

# 音声認識研究内容紹介

---

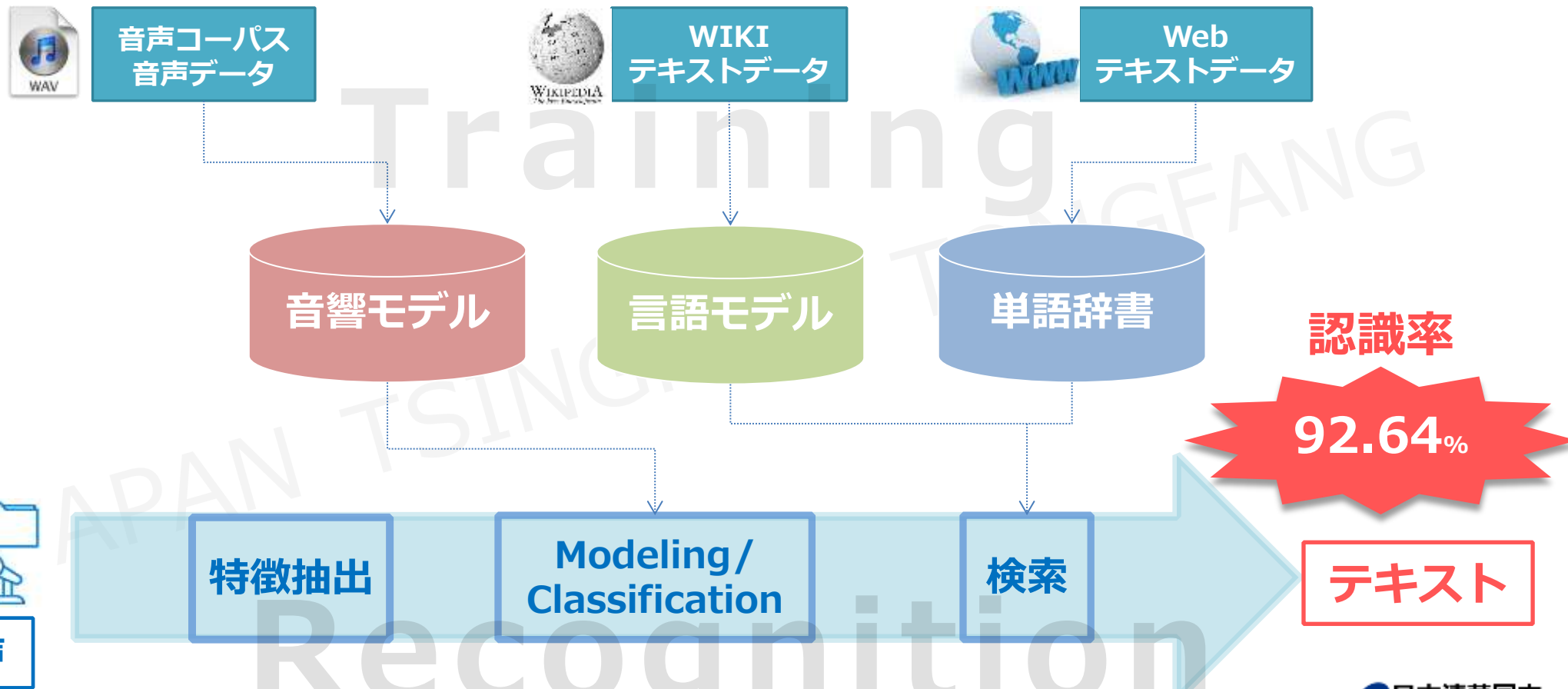


# 目次

---

- 音声認識の全体構成図
- 属性説明
- IBM Watsonとの比較
- DEMO構成

# 音声認識の全体構成図



# 属性説明

音響モデル	音声ドメイン	DeepLearning技術	モデルサイズ
AM	講演	○	76 MB

言語モデル	学習用データ		辞書	LMサイズ (MB)
	データ名	ドメイン		
LM	WEB 07	—	vocab	457
	WIKI	百科事典		

辞書	単語の数	選定基準
vocab	89556	WIKIに出現頻度が高い順（総数約9万まで）

デコーディング	サイズ (GB)	スピード
decode	4.3	0.41

# IBMワトソン(Speech to Text)との比較

	IBM Watson(Speech to Text)	自主研究開発
開発元	IBM 長期蓄積	清華同方研究新規開発
プラットフォーム	IBM専用、非公開	Kaldi オープンソース
日本語音声認識	89.67% ※1 ※2	92.64% ※3
A Mモデル訓練用データ資源量	数万時間	数百時間(研究センターに)
訓練環境と時間	高性能のマシン数百~数千台で数日	Tesla GPU 1枚のマシン1台で 3週間
Knowledge Transfer技法	なし	中国語の音声認識ニューラルネットよりKT ※4

※1 IBMの公開したサービスAPIでWEBアプリをバッチ処理で音声ファイルwavをIBMサーバにアップして識別した結果

※2 IBM公式発表による音声認識率は93.1%、なおMSは94.1%(2016年9月)いずれもテスト集は音声評価用基準SWBベース

※3 IBM Watsonサービスで、テスト対象はアプリPCM録音で録音したファイル

※4 KT(Knowledge Transfer) は中国で特許提出済、現在は中国特許庁で審査中

# IBMワトソン(Speech to Text)との比較

	IBM Watson(Speech to Text)	自主研究開発
LMモデル訓練用データ資源量	不詳(無制限)	20GB(自主採取、カスタマイズ可) 訓練にngram
言語辞書	不明	Unidic (Mecab)+ドメイン単語
音声クラウドサービス運用	IBMクラウドにアップロード	お客様の自由
騒音対応	あり	あり(kaldiベースとKTによる)
サービス(I/F)	サービス公開 ※5 Watson Developer Cloud SDK公開	WebService (Socket/NodeJS) 非公開、受注開放可

