

AIによる最先端音声合成 ソリューションのご紹介

2021年1月12日

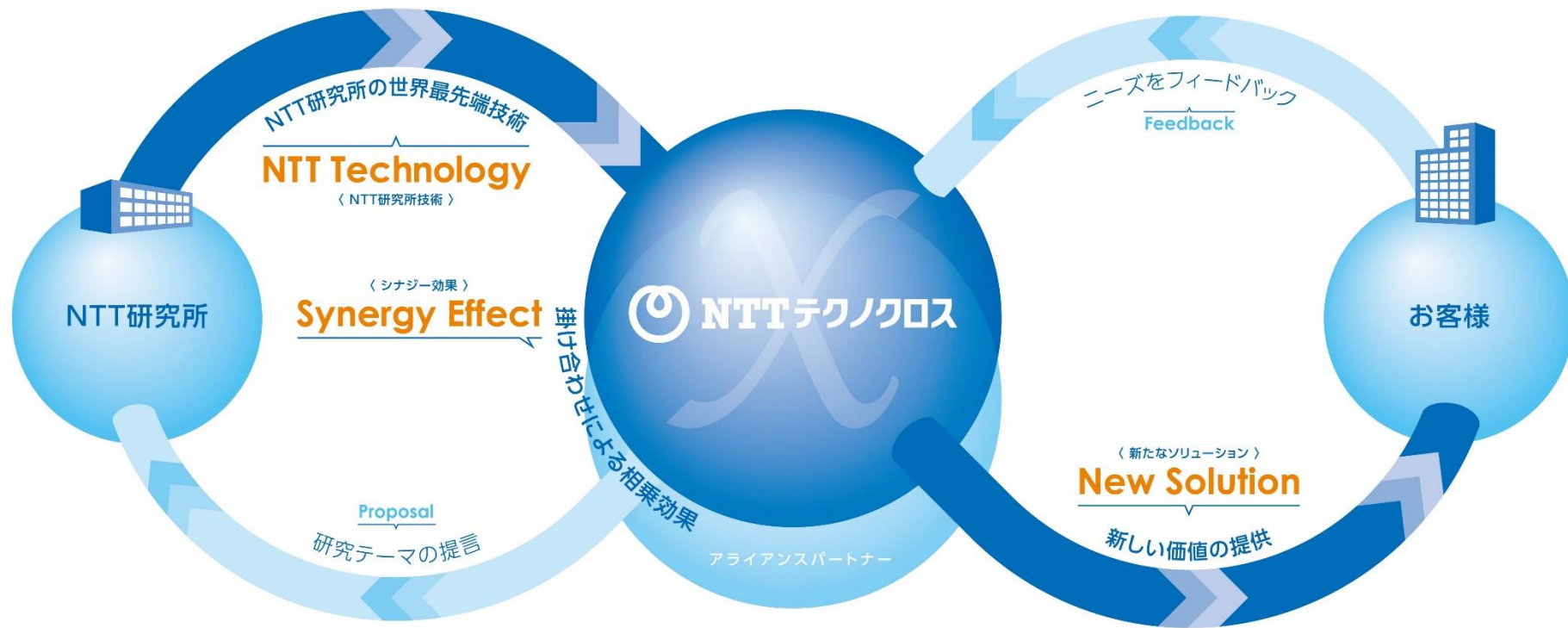
メディアイノベーション事業部

第二ビジネスユニット

マネージャ 鳥居 崇

NTTテクノクロス株式会社紹介

NTT研究所の世界に誇る最先端技術を、皆さまのビジネスに活用できるように、ソリューション化することが私たちの使命です。
NTT研究所に最も近い会社として、先進技術の研究試作開発を受けつつ、そこで培ったノウハウを活かして、皆さまに役立つソリューションを創出してまいります。



