

CONFIDENTIAL

VPA(対面AI)UI作成ツールの作成について



CODE Mitou Foundation
for JAPAN 一般社団法人未踏



Fairy Devices



Fairy Devices

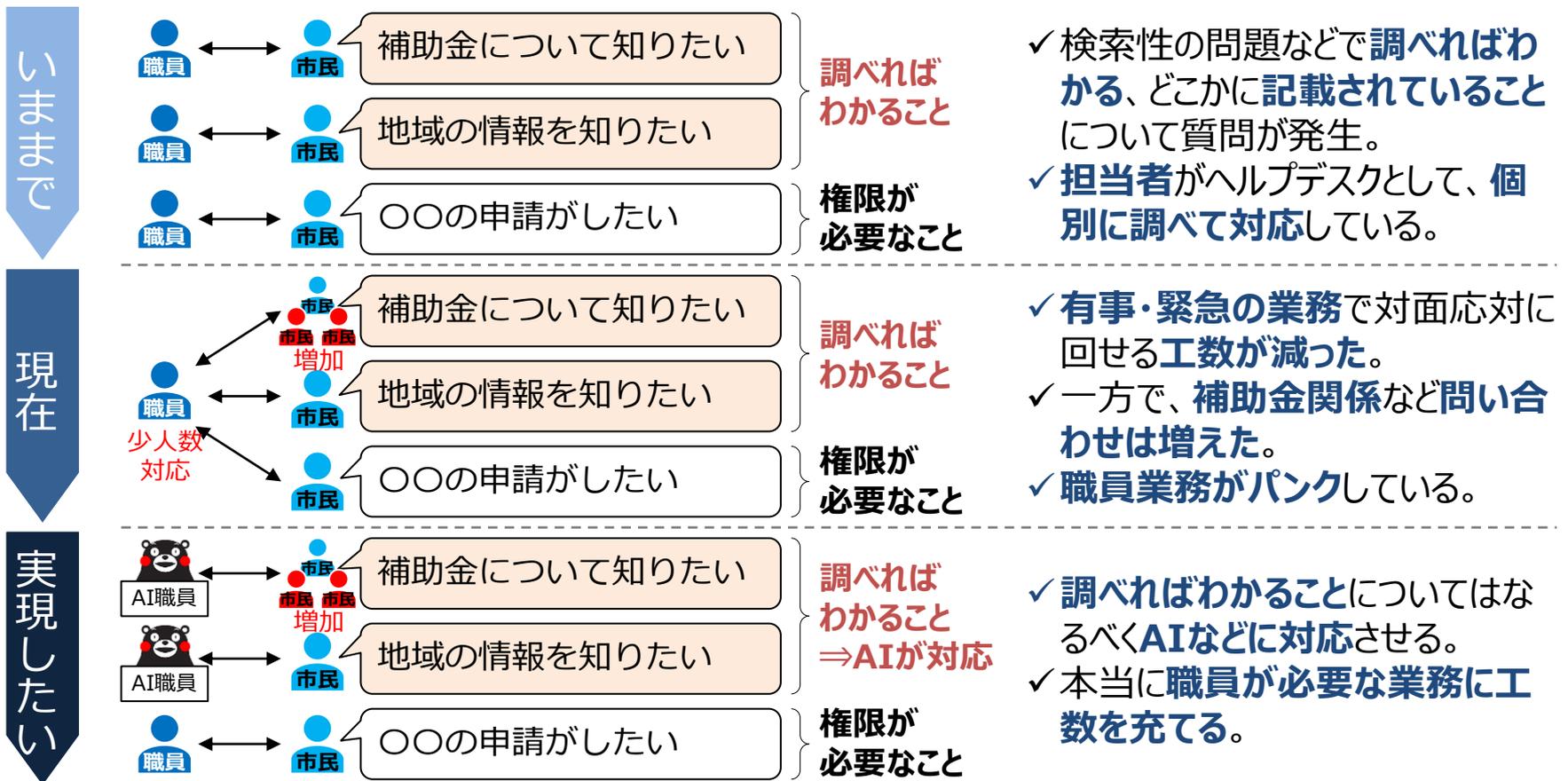
VPA(対面AI) UI作成ツールについて

会社紹介

対面応対業務のパンクをAIで解消したい。



COVID-19の流行で在宅勤務化や有事・緊急対応などが余儀なくされた結果、元々の窓口業務に割いていた工数が不足し、業務のパンクが各所で発生。
起きていることイメージ
何が発生しているか？



