政との対話だけでは不十分である。一般の人々が自然に体験し、関心を持つ機会がなければ、技術は「自分とは無関係のもの」として扱われてしまう。そこでCoeFontは、エンターテインメントの形を借りてAI音声为社会に解き放つ戦略を採用した。

代表的な試みが「おしゃべりひろゆきメーカー」である。著名人の声をAIで再現し、ユーザーが任意の文章を入力すると、ひろゆき氏がその内容を読み上げる動画が生成される。この企画はSNSを中心に爆発的な広がりを見せ、数日間にわたってトレンド上位を占めた。技術的には既存のモデルを応用した比較的シンプルな体験であるが、AI音声は「難しい技術」から「誰でも使える身近なツール」へと変貌する契機となった。

文化的浸透には、日常の中で触れられる文脈が必要だということが、この事例から読み取れる。AI音声はゲームやSNSといった文化圏の中で用いられることで、理解が感覚的に、そして自発的に獲得されていった。技術そのものを説明するより、実際に触れた人が自分の感覚で理解するプロセスのほうがはるかに強い説得力を持つことを示す重要な実践である。

ただし、著名人の声を扱う以上、倫理とリスク管理は欠かせない。CoeFontは同時に不正利用の防止、契約・ガイドラインの整備、検知アルゴリズムの強化などを行い、「楽しさ」と「安全性」を両立するための仕組みを整えていった。文化的普及と倫理的配慮は二律背反ではなく、両者を同時に進めることが技術の持続的な社会受容につながる。

#### 5. 法制度の空白を埋める試み

AI音声の社会実装が進むにつれ、「声の権利」をめ

ぐる新しい課題が浮上した。声は肖像権や著作権と部分的に重なるが、完全に保護する法律は整備されていない。そのため声優やナレーターなど、声を仕事にする人々にとってAI音声は大きな不安を伴う存在でもあった。

この問題に対してCoeFontは、声の権利保護を目的とした「ボイス・ライト・プロテクション」という自主的な基準を策定した。ここには、本人同意の徹底、公開・非公開の選択権、収益配分の透明性、利用用途の制御などが含まれている。このような基準を企業自らが設けることは、法制度が追いついていない領域において重要な役割を果たす。業界関係者からの信頼を獲得し、声優事務所やアーティストとの協業が進んだことも、制度の空白を埋める実践として評価できる。

さらにCoeFontは、行政や政党との意見交換会にも招かれ、AI音声の活用とルール形成に関する議論に参加した。企業として制度形成に関与する姿勢は、Public Engagementの新たな形であり、技術の健全な発展を支える基盤づくりでもある。

#### 6. CoeFontが示すPublic Engagementモデル

以上の取り組みを俯瞰すると、CoeFontの社会実装は三つの層が相互に作用する構造として整理できる。第一に、声を失う人に寄り添う取り組みに象徴される倫理的基盤の構築がある。技術の利用目的が「誰かの生活を支える」方向で明確に語られることで、社会からの信頼が形成されていった。第二に、エンターテインメントを通じた文化的普及がある。AI音声は日常の文脈の中で自然に体験されるようになることで、技術に対する敷居が下がり、大衆的理解が獲得された。第三に、自主基準の制定や行政との対話に代表される制度的整合がある。技術が無秩序に広がるのではなく、社会の価値観と調和しながら発展するための制度的枠組みが整備されつつある。

これらは直列的なプロセスではなく、同時並行的に進行することで相互補完的に機能した点に特徴がある。倫理、文化、制度が三位一体となって技術の社会実装を支えるという構造は、生成AI全般にとって有効なモデルとなり得る。

#### 7. おわりに

AI技術が進化し続ける時代において、重要な

は技術そのもの以上に、「社会が何を大切にするか」を見据えながら実装を進める姿勢である。声は、単なるデータではなく、人間の感情や記憶を伴う極めて個人的な存在である。だからこそAIが声を扱う際には、文化や倫理、制度との調和を慎重に設計する必要がある。

CoeFontの取り組みは、AI音声という新領域におけるPublic Engagementの実践例として、先端技術の社会実装がどのようになされるべきかを示す重要な事例となっている。倫理的立脚点から始まり、文

化的普及を経て、制度形成へと接続する一連の営みは、AIと社会の関係を新たに描き直す試みでもある。

AIと声が変わるとき、社会は「人間らしさとは何か」という問いに改めて向き合うことになる。その問いにどう向き合い、どのような未来を形づくるのかは、技術者だけではなく、私たち社会全体に委ねられている。技術の進展を恐れるのではなく、人間の価値を中心に据えながら、AIとの共生を積極的に構想していくことが求められている。