本日の内容

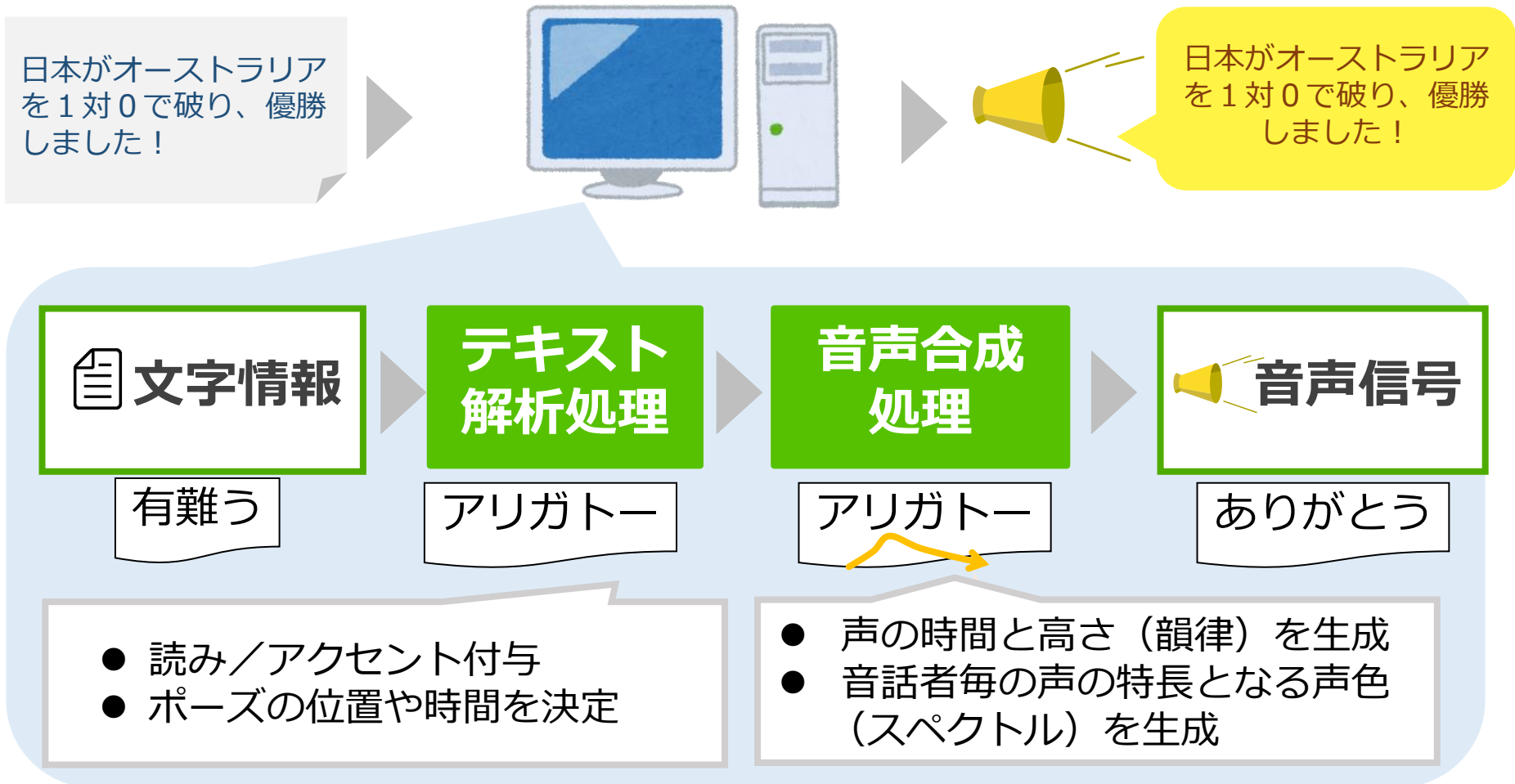
| | | |
|---|---------------|-------|
| 1 | 音声合成とは | 3min |
| 2 | DNN (AI) 音声合成 | 5min |
| 3 | 最新事例【1】～【5】 | 15min |
| 4 | 音声合成導入の課題 | 3min |
| 5 | 今後の展開と課題 | 3min |

本日の内容

| | | |
|---|---------------|-------|
| 1 | 音声合成とは | 3min |
| 2 | DNN (AI) 音声合成 | 5min |
| 3 | 最新事例【1】～【5】 | 15min |
| 4 | 音声合成導入の課題 | 3min |
| 5 | 今後の展開と課題 | 3min |

