

# 対話型アシスタント技術「netpeople」の実用事例

## Case studies of netpeople, a cloud-based conversational digital assistant

風見 清司\*<sup>1</sup>、樋口 恵一郎\*<sup>1</sup>、李 惠琳\*<sup>1</sup>、森 祥二郎\*<sup>2</sup>

Kiyoshi Kazami, Keiichiro Higuchi, Hui-Lin Lee, Shohjiroh Mori

\*<sup>1</sup> イナゴ株式会社 iNAGO Inc.

\*<sup>2</sup> 富士通テン株式会社 FUJITSU TEN LIMITED

**Abstract:** iNAGO's platform product netpeople is the launch pad that enables businesses to offer a smart assistant in the cloud for any device and service. netpeople provides actual human-like conversation and interaction with exclusive Content-Aware Goal-Oriented technologies that work together in unison.

### 1. はじめに

「Ok, Google」<sup>[1]</sup>「Hey, Siri」<sup>[2]</sup>などと呼びかければ、自動で音声認識が起動し、ユーザーの言葉を理解する音声アシスタントは、スマートフォンだけでなく、コネクティッド・カーやスマートホームなど、その領域を拡げている。音声認識の世界市場は、2017年までに1130億ドル（約13.5兆円）規模に成長するとの調査結果<sup>[3]</sup>も出ている。しかしながら、人間の言葉の意図を理解し“対話”ができるアシスタントサービスはまだ少ない。イナゴでは、独自の HCI (Human-Computer Interaction) 技術により、ユーザーの発話に動的に返答することが可能な音声アシスタント技術・netpeople の開発を行っている。この論文では、netpeople の概略および実用事例を紹介する。

### 2. netpeople とは

netpeople は、自然な会話から利用者の意図を理解し必要な対応を実行する、対話型音声アシスタント技術である。端末に実装された音声認識エンジンから、テキスト化されたユーザーの発話内容を受け取り、netpeople が発話意図の解析後および必要な情報検索処理を行い、最適な返答を音声合成エンジンに引き渡す。従来型の音声アシスタントとの大きな違いは、動的に返答を生成できることである。

多くの音声アシスタントは、システム側で設定された会話フローでユーザーの発話を機械処理するため、会話フローが存在しない発話には対応できない。人間のあらゆる発話に対応するには、会話フローを追加し続けていくしかなく、実用化には適さない。

これに対し netpeople は、会話の内容や状況に応じて柔軟に対応ができる、ユーザー主導型の対話シ

テムを構築した。独自開発の Goal-oriented Technology (ゴール共有型会話技術)<sup>[8]</sup>、Context-Aware (文脈把握技術)、Natural Language Understanding (NLU:自然言語理解)により、必要に応じてシステム側から積極的に質問を行うことでユーザーの発話目的を把握し、最適な情報へ効果的に誘導する。

