

自然言語処理技術に関する動向

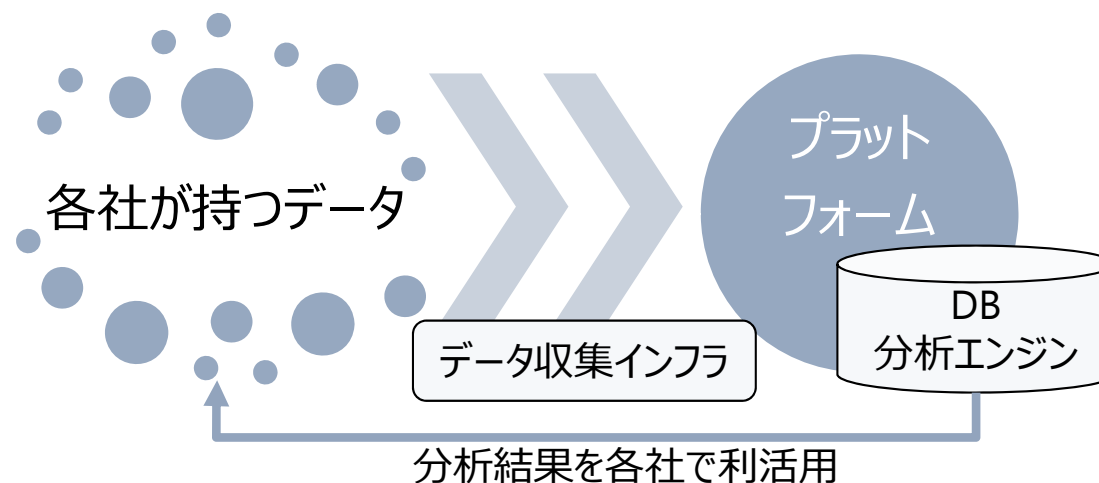
技術戦略委員会向け資料

2016年12月13日版

MRI 株式会社三菱総合研究所

データ自体のプラットフォーム化

- Google, Amazon等の出現以降、データを集め、持つこと自体が価値となり、データを持つものがプラットフォーマーとしての地位を確立しつつある。こうした流れは、所謂インターネットの世界に留まらず、IoT/CPSの世界にも拡大しつつある。例えば、GEなどは、航空機をリースとして貸出し、そのエンジン稼働データを一元的に集約・管理することにより、アフターマーケット市場で高いシェアを確保している。
- こうした事業者は、従来データを他社より早く、多く集めることで競争優位を確立してきたが、近年では、集めたデータを活用可能な基盤としてAPIの開放を進めてきており、エコシステムの拡大、取り込みを進めている。さらに、APIを通じて利用されるサービスに係るデータも、自社データとして取り込むことでデータの拡大・拡充を図っており、更なる強固なプラットフォーム構築を進めている状況である。



ユーザ	データ種別	効果	プラットフォーマー
航空会社	エンジンの利用状況 (センサデータ)	故障予知 燃費改善	GE
販売店・メーカー	閲覧・購買履歴 コメント	潜在的な購買者への推薦	Amazon
利用者	検索履歴 メール	適切な広告表示 検索キーワード候補	Google