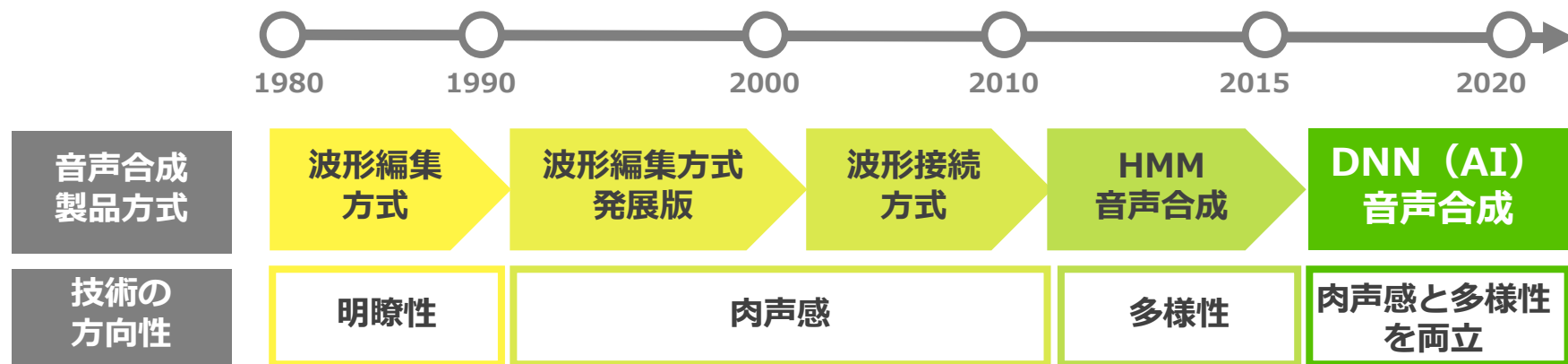
1 音声合成とは : 概要

計算機上で文字を音声に変換するメディア変換



1 音声合成とは : 技術発展・適用分野

NTTの音声合成技術発展の歴史



拡大した適用分野

従来の電話サービスの領域を超え、様々なサービスに活用が拡大

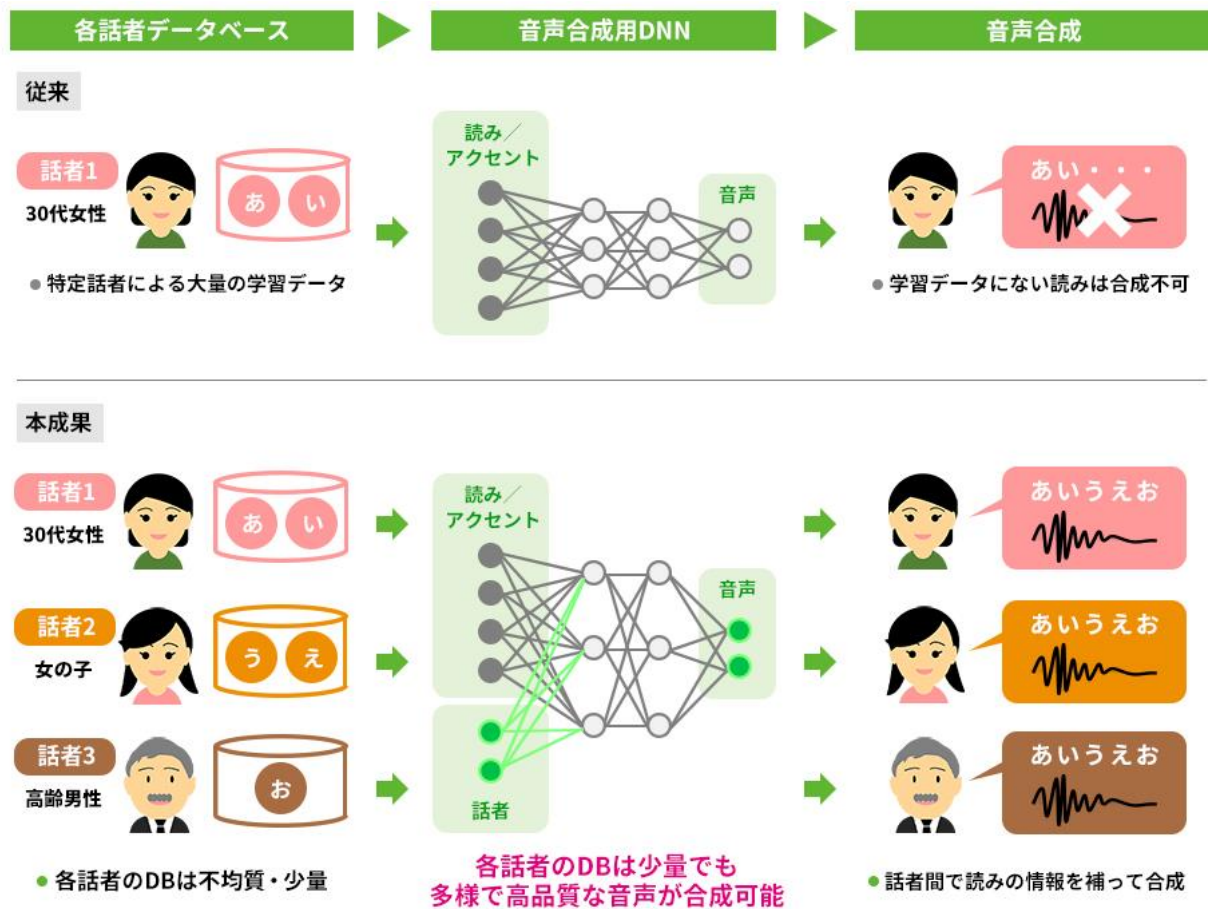


本日の内容

| | | |
|---|---------------|-------|
| 1 | 音声合成とは | 3min |
| 2 | DNN (AI) 音声合成 | 5min |
| 3 | 最新事例【1】～【5】 | 15min |
| 4 | 音声合成導入の課題 | 3min |
| 5 | 今後の展開と課題 | 3min |

2 DNN (AI) 音声合成 : 概要

再現する話者の音声が少量でも、膨大な音声データを統合して学習に活用することにより、**極めて類似性の高い音声合成**を実現



2 DNN (AI) 音声合成 : ボイス (声) 作成

再現する声の確定

- 声のイメージのヒアリング
- 発声者(話者)の選定(確定済みの場合は不要)



収録スケジュール調整

- 発声者(話者)とスタジオのスケジュール調整
- 収録用原稿の準備(声のイメージに合う原稿がベター)



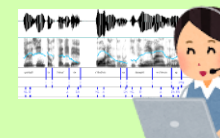
音声収録

- 収録回数: 1、2回
- 収録時間: 2～3時間/回



音声データ整備作業

- 音声切出し・アノテーション作業・ラベリング作業
- 作業期間: 2～6週間



DNNによる機械学習

- GPUマシンを利用して機械学習
- 学習期間: 数日～1週間程度



本日の内容

| | | |
|---|---------------|-------|
| 1 | 音声合成とは | 3min |
| 2 | DNN (AI) 音声合成 | 5min |
| 3 | 最新事例【1】～【5】 | 15min |
| 4 | 音声合成導入の課題 | 3min |
| 5 | 今後の展開と課題 | 3min |

3 最新事例【1】 コミュニケーションロボット①

概要

- アンドロイドロボット
- ボディは、黒柳徹子さんを3Dカメラでスキャンして、等身大の型を作成、表情やしぐさ、癖なども研究し尽くし再現
