NUANCE

進化するAI音声アシスタントの最新技術

2019.05.27　ニュアンス・コミュニケーションズ・ジャパン株式会社

シニア プリンシパル テクニカル エクスパート　石川　泰

© 2019 Nuance Communications, Inc. All rights reserved.

Contents

Introduction: Nuance

音声インタフェースのこれまでと今

AI音声アシスタント

MaaS時代のHMI

まとめ　今後の展望

6月6日-7日のイベントで

NUANCE

© 2019 Nuance Communications, Inc. All rights reserved.

Nuance Communications

音声認識・合成

自然言語理解　　手書き文字認識

革新的対話AI音声アシスタント

音声認証　　AI

2億台以上
Nuanceの技術が搭載された
自動車台数（グローバル）

50言語以上
サポート言語

#1
世界の自動車メーカおよび Tier 1
サプライヤーの ナンバーワン
OEM パートナー

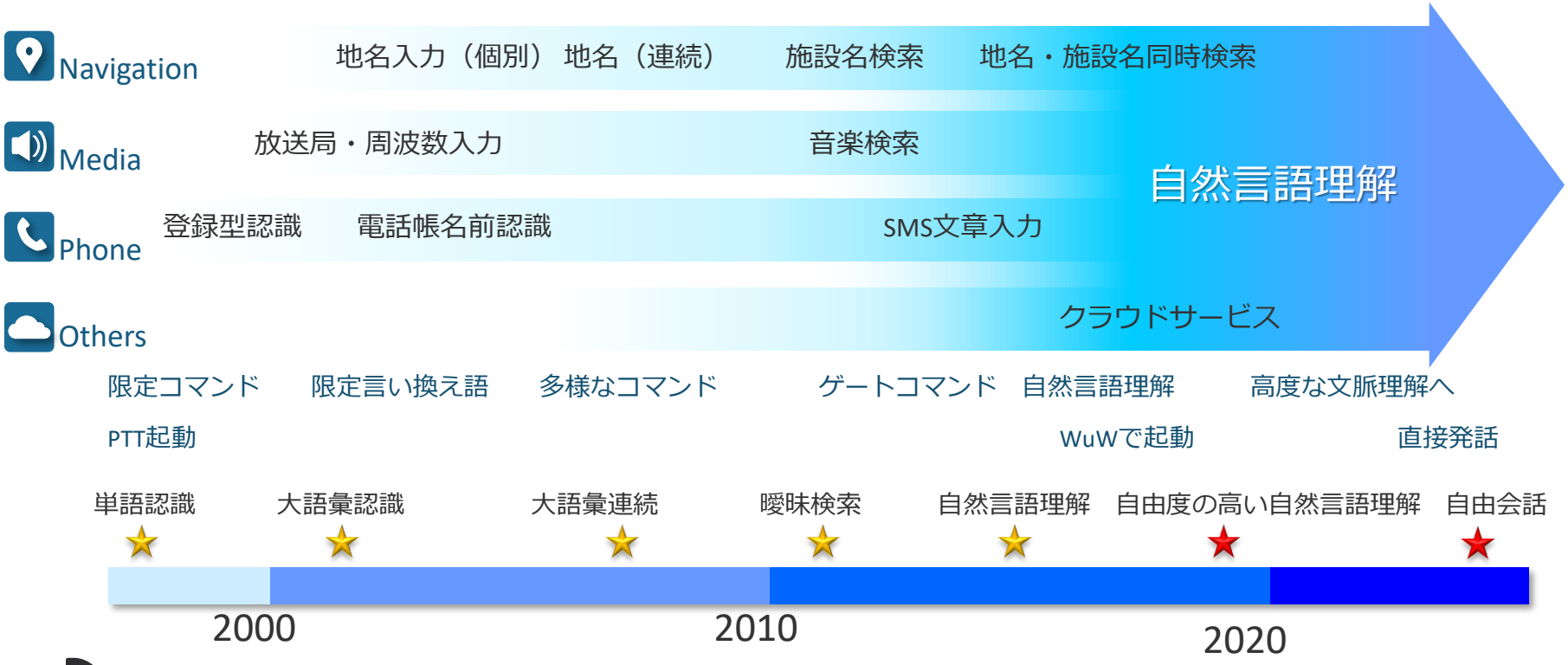
© 2019 Nuance Communications, Inc. All rights reserved.