Web（キーワード） ⇒ IoT/CPS ⇒ 自然言語（意味解釈）へ

データの収集・分析の主戦場が、2000年代のWeb、2010年代前半のIoT/CPSから自然言語にシフトしつつある状況。

企業名	最近の動向概要
Amazon	<ul style="list-style-type: none"> • Alexaで音声認識、自然言語理解の機能を提供。 • 総額1億ドルのAlexaファンドを用意して、対話AIを活用するベンチャ企業に投資。 • クラウドプラットフォームAWSのAmazon AIサービスとして、画像認識Rekognition、テキスト音声変換Polly、音声認識と自然言語理解Lexを2016年12月に発表。
Google	<ul style="list-style-type: none"> • 従来より機械学習を用いてフレーズベースでの翻訳を実現していたが、文やフレーズの特徴量を捉えて翻訳を行うグーグルニューラル機械翻訳（GNMT）を2016年9月に中国語から英語の翻訳で導入し、11月には日本語を含む複数の言語から英語の翻訳にも導入した。独特の誤訳の問題は指摘されるも、総じて精度は向上したほか、雑音環境下での音声認識の誤認識が減少する効果もあった模様。 • 同社のクラウドサービス上で、クラウド自然言語API（Cloud Natural Language API）のベータ版の提供。センチメント分析（ネガポジ分析）、表現抽出、シンタックス解析などの機能が利用できる。 • その他、クラウドサービスで学習済みの機械学習API、クラウド・スピーチAPI（Cloud Speech API）や視覚API（Vision API）、翻訳API(Translate API)なども提供している。クラウドサービスを利用する企業単位にカスタマイズして利用することが可能となっている。
Facebook	<ul style="list-style-type: none"> • 2013年9月、シリコンバレーに研究所「Facebook AI Research」を設立。所長はYann Lecun氏。2015年6月パリにも研究所を設立。 • テキストメッセージサービスMessengerを通じてパーソナルアシスタント「M」の試験提供。 • チャットボットが接続できるようにAPIを開放。
Baidu	<ul style="list-style-type: none"> • 2014年5月、約300億円を投じてシリコンバレーに研究所「Baidu Research」を設立。所長はAndrew Ng氏。 • 2016年1月14日、深層学習の簡易化・高速化を目指したオープンソースの人工知能ソフトウェア「Warp-CTC」をリリース。 • 医者にコンタクトする前に簡単な診察、診察の予約を会話で行えるチャットボットMelodyを開発。現在は中国向けにサービス。
Microsoft	<ul style="list-style-type: none"> • 人工知能研究プロジェクト「Project Oxford」として画像、音声、テキスト用の各種認識APIのサービス提供を始める。 • Windows 10、スマートフォン用に音声認識によるアシスタント機能Cortanaを提供。使いこむことでより個人向けにカスタマイズされる。 • 高校生の人格を持つチャットボット「りんな」は、中国語用対話エンジンXiaoIceをベースにしており、過去の膨大な会話データや、約1000万冊分の書籍データ、ファッションの種類や素材などの情報を学習。
IBM	<ul style="list-style-type: none"> • 人工知能「Watson」の商業化を推進し、ユーザ企業との連携を進めている。主に、窓口やコールセンターの受付支援等に利用。 • Watsonの日本語化も進めており、コグニティブ・サービスと呼んでいる6つの機能、自然言語分類、対話、検索およびランク付け、文書変換、音声認識、音声合成のAPIが日本語化されている。 • Watsonの機能を含むAI関連機能Bluemixとして、各種APIをPaaSで提供。