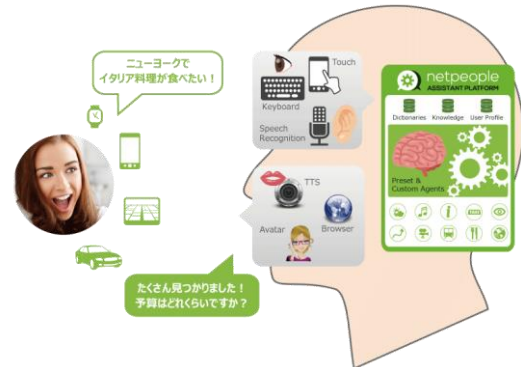


図1：音声認識を「耳」、音声合成を「口」とすれば、netpeople は「脳」にあたる。

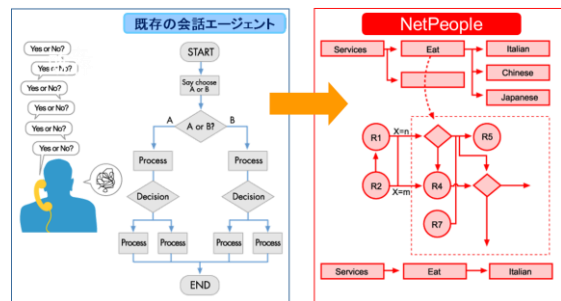


図2：あらかじめシステムに設定された固定型会話フローではなく、ユーザー主導型の臨機応変な会話対応が可能

## 2.1 Goal-oriented Technology (ゴール共有型 会話技術)

ユーザーが発話した言葉をその都度認知するだけでなく、前後の発話内容と合わせて、利用者の意図(ゴール)を総合的に理解する。同時に適切なタイミングで適切な質問を動的に行うことで、ゴール到達に必要な条件を追加取得し、利用者を効率的にゴールに誘導する。



図3: Goal-oriented Technology による会話例

## 2.2 Context-Aware (文脈把握技術)

ユーザーの発話内容をコンテキスト(文脈)として記憶することで、会話の途中で話題が変わっても、次の話題に必要な情報を引き継いで会話を行い、新しいリクエストに対応する。



図4: Context-Aware による解析イメージ

## 2.3 Natural Language Understanding (NLU:自然言語理解)

一般的な音声認識は、単語がどのように使われているか判断できないため、言葉の意味を理解するこ

とはできない。netpeopleのNLUエンジンは、単語が会話の中でどのように使われているかを分析し、ナレッジとともにコンテキストに保存することで、言葉の意味を理解する。



図5: NLUによる解析イメージ

## 3. netpeople アシスタントプラットフォーム

音声アシスタントは、大手IT企業が個別に開発し、サービス提供しているケースがまだ多い。イナゴでは、一般企業が音声アシスタントを自社サービスに導入する際の負担低減を目指し、開発ツール「netpeople アシスタントプラットフォーム」を用意している。レストラン検索など標準アシスタント機能がプリセットで搭載されたクラウド型プラットフォームで、APIおよび各種プロトコルがあらかじめ定義されており、外部システムとの統合が自由にできる。GUIベースで、プログラマーでなくても目的に応じた独自のデジタルエージェントの開発ができる。また、組み込みのプログラミング言語を通して高度な機能を追加することも可能である。プラグインシステムを使用すれば、自由にコア機能を拡張することもできる。



図6: netpeople アシスタントプラットフォーム

## 4. netpeople 導入事例

### 4.1 mia (ミア) | 音声対話アシスタント<sup>[4]</sup>

イナゴの Android 向けショーケースアプリケーション。音声アシスタント mia が会話を通して、POI・情報検索、メール、電話などの操作をサポートする。ドライブモードがあり、運転中での操作が可能。

主な機能：

- グルメ店舗会話絞り込み検索
- 乗換会話案内
- 地域情報会話絞り込み検索
- イベント会話絞り込み検索
- 目的地設定機能
- メール/Twitter 等との連携 ※モバイルモードのみ
- Web 検索 (Google、Wikipedia 等) ※モバイルモードのみ
- 音楽再生 (Beta)
- 天気予報
- スケジュール作成
- 電話 (架電) 機能
- Gmail 送信機能
- SMS 送信機能
- モバイルモード、ドライブモード切り替え対応
- 雑談 (スモールトーク) 機能



図 7：mia powered by netpeople 画面イメージ (左：モバイルモード、右：ドライブモード)



図 8：mia powered by netpeople サービスイメージ

### 4.2 富士通テン株式会社

【カーナビ用対話型エージェントアプリ「CarafL (カラフル)」】<sup>[5]</sup>：「CarafL」をインストールしたスマートフォンと「ECLIPSE」カーナビを Wi-Fi®で連携し、音声で施設検索や目的地設定ができる。カーナビ専用のマイクに話しかけると、発話した音声をセンターサーバーで解析する。netpeople は意味理解処理を実行。