⇒ 窓口業務のうち、調べればわかる質問系についてはAIに対応させたい。IAL

VPA(対面AI)の普及がなぜ難しかったか？

窓口対応などの業務付加低減に向けて、VPA(対面AI)が注目されているが、複数の技術連携が必要だったため、これまで構築コストが高かった。

VPA(対面AI) UIの例



※ビックカメラ社のAI店員(試験中) OXYBOT社製 Virtual AI Crew

✓ 技術的には実現可能でGoogle HomeやAlexa含めて便利なのはわかっているはずなのに、VPA(対面AI)は街角で、たまに見かける程度……。実は作るのが大変。

実現しなければならないこと

普通の人
が利用可能にするには？

使いやすい
見た目にするには？

何を話してるか
聞き取るには？

聞き取った内容を
理解するためには？

理解した結果を
適切に対応するには？

必要な技術

App/Web
技術



UI&CG技術



音声認識技術



自然言語
処理技術



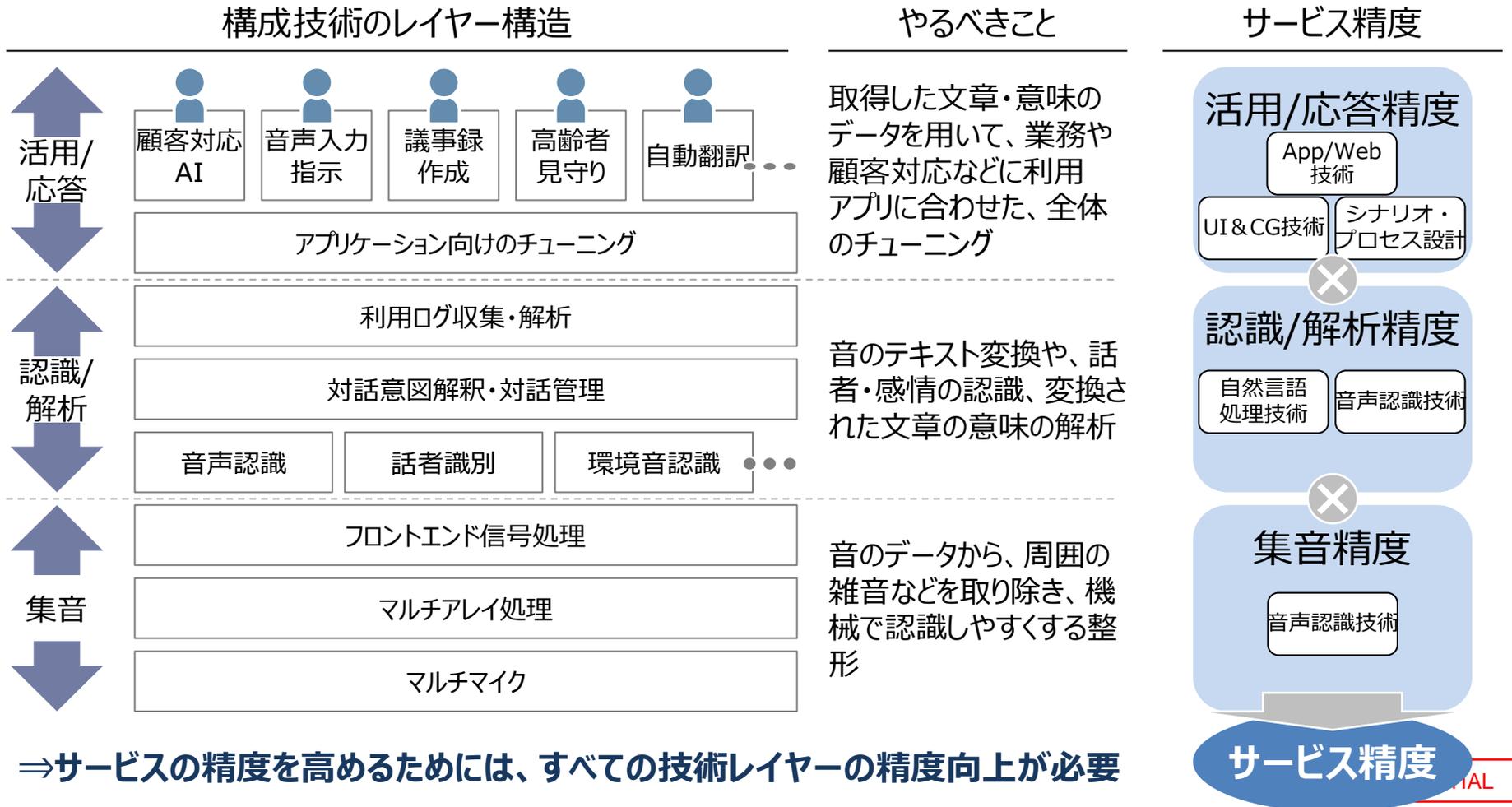
シナリオ・
プロセス設計

一方で、会津若松市のマッシュ君など成功例が存在。
技術を民主化し、人手不足を解消するAI普及を実現したい。

(参考)自然言語サービスの精度を高める為に



自然言語アプリケーションを構築する際に目に見えるアプリケーションだけではなく、マイクによる集音や、認識/解析エンジンの精度が重要な要素



今回実現する内容“VPA-UI作成ツール”

VPAの実現に必要な技術をソフトイーサ社とフェアリーデバイス社で持ち寄り、簡易に実装可能なツールを作成。自治体別の導入はCode4Japanのエンジニアが実装

VPA UI作成ツールの概要



実現する内容

担当する組織(想定)

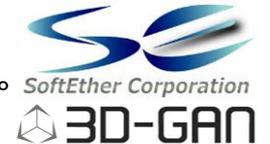
App/Web技術

HTML5などで簡易にWebサイト、WebAppに実装可能な形にすることで、各地域のエンジニアが実装可能



UI & CG技術

ソフトイーサ社の既存技術を活用し、キャラクタデータを簡易に差し替え可能。データの作成は(社)3D-GANが支援