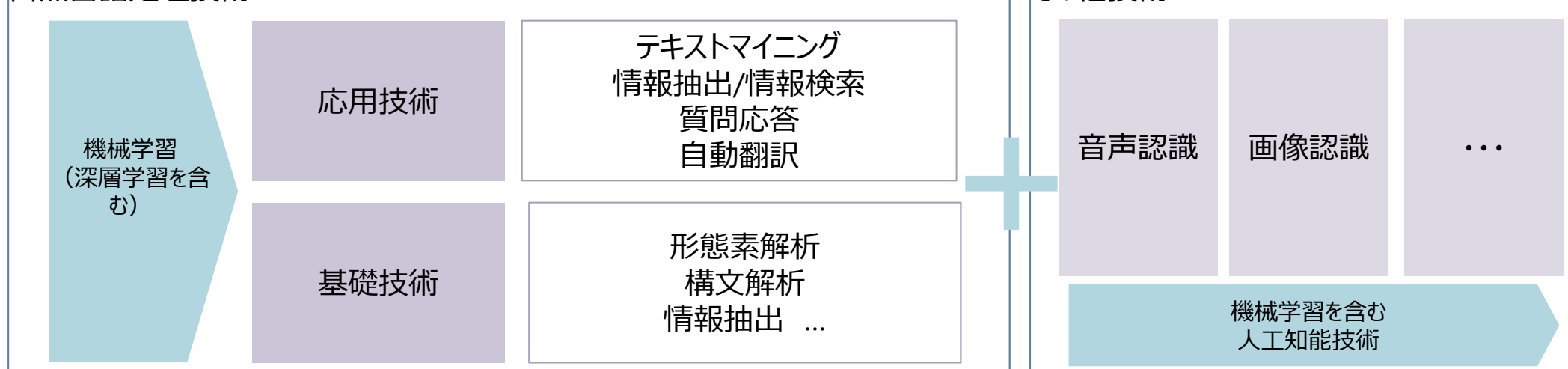
自然言語処理・理解技術に基づく関連サービス

- 自然言語処理・理解技術、その他技術に基づいて開発が進んでいるサービスの俯瞰を以下に示す。

実用化されているサービス

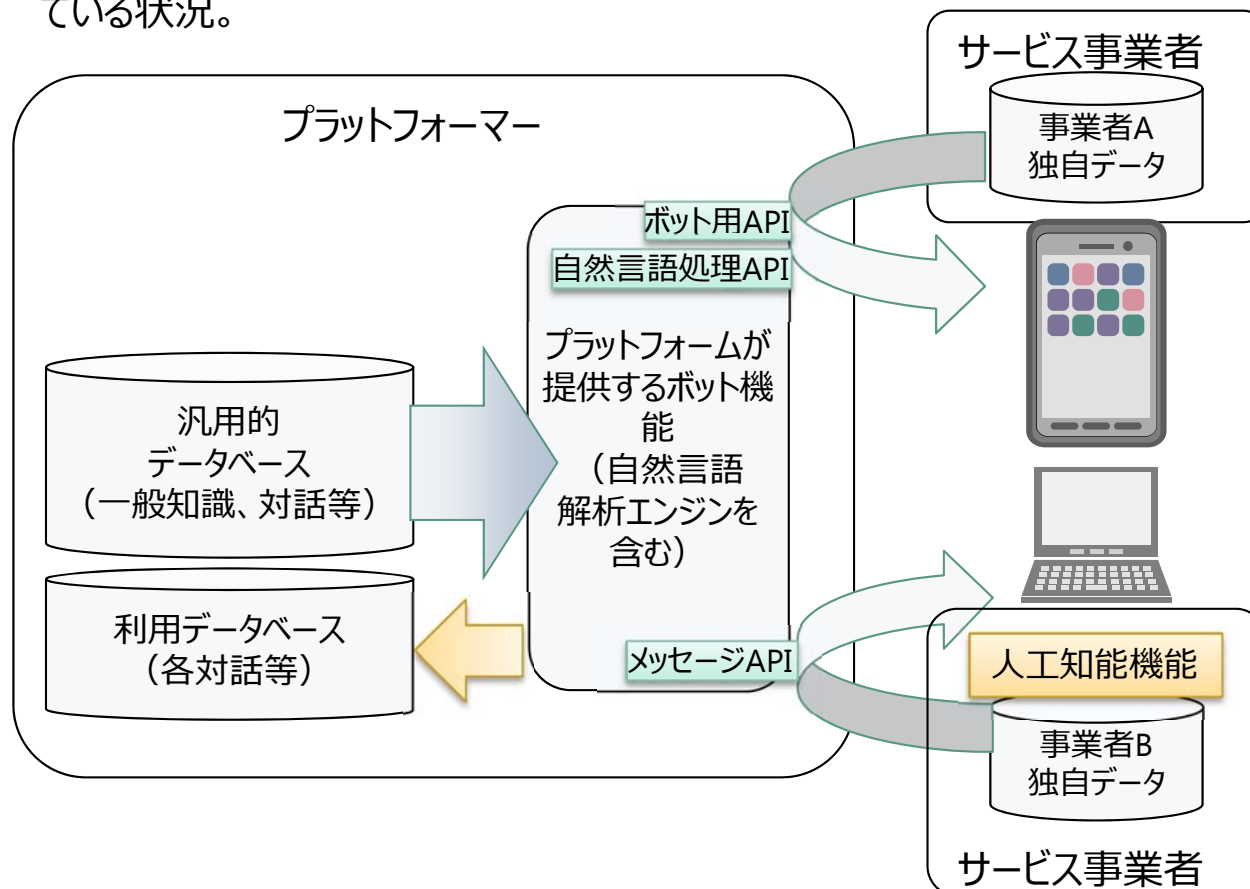
チャットボット	パーソナルアシスタント	質問応答	その他
人間に代わって自動でコミュニケーションをとることができるサービス。	ユーザーとの会話あるいは行動履歴等により、ユーザーの求める情報を検索・レコメンドするサービス。	医者による医療診断支援、受付業務対応、コールセンターVOC解析など、顧客や業種に特化した応答を行うサービス。	マーケティング等への応用 ...
【サービス事例】 <ul style="list-style-type: none"> ● りんな/日本マイクロソフト ● パン田一郎/LINE ● SwingBot/BHI 	【サービス事例】 <ul style="list-style-type: none"> ● Echo/Amazon ● Google Assistant/Google ● Cortana /Microsoft 	【サービス事例】 <ul style="list-style-type: none"> ● Watson/IBM ● Web受付/metadata 	【サービス事例】 <ul style="list-style-type: none"> ●