補足

- ニュアンス・コミュニケーションズ・ジャパン株式会社
  - https://www.nuance.com/ja-jp/index.html
  - 本社：Nuance Communications Inc. （米国マサチューセッツ州）
  - ドイツ，ベルギー，カナダ，日本，中国，韓国などに35以上の地域拠点．70ヵ国以上で製品・サービスを提供
  - 音声・イメージング技術
  - 自動車，エンタープライズ，ヘルスケア等の事業分野
  - 開発製品例
    - MBUX　https://www.youtube.com/watch?v=v5nwuHmkbaw　（英語）

© 2019 Nuance Communications, Inc. All rights reserved.

音声認識技術の進歩



© 2019 Nuance Communications, Inc. All rights reserved.

補足

- 音声認識の基本的起動方法
  - PTT　　Push To Talk
    - 音声認識を起動するためのスイッチ，スイッチにより起動する音声認識方式を示します．多くの車載システムで使われてきました．
  - WuW　　Wake up Word
    - 特定のキーワード（例：”Hey Mercedes”）の検出によって音声認識・対話を起動する方法です．常時キーワード検出が行われます．その間，音楽やナビの案内音声などシステムから出力音声がマイクに入力されるため，スピーカからマイクまでの音響特性を推定し，入力音声から除去し，ユーザの声だけを抽出する技術（エコーキャンセリング）が必要です．最新のシステムで使われはじめました．
  - 直接発話方式：JustTalk
    - 常時音声認識を実行し，その内容，発話特徴からシステムに向けての発声かどうかを判断し，スイッチ，キーワードなしに対話を行うことができる方法です．今後のシステムで使われることになります．ただ，この方式では，常時音声を認識するため，プライバシーの保護が必要です．Nuanceでは，認識サーバをつかわず車載機器内でこれを実現しています．

© 2019 Nuance Communications, Inc. All rights reserved.

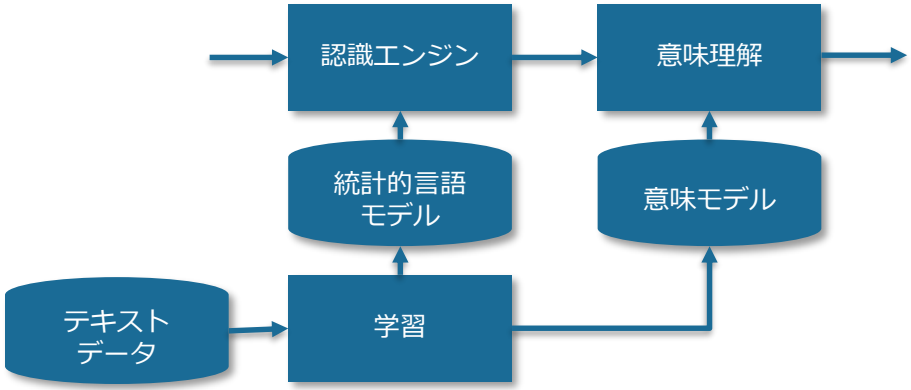
補足

- 車載システムの音声認識に要求される技術
  - 地名認識
    - 数万以上の語彙の連続となるため，きわめて大規模な対象の認識となります．かつては組み込み機器での実現が困難な技術でした．ナビゲーションでの目的地設定として要求が高く，2000年代半ばから，単語地名認識，連続発声地名認識をいち早く実現してきています．
  - 施設名認識，音楽検索
    - ユーザが正式名称を知らない場合もあり，「ヒルトン東京お台場」に対して，「東京お台場ヒルトンホテル」，「Sgt. Pepper’s Lonely Hearts Club Band」に対する「Sgt. Pepper’s」のように，倒置，部分発話，不要語の挿入があっても，認識し，検索ができる必要があります．Nuanceは単語検出とあいまい検索技術によるFuzzyMatchでこれを車載器でいち早く実現しました．
  - SMS ディクテーション
    - ショートメッセージの文章入力．任意の文を認識できる必要があり，統計的な言語モデルを用いる音声認識がになります．