音声認識技術

複数の企業用VPAで導入実績のある国内最高水準の音声認識技術をフェアリーデバイス社が提供



シナリオ・プロセス設計

Google DialoFlowなど標準的なシステムとAPI連携し、各地域のエンジニアが実装可能



自然言語処理技術

日本語での応対に留まらず、NICTの技術を用いて、主要4言語(日・英・中・韓)を含む10言語に対応



PoCを通じて自治体が低コストで保守・運用可能な実装を目指す **IDENTIAL**

動作イメージ（参考）

下記は本件に際して作成したプロトタイプとなります。



App/Web
技術

ブラウザで動作します。
HTML5で実装され、容易にクロスプラットフォーム利用が可能です。

音声認識技術

選択肢、文字入力による対応の他、
音声認識による対応も可能です。

UI & CG技術

キャラクターの3Dモデルは切り替えられます。
自治体などではゆるキャラを用いることが想定されます。

シナリオ・
プロセス設計

プログラミングの知識無く、分岐やキーワードを設定することで質問結果を反映することが可能です。

自然言語
処理技術

あいまいなキーワード入力になっても、
自然言語処理技術を用いて候補を提示できます。多言語にも対応します。

※画像はソフトイーサ社製 3Dソフトウェア「キャラミン」のキャラクター「許斐まい」

<https://www.charamin.com/blog/post/0008>

CONFIDENTIAL

VPA(対面AI) UI作成ツールについて

会社紹介

Comprehensive Audio AI Solution Provider for Enterprise Customer



2M+
Licenses

600K
Daily
Access

Fairy
I/O
Edge AI

Company Overview

- 設立 2007年4月
- 本社 東京都文京区湯島2-31-22
- 取締役

藤野 真人
(代表取締役)

名井 哲夫
(元情報通信事業
副本部長)

山田 純
(元クアルコム
ジャパン会長)