自然言語処理技術

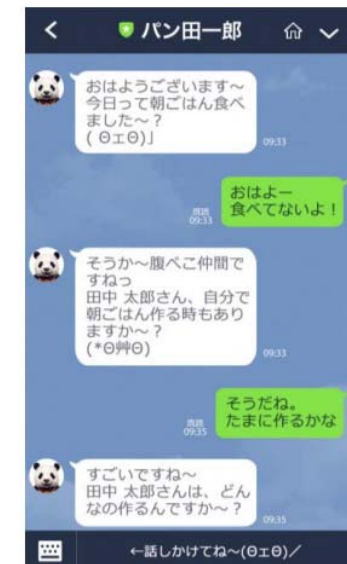


自然言語処理活用サービス：チャットボット（チャット+ボット（ロボット））

- 人間の代わりに自動で対話型コミュニケーションを行うサービスの総称であり、様々な事業者が、様々な形で商用化を推進中。対話型サービスエンジン提供企業によって構築されたデータベースと自然言語解析エンジンに、ユーザー企業が独自のデータを投入することで、企業独自のサービスを構築可能に。
- 多くの事業者は、プラットフォームとしてのエコシステムを構築するために、APIを第三者に対して開放・提供しており、第三者は容易にボットをサービスに組み込むことが可能になり、様々な応用分野への展開が急速に進んでいる状況。サービス事業者は人工知能機能を使うことも可能。
- なお、APIを介してやり取りされたデータはプラットフォームにも蓄積され、結果的にプラットフォームの更なるデータ拡充に貢献している状況。



チャットボット事例
リクルートジョブズ LINE公式アカウント「パン田一郎」



今日の天気など、何気ない普段の話をしながら、アルバイト探しがLINEやFacebookのインタフェースを通じて友達と話すようにできる。

自然言語処理活用サービス：質問応答

- あらかじめ特定分野に関して、知識となるデータ（製品に関するFAQや論文など）を学習しておくことで、入力された質問に対して意図を理解して質問を変えるなど対話を誘導するなどして、対象範囲を限定していくことにより適切な答を導き出す。
- コールセンターでの受付業務支援や医療分野の診断支援などの適用事例が増えている。
- 診療等の業務での活用を繰り返すことにより、更なる高機能化とデータの収集を進めている。
- 外国のAIシステムの場合は、個人情報保護（情報の国外送付）、日本語化されていない等の検討課題がある。