© 2019 Nuance Communications, Inc. All rights reserved.

補足

- 自然言語理解 NLU（Natural Language Understanding）



文例とその意味を記述したテキストデータから，言葉の局所的な並びかたの統計的性質を学習し，言語モデルを構築します．通常は，3単語程度の単語列の出現確率を用います．これにより，任意の文で，もっとも言語的に妥当な認識結果が得られます．また，単語列と文の意味の関係を機械学習し，単語の系列，共起関係などから意味を推論する意味モデルを構築し，自然言語理解を実現しています．この方法では汎用の言語理解ができるわけではありません．サービスに合わせて学習データを構築し，そのサービスにあわせた自然言語理解を実現することになります．

© 2019 Nuance Communications, Inc. All rights reserved.

補足

- 車載システムの音声認識に要求される技術
  - 高騒音下音声認識
    - 走行騒音，同乗者の話声など騒音は大きな問題です．マイクロフォンが口もとにあるスマートフォンとは異なり種々の技術が必要です．
      - 複数のマイク（マイクアレー）の信号処理により，高い指向性を実現し，SN比を改善するビームフォーミング技術が多くのシステムで利用されます．
      - 長年の車載機器での音声認識の蓄積により，信号処理でのノイズ低減，ノイズ耐性の高い音声認識技術を保有しています．
  - バージイン（Barge-in）
    - かつてのシステムでは「ピッと鳴ったらお話しください」とメッセージが出ていましたが，ユーザはこのビープ音を待たずに話してしまう傾向があります．また，使い方がわかっているにもかかわらず「続きの住所を言う場合は，住所を，この場所を目的地と設定する場合には…」などと長いプロンプトを待つのも非効率です．7ページのWuWで説明したエコーキャンセルの技術で，ビープ音やプロンプトの出力中に話し出しても，正しく認識し，出力中のプロンプトを止めることができます．これをバージインと呼びます．

© 2019 Nuance Communications, Inc. All rights reserved.

HMIの視点から

自然言語理解技術

Hello Dragon, 近くのイタリアンレストランを教えて

要求の直接表現によるHMI

過去の音声操作

ピッと鳴ったらお話ください

なにを行いますか

ジャンルをお話しください

もう一度お話ください

ナビゲーション

施設検索

イタリアン

CANCEL !!!!

© 2019 Nuance Communications, Inc. All rights reserved.

HMIの視点から

自然言語理解技術

要求の直接表現によるHMI

Hello Dragon, 近くのイタリアンレストランを教えて

Navigate to 103 Broadway Cambridge

ちょっと暑い！

目的地の天気はどう？

Send message to John, “sorry, I’ll be 15 minutes late”