Selected Corporate Partners

LINE



Selected R&D Partners



東京大学
THE UNIVERSITY
OF TOKYO

CONFIDENTIAL



藤野 真人

1981年 埼玉出身、北海道、青森、山形等で育つ

2001年 東京大学理科I類入学

2005年 同大学院医学系研究科医科学専攻入学

2006年 株式会社RNAi 代表取締役

2007年 同大学院中退

2007年 東大構内にてフェアリーデバイス株式会社設立

2009年 株式会社RNAi 退職

■ その他の活動（太字は現職）

2009年 **IPA 未踏ソフトウェア開発事業**「弦楽器自律演奏システムの開発」

2015年 **日本電産顧問（非常勤）**

2018年 総務省「先駆的ICTに関する総務大臣懇談会」メンバー

⇒ **従業員のうち3名が個別にIPA未踏事業に採択経験**

CONFIDENTIAL

ビジョン 人とAIをつなぐ

ミッション コネクテッドヒューマンを実現する

会社沿革

2007年 創業登記（社員2名）

2010年 音声処理技術の研究開発を開始

2012年 音声認識クラウドサービス mimi の開始

2016年 Google テクノロジーパートナーに認定

2017年 mimi powered by NICT の発表

2018年 mimi 多言語対応（10カ国語）の開始

2018年 業務用スマートスピーカー「Tumbler」の発表

2018年 第三者割当増資（シリーズA）

2019年 mimi 利用台数 200万台を突破

2020年 第三者割当増資（シリーズB）実行中



CONFIDENTIAL

Why US?: Key Technologies

様々な現場の騒音環境で、クリアな音声認識・翻訳を実現。長年に亘り大手企業各社に使用されてきた、オンリーワン・ソリューション



- 音声エッジAIによるフロントエンド処理