- 「totto」の声は、テレビ朝日が保有する42年にわたる「徹子の部屋」の会話データをもとに、最新の音声合成技術を用いて作成、徹子さんご本人のような声を再現



totto制作委員会様 (totto)

デモ動画：

<https://www.youtube.com/watch?v=LvmGeJl5yjA&t=6s>

導入効果

- 特長ある本人の声を最大限に生かし、音声合成により再現し、最新の音声対話技術を組み合わせることで、徹子さんのキャラクター性が活かされた自律的な会話を実現
- 既に収録ができない方、収録スケジュールが組めない方についても音声収録を行わずして、音声合成として声を再現

3 最新事例【2】 コミュニケーションロボット②

概要

- ロボット型の通信端末（コミュニケーションロボット）向けのクラウド型ロボットプラットフォームサービス
- **コミュニケーションロボットの声**として利用

導入効果

- 企業独自のオリジナルロボットに対し、**音声収録が不要になり、会話内容を自由自在に作成**する事が可能
- **オリジナルロボットの声向けに所望の話者**（発声者）を低コストで、自然かつ本人に近い音声合成を実現



NTT東日本様（ロボコネクト）
URL : https://business.ntt-east.co.jp/service/roboconnect/download/roboconnect_character.pdf

3 最新事例【3】 バーチャルアナウンサー①

概要

- ニュースの内容に応じて、通常・明るい・暗いなど**感情を含めた読み上げ**
- 複数人の女性アナウンサーの声から**新たな声**
- 発話タイミングと**CGアナウンサーの唇の動きを連動**し、自然な発話



テレビ朝日様（花里ゆいな）

デモ動画：

<https://www.youtube.com/watch?v=GWJ45dHhQhM>

導入効果

- **働き方改革の時代**にあわせて、深夜帯やボリュームのあるインターネットニュース等の番組に出演することで、簡易に「顔出しニュース」が作成できる為、**コンテンツ制作の効率化を実現**すると共に、**コンテンツ演出の向上**にも貢献
- 多忙なアナウンサーの方々のサポートとして**業務負担を軽減**
- **緊急時放送**にも強みがあり、深夜帯の突発的なニュースや災害時の迅速かつ正確な報道での活用