Watsonの医療診断への活用事例

導入事例	大塚デジタルヘルス/桶狭間病院	東京大学医科学研究所
導入時期	2016年 7月	2015年 7月
概要	<ul style="list-style-type: none"> ● 精神科に特化した電子カルテ分析ソリューション「MENTAT（メンタット）」をリリース ● MENTATでは、精神科の電子カルテのデータを、匿名化したうえでクラウドで解析し結果を返す。 ● IBM Watsonの自然言語処理技術により、電子カルテの自由文という非構造化データから、患者の入院長期化や再発に影響を及ぼす情報を抽出。 ● ファーストユーザーで開発にも参画した桶狭間病院（愛知県豊明市）では、すでに診療現場で活用。患者の状態がわかりやすく可視化されたことにより、医療スタッフと患者とのコミュニケーションが促進され、患者満足度も向上した 	<ul style="list-style-type: none"> ● 急増する論文とシーケンスデータに対応するために、「Watson for Genomics（WfG）」を導入 ● 2016年8月、Watsonが60代の女性患者の正確な白血病の病名をわずか10分で見抜き、病名から割り出した適切な治療法によって患者の命を救ったと発表 ● この患者は急性骨髄性白血病と診断され、半年間は2種類の抗がん剤の治療を受けていたが、病状は改善しなかった。Watsonは、2000万件以上のがんに関する論文を学習した結果、治療法を模索していた医師に正しい病名（二次性白血病）と治療法をアドバイスし、数カ月で患者の病状を快方に向かわせた。

自然言語処理を活用したサービス

- プラットフォームにデータを収集していく仕組み（PFが提供する機能をAPIを通じて利用）を構築することで、データの蓄積→人工知能高度化のエコシステムが構築。特に自然言語処理の分野では、海外事業者が構築、提供するプラットフォーム上のサービスを利用することで、属性情報を含む個人の大量データの収集することによる人工知能の開発が急速に進みつつある。
- 一方、我が国においては、NICTのVoiceTra等、従来より分野特化型で高品質なデータを整備しており、これをベースとした人工知能の開発が進められている。

事例：音楽再生やコンシェルジュ機能など、生活全般に対応したアシスタント機能を提供。

Alexa【Amazon】

- 家庭用音声アシスタント端末「Echo」に採用しているクラウドベースの音声認識。
- Echoは無線LANやBluetooth機能を搭載した、円筒型デバイスであり、上面に7基のマイクを搭載し、常時オン状態となっている。音声で「Alexa」と話しかけることにより、Amazon MusicやPrime Musicなどと連携した音楽の再生や、オーディオブックの読み上げ、Googleカレンダーに登録したスケジュールの確認、Amazonの買い物履歴を参考にしたAmazonの再注文等が行える。
- Alexa Skills Kitを活用することによって、Alexaのプラットフォームに独自の対応機能（skill）を追加することが可能。2016年10月時点で3,000以上のskillが利用可能である。

売上台数
5百万台（発売開始2年間）



【出典】
<https://developer.amazon.com/public/solutions/alexa/alexa-voice-service/getting-started-with-the-alexa-voice-service>
<http://www.geekwire.com/2016/amazon-echo-sales-reach-5m-two-years-research-firm-says-google-competitor-enters-market/>
<http://phx.corporate-ir.net/phoenix.zhtml?c=97664&p=irol-newsArticle&ID=2216758>

事例：観光分野における固有名詞・固有表現にも対応し、高度な音声翻訳機能を提供。

VoiceTra【NICT】

ダウンロード 46.3万回
利用回数翻訳1046万回
(2016.11月末)