- 16~64個のマルチマイクを同時制御
- エッジとHuman Data CloudがDNN連携

- ヒトが生成する非定型データ（音声・映像等）に特化したクラウドプラットフォーム
- 音、言語、感情、環境まで、音声の総合テクノロジースタック

- スマートスピーカーからウェアラブルまで、「ハードウェア+ソリューション」を実導入
- 業界毎にグローバル#1パートナーと共同でサービスを開発

SHARP

DAIKIN

OMRON

 **TIS**
TIS INTEC Group

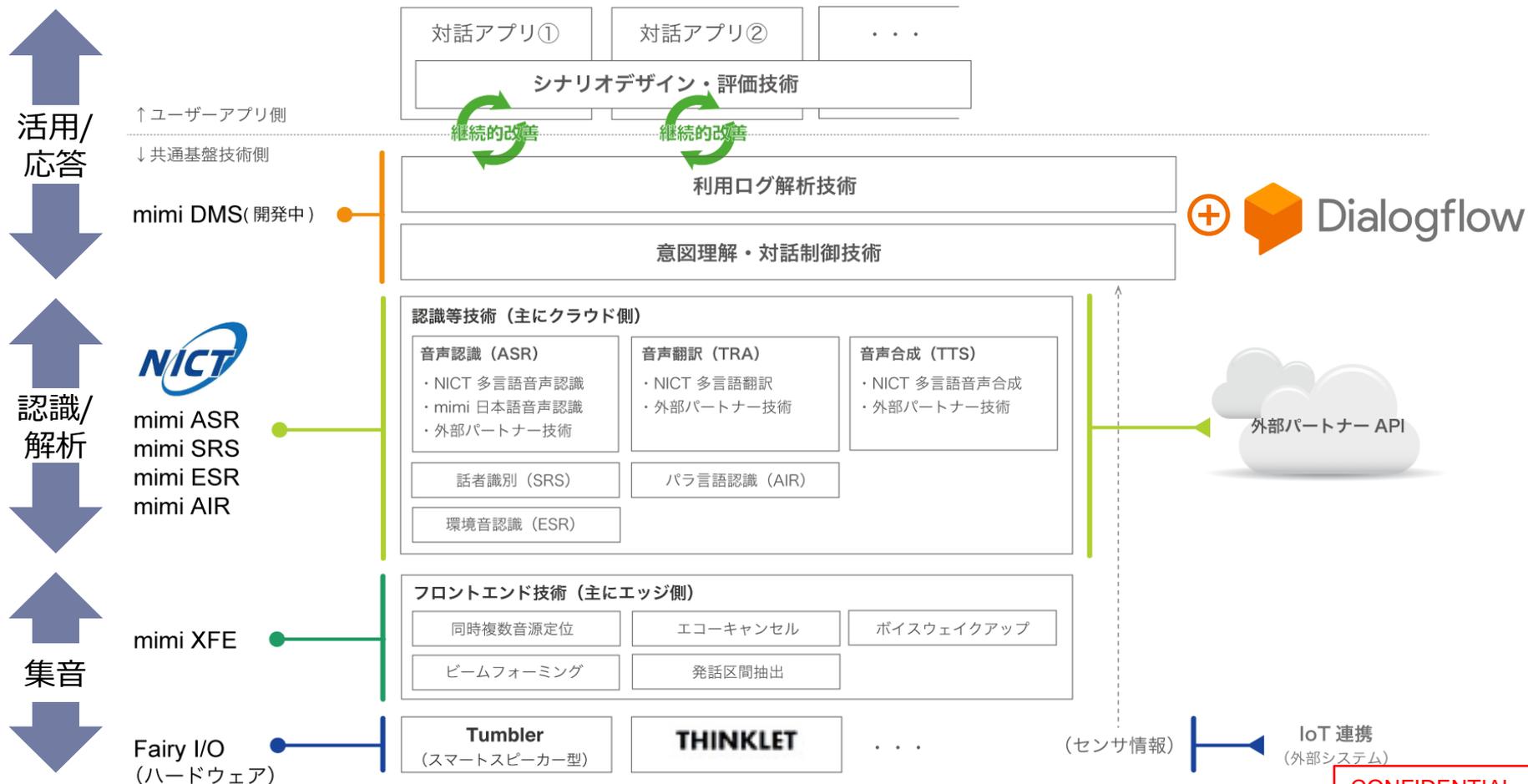
NEC

CONFIDENTIAL

音声認識分野で世界有数の技術力



弊社では音声認識に関する各テクノロジースタックを柔軟に提供し、お客様のニーズに合わせた音声認識ソリューションを実現可能です。



CONFIDENTIAL



mimi 音声認識クラウドサービス

多言語音声認識、自動機械翻訳、自動言語判定、音声合成、感情認識、環境音認識、話者識別等を搭載した、音声AIの総合クラウドサービス。

mimi XFE

mimi XFE 音声処理ライブラリ

屋外等の高騒音環境下でも正確な音声認識を実現するためのエッジ側で動作するソフトウェア。世界的にもオンリーワンの技術。



Tumbler

Tumbler T-01 業務用スマートスピーカー

高性能マルチマイクハードウェア。mimi、XFEと組み合わせ、同時通訳、発言録、窓口対応補助業務等に利用されている。



共同研究開発事業

音声認識、画像認識、その他機械学習に関する共同研究開発、及び応用ソリューションの開発事業。

CONFIDENTIAL



Tumbler

Tumbler™ (業務用高性能スマートスピーカー)

窓口での同時通訳や発言記録に加え、対話型ロボットやデジタルサイネージ等にも利用される、高性能スマートスピーカー

mimi®

XFE

mimi® XFE (音声処理エッジAI)