3 最新事例【4】 バーチャルアナウンサー②

概要

- **Vチューバーと音声合成技術**を組み合わせ、放送局独自でバーチャルアナウンサーを製作し、**番組で実際に運用**
- 音声合成技術以外にも音声波形編集ツールを駆使することで、更にリアルな発声を実現、またモーションキャプチャーシステムも独自開発することで**自然なキャラクターを創造**



RKB毎日放送様（百道桃）
URL : <https://rkb.jp/e-sekai/>

導入効果

- 番組費圧縮や働き方改革によるアナウンサー稼働可能時間の制限等に対し、バーチャルアナウンサーの出演で、**ロケ費用の削減とアナウンサー稼働の最小化を実現**

3 最新事例【5】 プロ声優の合成音声サービス

FutureVoice Actors



国内最大級

約30種類（※）の人気声優・現役プロ声優の「声」を利用可能

※2020年1月12日現在

ハイクオリティ

最先端の音声合成技術を活用し、声優の「声」を忠実に再現

24時間×365日

声優のスケジュールに関係なく「声」をいつでもどこでも作成

かんたん契約

契約時の権利調整手続きの期間や手間を大幅に削減

「声」の提供

音声合成「声」モデル

音声利用権利調整

音声合成クラウドサービス



本日の内容

| | | |
|---|---------------|-------|
| 1 | 音声合成とは | 3min |
| 2 | DNN (AI) 音声合成 | 5min |
| 3 | 最新事例【1】～【5】 | 15min |
| 4 | 音声合成導入の課題 | 3min |
| 5 | 今後の展開と課題 | 3min |

4 音声合成導入の課題 : 注意点編

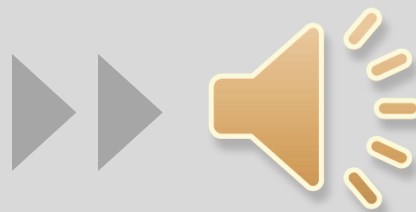
まず、テキストに注意しましょう

注意

1. 新語、特殊な専門用語、きらきらネームを含むテキスト
2. 読みを間違えないようにと、平仮名ばかりのテキスト
3. 文章の校正がイマイチなテキスト（句読点の入れ方等）

【例文】 上記のような注意すべきテキストが入っていた場合

紅葉君がDXとは企業が避けて通れない
でじたるぎじゅつによるぎょうむや
ビジネスのへんかくと言っていました。



サンプル音声

4 音声合成導入の課題 : 解決編

テキストの校正を整えるだけでも聴き易くなります！

【例文】 読点を入れ、読みづらい単語はひらがなに、一般的な単語の平仮名は漢字に

めーぷる君がデジタルトランスフォー
メーションとは、企業が避けて通れな
い、デジタル技術による業務や
ビジネスの**変革**と言っていました。



サンプル音声

実際の運用では、**テキストの校正を整える**だけではなく、**正しい読み方にするための辞書やチューニング機能を****活用**した運用が多いです。

4 音声合成導入の課題 : チェック編

導入前には、下記の点を確認してください。

- 音声合成の必要があるのか、**収録で十分**ではないのか？
- 音声合成のクオリティで**サービスレベル**は問題ないのか？
- 読みやイントネーションを正しくする運用は必要なのか？
必要な場合の**運用方法**をどうするのか？

本日の内容

| | | |
|---|---------------|-------|
| 1 | 音声合成とは | 3min |
| 2 | DNN (AI) 音声合成 | 5min |
| 3 | 最新事例【1】～【5】 | 15min |
| 4 | 音声合成導入の課題 | 3min |
| 5 | 今後の展開と課題 | 3min |

5 今後の展開と課題

- 音声合成技術の技術の進化は速く、より少量で人の声を再現し、テキスト情報を理解し、リアルな声を表現するようになるのも遠い将来ではなさそうです。
- 声を必要とするサービス・業務は数多く存在し、よりこの技術及びビジネスが広まる見込みです。
- 技術の進化に伴い、何等かの問題が生じるのは世の常ですが、それまでに解決すべき課題があります。

課題

- 音声合成利用に関する犯罪及び不正利用防止の技術と法整備
- 声優及びアナウンサーとキャラクタービジネスの共存