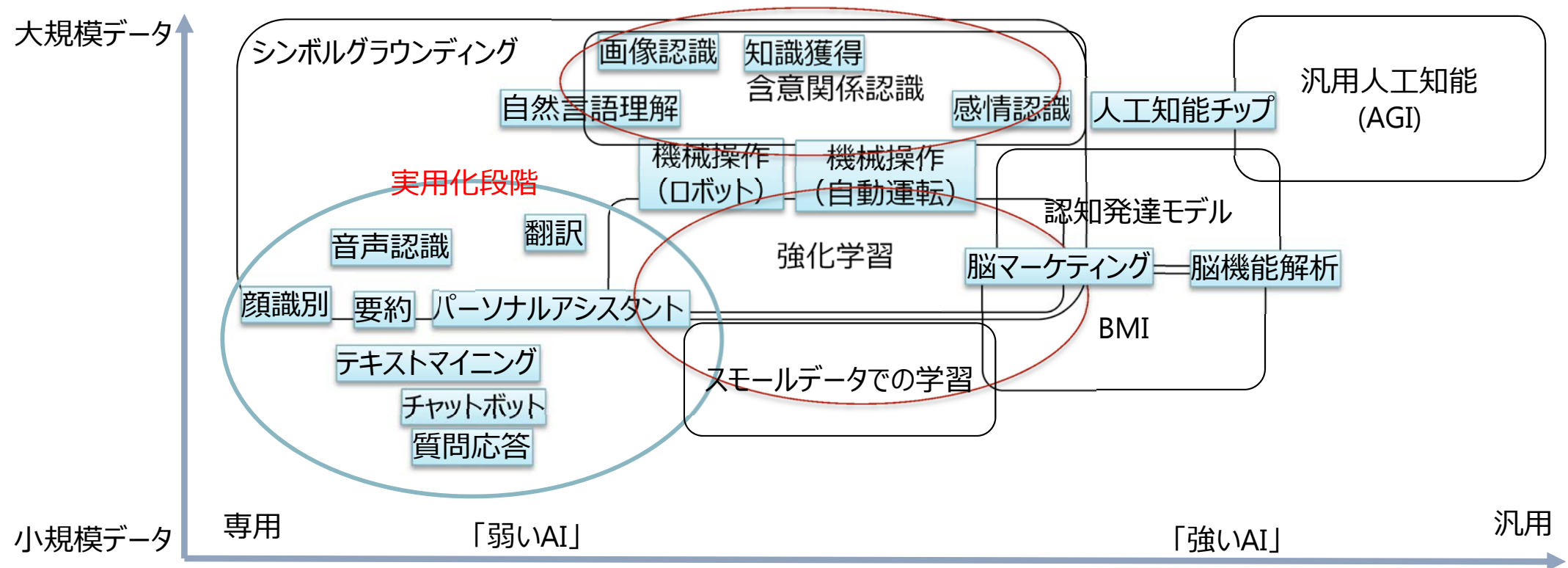
- 多言語音声翻訳アプリであり、旅行会話を中心として日本語を含む31言語の翻訳に対応。
- 全国の駅名や観光に関する表現など、固有名詞・固有の表現を充実させている。
- 2016年10月には、ヤマハ株式会社の提供する音のユニバーサルデザイン支援システム「おもてなしガイド」と連携し、「おもてなしガイド」のアプリからVoiceTraへのシームレスなアクセスを可能にする実験的な連携を開始。
- NICTの多言語音声翻訳システムは民間企業等にライセンスを実施しており、技術移転が進んでいる。目的に応じた単語・表現を翻訳システムに追加することや、秘匿性や防災等の観点からの独自サーバーの設置など、様々な形態での製品・サービスが実現しはじめている。