騒音環境でもクリアな音声を録音を可能にするエッジ音声AI。



多言語対応駅案内ロボット



日英中韓で、駅周辺情報をロボットが自動案内



TIS

TIS INTEC Group

議事録自動作成サービス



「日々のことば」を資産とし、ビジネスに新しい価値を創出する音声・対話AIサービス



会議の議事録を自動作成
対話データを分析・活用

※OXYBOT社 Virtual AI Crew向け
サイネージ向けマイク※



駅や商業施設の騒音環境で、
会話可能なAIエージェントを実現

CONFIDENTIAL

多言語地域案内ロボットによる省人運用化



多言語案内ロボット

<特徴>

「多言語・自動の地域案内」によって郵便局を街のホットスポット化。

「リアルタイム送客」による収益化が可能

参考事例



OMRON

多言語「駅」案内ロボット



新宿駅



下北沢駅

ラグビーワールドカップにて活躍！政府も海外にPR

➡ **大好評につき、全国に拡大中！**

参考TV動画：

https://www.youtube.com/watch?v=PR90IgjD_W8

活用イメージ



外国人



子供連れ



カップル



修学旅行

顧客対応
機械翻訳

(A)職員無しで自動的に、地域の周辺情報を音声+映像で案内



(B)訪問ユーザーをリアルタイムで最適施設へ送客



観光名所



商業施設



公共施設



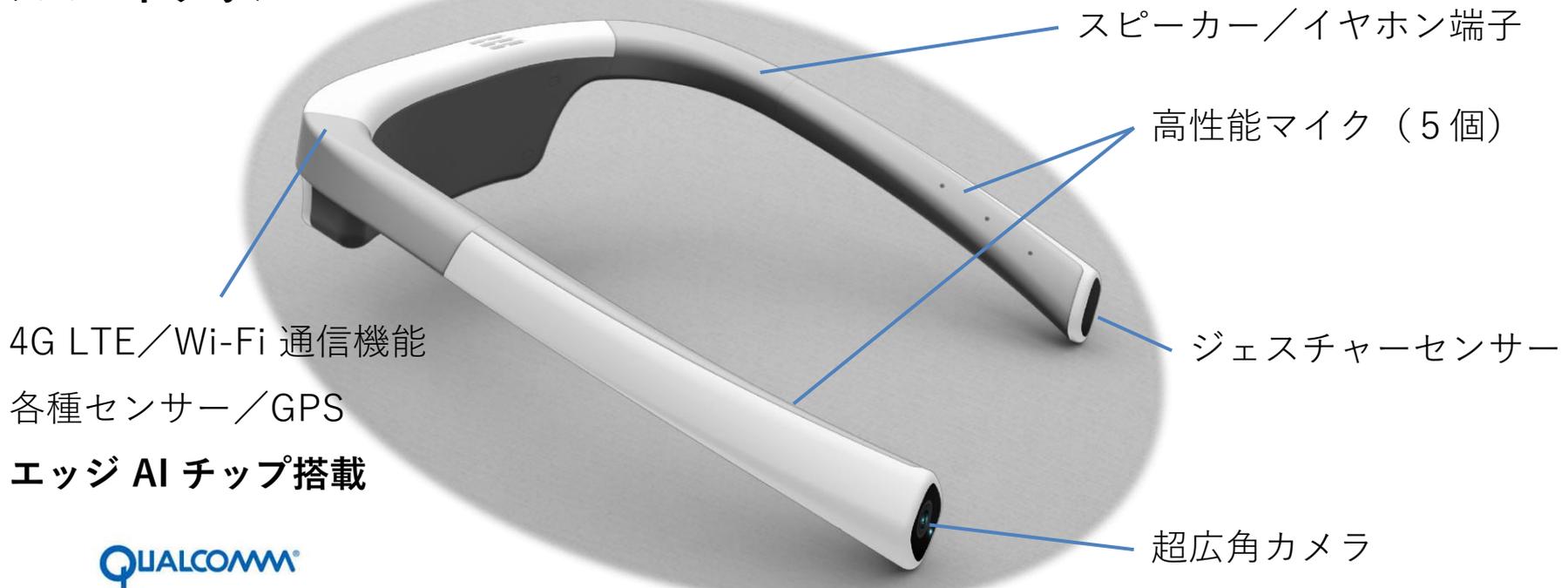
イベント

送客
混雑回避

装着者の見ているもの、聴いているものをリアルタイムにデータ化し、今までにない情報活用を可能にすることができるワーカー支援デバイスを開発

人とAIをつなぐ機械 THINKLET[®]

人が得意なことは人に、AIが得意なことはAIに任せる役目を担うウェアラブルスマートフォン



CONFIDENTIAL

世界初。現場業務のデジタルトランスフォーメーションを実現する「セカンドブレイン」



- **「ヒトが作り出す非定型データ」**のデジタル化を実現する首掛け型 **「ウェアラブルAI」**
- XFEによる高精度の音声データ収集と、超広角カメラによる一人称視点の画像データ蓄積を実現。固定マイク/カメラでは成し得なかった、**「ヒューマンビッグデータの可視化」**を実現
- LTEを内蔵し、遠隔作業支援のみならず、**「解析データに基づく現場支援ソリューション」**を実装可能

代表的な用途

DXレベル0

業務活動情報のデジタル化
業務現場の情報を、安全に収集しデジタル化・保存。現場情報の分析を可能に

DXレベル1

データを活用した業務支援
技能伝承、訓練の質向上、エビデンスで状況判定、ナレッジマネジメント

DXレベル2

AIを活用した業務改革
業務の自動化、効率化、ヒューマンエラー対策

CONFIDENTIAL

ご参考：業界別THINKLET™導入の発表(例)

すでに2社と業務用ソリューションの共同開発を発表しており、今後B2Bソリューションを共同開発するパトロン企業を募集中



DAIKIN

11月21日共同発表



対象領域/活用先

空調機の設置・メンテナンスのフィールドワーカーの遠隔支援、業務デジタル化

ユーザ規模(TAM)

国内
150万人

世界
3,900万人



OMRON

11月27日発表



対象領域/活用先

駅務・ドライバー・倉庫従業員向けの業務デジタル化、AI活用(自動翻訳など)

ユーザ規模(TAM)

国内
120万人

世界
2,400万人

CONFIDENTIAL