※5 このサービスの有効期限内や音声量の制限値以内は無料、一定の音声認識量や試用時間を超えると有料

# IBMワトソンとの音声認識実例の比較 1

## ● 音声ラベリング



ECO\_0150

**音楽コンサート**や舞台など演者と観客が同じ空間で生の体験を共有できるライブエンターテイメントが日本で今活況を呈している

## ● 音声認識結果

IBM-Watson

**今後今想定**や舞台など演者と観客が同じ空間で生の体験を共有できるライブエンターテイメントが日本で今活況を呈している

現在の研究結果

**音楽コンサート**や舞台など演者と観客が同じ空間で生の体験を共有できるライブエンターテイメントが日本で今活況を呈している

# IBMワトソンとの音声認識実例の比較 2

## ● 音声ラベリング



ECO\_0159

**総務省**が都市部の若者が長期休暇を利用して地方で働くふるさとワーキングホリデー制度を二千十七年にスタートするとマスコミ各社が報道

## ● 音声認識結果

IBM-Watson

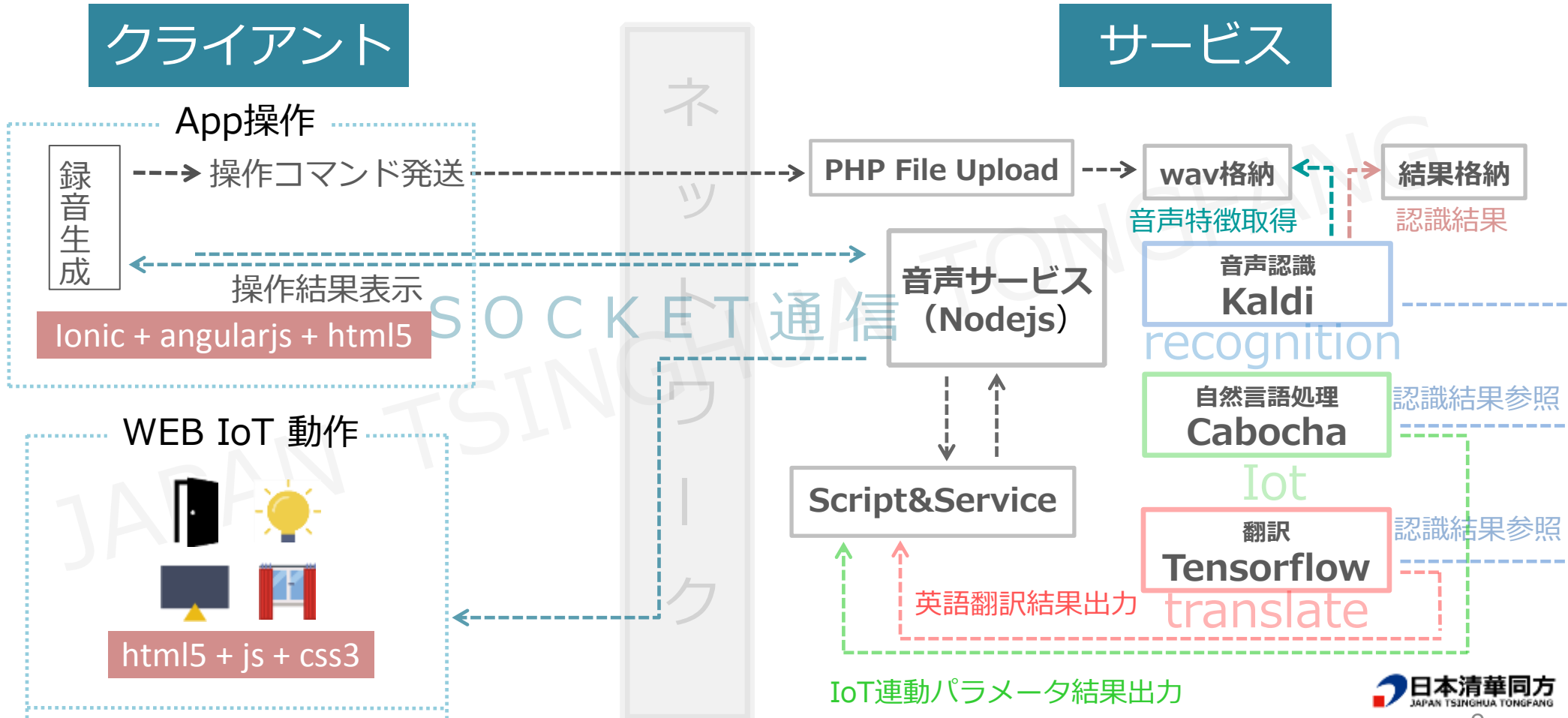
**フォーム性**が都市部の若者が長期休暇を利用して地方で働くふるさとワーキングホリデー制度を二千十七年にスタートするとマスコミ各社が報道