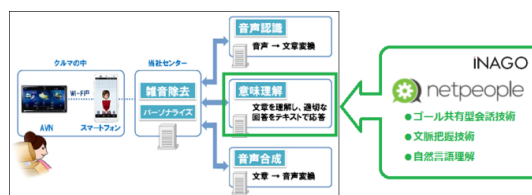


図 9：「CarafL」・netpeople 連携図

### 4.3 NEC パーソナルコンピュータ株式会社

【音声操作アプリ「LAVIE ボイス」】<sup>[6]</sup>：パソコン画面内に現れるアシスタントキャラクタ「パイロ」が Web 検索をサポート。音声によるアプリ起動や文字入力も可能。ネットワーク接続がない場合にも動作するローカル組込み型、および高性能なクラウド型の 2 種の音声対話システム方式を連携させたハイブリッドアプリケーションシステムを実現。



図 10：「LAVIE ボイス」音声対話システム方式

### 4.4 株式会社ヤマダ電機

【音声アシスタントアプリケーション「mia 音声検索」】<sup>[7]</sup>：ヤマダ電機のプライベートブランドタブレットのおでかけアプリケーション。netpeople アシスタントプラットフォームの標準機能を利用。GUI

にオリジナルキャラクターのデンちゃんを使い、ブランドイメージを訴求。

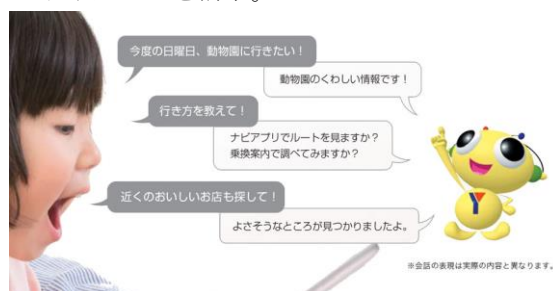


図 11 : 「mia 音声検索」 会話例

## 5. おわりに

Apple による Siri の登場で、音声認識が技術開発分野から音声認識サービス市場に転換したと言っても過言ではないだろう。現在では、Google、Microsoft、Amazon、Facebook など、IT 業界をリードする企業が音声アシスタントサービスの提供を開始している。また、先に述べた音声認識の世界市場予測では、今後、バイトメトリクス、ヘルス、オートモーティブ分野での成長が期待されている。技術の普及、市場の拡大には、一般企業が莫大なコストをかけずに開発できる環境が必要である。イナゴでは、今後も HCI 技術のビジネス展開を視野に入れた、開発プラットフォームの研究開発に取り組む。

## 参考文献

- [ 1 ] Google Inc., Google 音声検索,  
<http://www.google.co.jp/dekiru/android/index.html>  
(accessed: Oct. 7, 2015)
- [ 2 ] Apple Inc., Siri, <https://www.apple.com/jp/ios/siri/>,  
Copyright © 2015, (accessed: Oct. 7, 2015)
- [ 3 ] Technavio, “Global Voice Recognition Market Forecasted to Reach \$113 Billion by 2017”,  
<http://www.technavio.com/news/global-voice-recognition-market-forecasted-to-reach-113-billion-by-2017>,  
(accessed: Oct. 7, 2015)
- [ 4 ] mia powered by netpeople, <http://mia-netpeople.com/ja/>,  
(accessed: Oct. 7, 2015)
- [ 5 ] 富士通テン株式会社、CarafL (エージェント)、  
<http://www.fujitsu-ten.co.jp/eclipse/product/wifi/carafL/index.html>, (accessed: Oct. 7, 2015)
- [ 6 ] NEC パーソナルコンピュータ株式会社、プレスリリース：インフォボードをはじめとするソフト面を拡充した 2015 年夏モデルを発表、  
<http://www.necp.co.jp/press/ja/1505/1201.html>, 2015 年 5

月 12 日、(accessed: Oct. 7, 2015)

- [ 7 ] 株式会社ヤマダ電機、EveryPad,  
<http://www.yamada-denki.jp/service/everypad/>, 2015 年 5  
月 12 日、(accessed: Oct. 7, 2015)

## 注釈

- [ 8 ] 特許出願中 (2015 年 10 月 7 日現在)