クイーンのアルバムで，オペラ座の夜をかけて

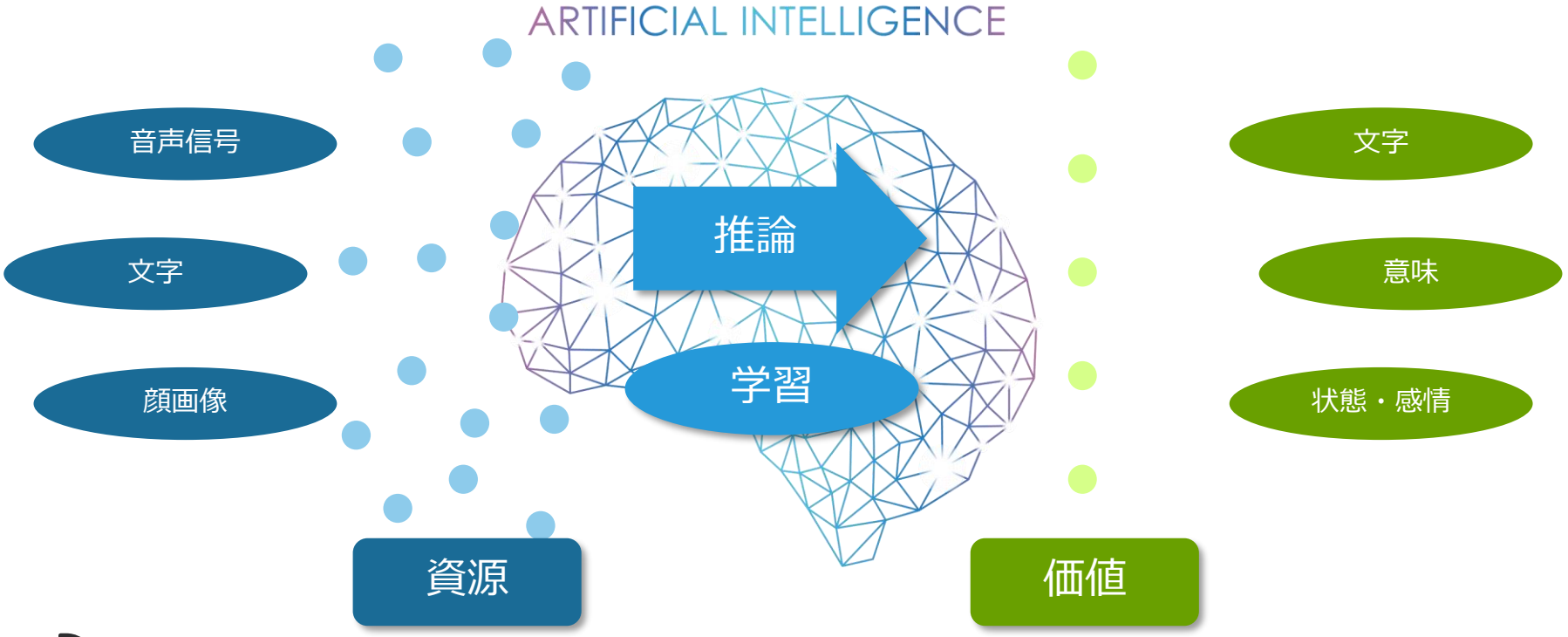
NUANCE 自然言語理解
＋
自動車走行騒音下での高い認識性能
WuW
バージイン（プロンプト出力中の認識）

© 2019 Nuance Communications, Inc. All rights reserved.