現在の研究結果

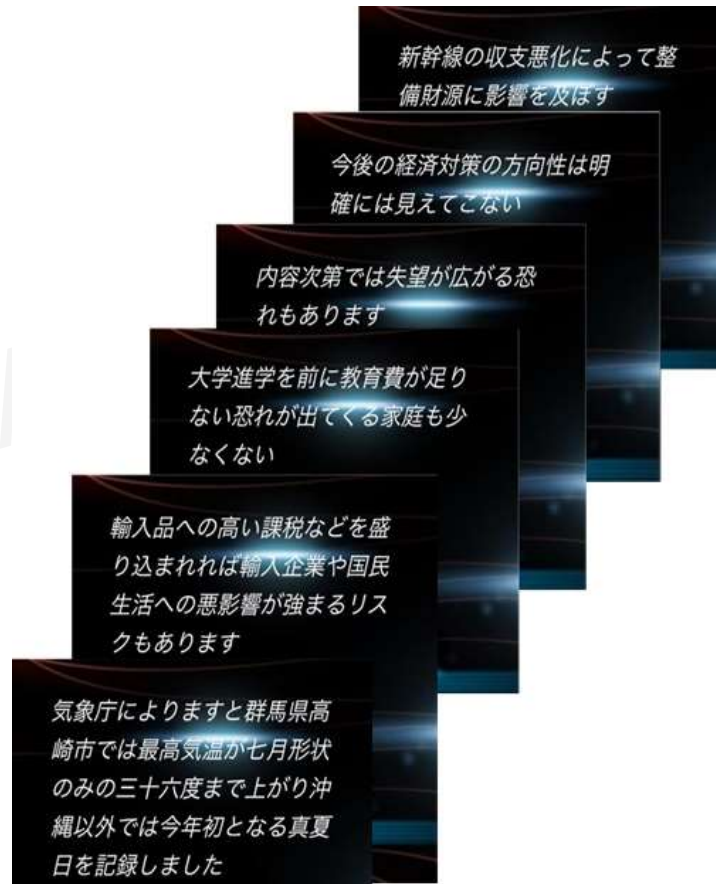
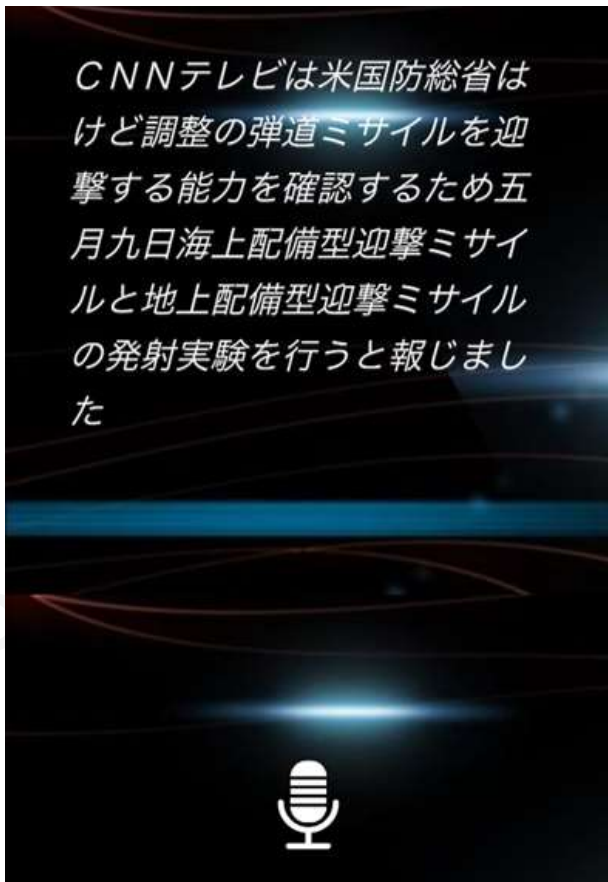
**法務省**が都市部の若者が長期休暇を利用して地方で働くふるさとワーキングホリデー制度を二千十七年にスタートするとマスコミ各社が報道



# DEMO構成図



# DEMO 1 | 識別実例



# Appダウンロード手順

---

1. お客様自身でiPhoneのUDIDを取得しておきます
2. メールで取得したUDIDを弊社担当 [shi.gt@tfsoftec.co.jp](mailto:shi.gt@tfsoftec.co.jp) に送信します  
弊社にて音声認識サーバーにAppleデバイスへお客様のUDIDを認証します
3. 弊社からお客様へパスワードをお送りして、  
お客様がURL(<https://www.pgyer.com/fWlq>)より  
App(s.t.t)をインストールします

ご清聴ありがとうございました

JAPAN TSINGHUA TONGFANG