AI技術

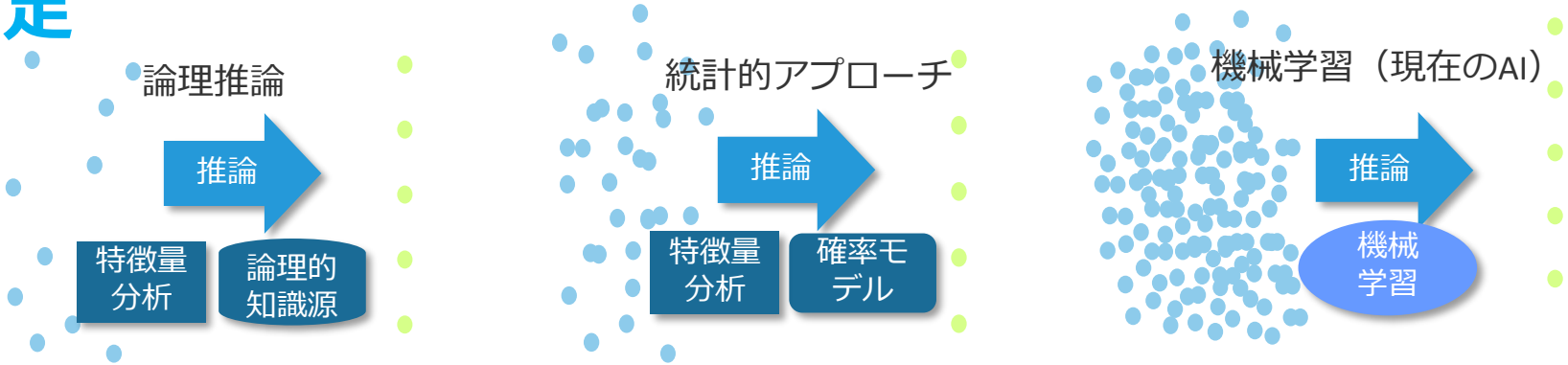
観測された値

得たい意味のある情報



© 2019 Nuance Communications, Inc. All rights reserved.

補足



AIは，学習と推論にかかわる工学的アプローチです．1980年代のAIは人間の知識を論理的に記述し，論理計算により推論を実現しようとしました．しかし，人間の知識はあいまいで状況依存性もあり，単純な論理式で表現できるものではありません．いくつかの成果は出したものの，AIは冬の時代を迎えることになります．

その後，観測データと結果の関係を確率モデルで表現すること，学習データ量を増やすことで，音声認識，画像認識は大きく進歩しました．しかし，確率モデルで扱える特徴量の抽出（たとえば，画像認識では，人は色が変わるところを見ているはずだから，エッジを抽出する）は人間の知識によるモデル化でもあり，確率モデルにも限界がありました．

それに対して，脳の情報処理を数理モデル化したニューラルネットを多層に構築するディープニューラルネットワークの技術が発展し，特徴抽出自体も機械学習に組み込む方法（深層学習）で，この推論技術が大きく発展しました．AIにはさまざまな分野・方法がありますが，深層学習の進歩が大きかったことから，AI＝深層学習で語られることも多々あります．

© 2019 Nuance Communications, Inc. All rights reserved.

音声技術が切り拓くユーザ体験

スカイツリーの近くの駐車場を探して

対話型AIアシスタント

一貫したサービス

目的地を選ぶ

状況に依存したサービス

快適なルートを選ぶ

状況に適合したHMI

一貫したHMI

車内で楽しむ音楽・情報

安全・快適に移動する

状況に依存した意味

基本的な意味

Intelligent Mobility

つながる伝える

ドライバーモニタリング

快適な車内環境を制御

MaaSの影響

自動運転の影響

Home

Train

Ride sharing

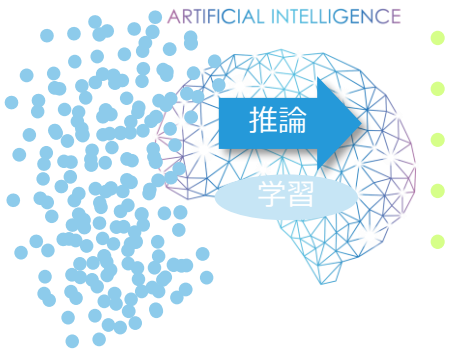
Walk

Office

Car

© 2019 Nuance Communications, Inc. All rights reserved.

Multi-Modality & Context Awareness

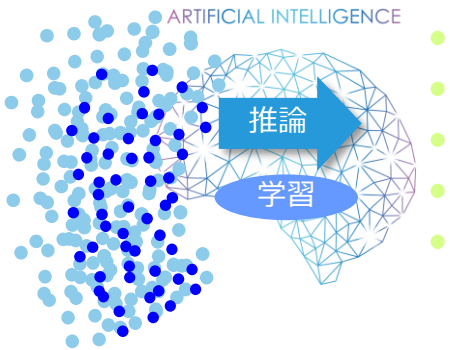


入力をリッチにしよう

今観測できること（ユーザ，環境）
過去に観測したこと（文脈）

© 2019 Nuance Communications, Inc. All rights reserved.

Personalization



個人個人の満足を最大化しよう

データの所有権
プライバシー保護
MaaS時代

ユーザの理解 → ユーザ体験

どんな表情？　気分？

何を見ている？

どこに向かっている？

天気は？

これまで何を話た？

道路の状況は？

車の状態は？

何をしてきた？

自動運転レベルは？

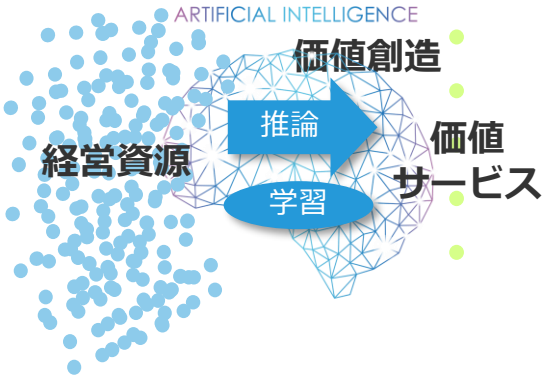
次に何をする予定？

ユーザのシステム・環境理解 ⇔ システムのユーザと環境理解

NUANCE

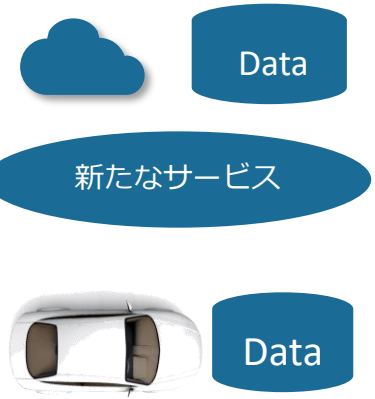
© 2019 Nuance Communications, Inc. All rights reserved.

MaaSとビジネスインテリジェンス



ユーザと，環境・サービスの接点としてのHMI

例：バーチャルドライバー



MaaS

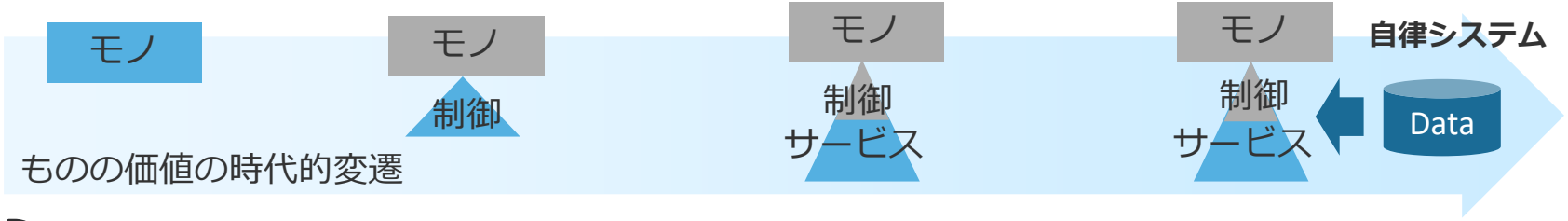
Data Management Eco-System

プラットフォーム

自動車

広義のHMIに求められること

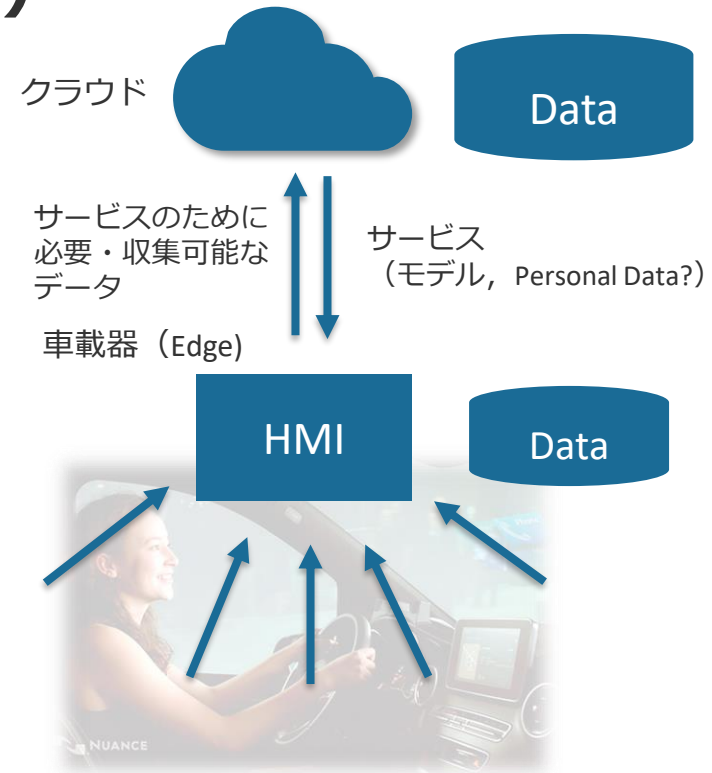
身体的使いやすさ　操作の認知負荷の軽減　ユーザの体験　人工物と人の相互理解とコミュニケーション



© 2019 Nuance Communications, Inc. All rights reserved.

補足（データ流通とサービス）

- データ流通の視点
  - 所有者と，プライバシー
  - どこでデータを活用するか，どのようにそのデータから学習したモデルを利用するのか
  - 経営資源としての流通とEco－System構築
- サービスと必要なデータの関係の視点
  - リアルタイムモニタリングデータ（環境，ユーザを知る）を活用するサービス・HMI
    - 特に自動運転にかかわる処理・サービスは車載システムで実現する必要
- サービス統合の視点
  - 自動車のデータを処理し，自動運転・MaaSを実現するプレイヤー，複数のサービスを統合して提供するプレーヤ



© 2019 Nuance Communications, Inc. All rights reserved.

これからのAI音声アシスタント

MaaS時代のサービスを革新的HMIで実現する

NUANCE

© 2019 Nuance Communications, Inc. All rights reserved.

補足

- Nuanceが今後実現していく技術 ➔ “Cognitive Mobility Assistant”
  - 自然言語理解　➔ より複雑な文の理解へ
    - Multi-Intent （複数の意図の一文表現）
      「野村さんのオフィスに目的地を設定して，到着予定時間を野村さんにメッセージで送って」
    - Not / Or などの検索条件理解
      「この近くで，すし屋以外の和食レストランを検索して」
  - 自動運転時代のAIアシスタント技術
    - ドイツ人工知能研究センター（Deutsches Forschungszentrum für Künstliche Intelligenz : DFKI）と提携し，研究開発を推進しています.
  - 感情を認識するAIアシスタント
    - Affectiva社（ https://affectiva.jp/)とともに，ドライバの状態，感情に応じたアシスタンスの研究開発を推進しています.
  - AR（拡張現実）の活用
    - 情報の提供方法として期待の高いARの活用と，AR環境でのユーザのふるまいを把握するインタフェースの開発を進めています

© 2019 Nuance Communications, Inc. All rights reserved.

補足

- Nuanceの最新技術は，
  - 最新技術は，社内で実証システムを構築し，お客様にお見せしています．
  - CES2019で紹介したNuanceの技術　（日本語字幕つき）
    - https://www.youtube.com/watch?v=v7ySRsxRX9Y&feature=youtu.be

- Nuanceの考える将来像
  - 近未来のモビリティ・エクスペリエンス
    - https://www.youtube.com/watch?v=K8FlwpaTSl4&feature=youtu.be

© 2019 Nuance Communications, Inc. All rights reserved.

Thank you.

© 2019 Nuance Communications, Inc. All rights reserved.