<https://www.nict.go.jp/press/2015/10/22-1.html>
<http://omotenashiguide.jp/>

人工知能技術の整理（2016年現在）

- 人間の脳機能を実現することで人工知能を実現する「強いAI」、機械学習等により人間の能力を拡大・補完する人工知能を「弱いAI」と呼ばれている。現在、実用化段階にあるサービスの多くは、「弱いAI」による専用AIであり、「強いAI」は今後の活発な開発競争が想定される。
- 現在の主な技術的な課題
 - シンボルグラウンディング（記号接地）：記号的に表現された概念を実世界の観測データと結びつける問題
 - 含意関係認識：ある概念が、別の概念を論理的に含意するか否かを判断する意味処理問題。
 - スモールデータでの学習：少ない量のデータから複雑な課題を学習するために、課題についての事前知識や、類似した課題の学習結果を利用する。

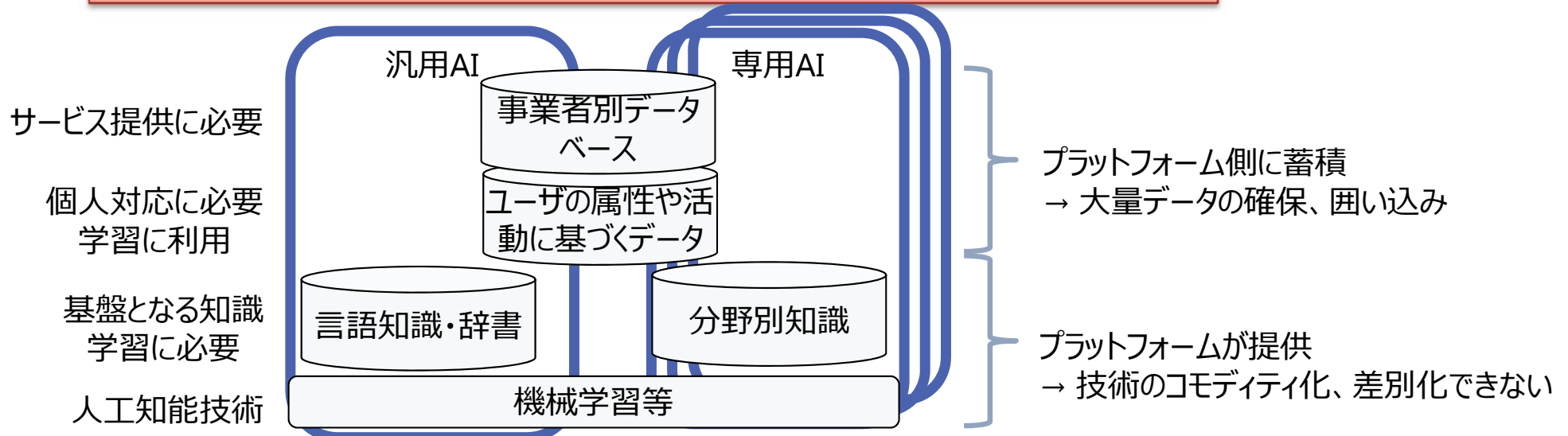


自然言語処理技術の社会実装に向けた課題

- AIの性能向上には、品質の高い学習データ、あるいは、大量のデータが必要である。
 - これまではコストをかけて品質の高い学習データを作ることによってAIの性能向上が行われてきた。
 - 現在、海外の大企業が低コストでプラットフォームに集まる大量データを収集、蓄積している。日本は、プラットフォームを取れていないため自動的にデータを蓄積できる環境になく、蓄積されるデータ量の観点で海外プラットフォーム事業者と圧倒的な差がある。
 - また、データが特定のプラットフォームに蓄積されることで、プラットフォーム事業者の方針に左右される可能性がある。
- 汎用AI技術は、海外事業者がオープンソース化やクラウド上のAPIとして提供することで活用が容易に（コモディティ化）なり、性能だけではなく、入手しやすさや使いやすさでの選択（デファクト化）される可能性がある。

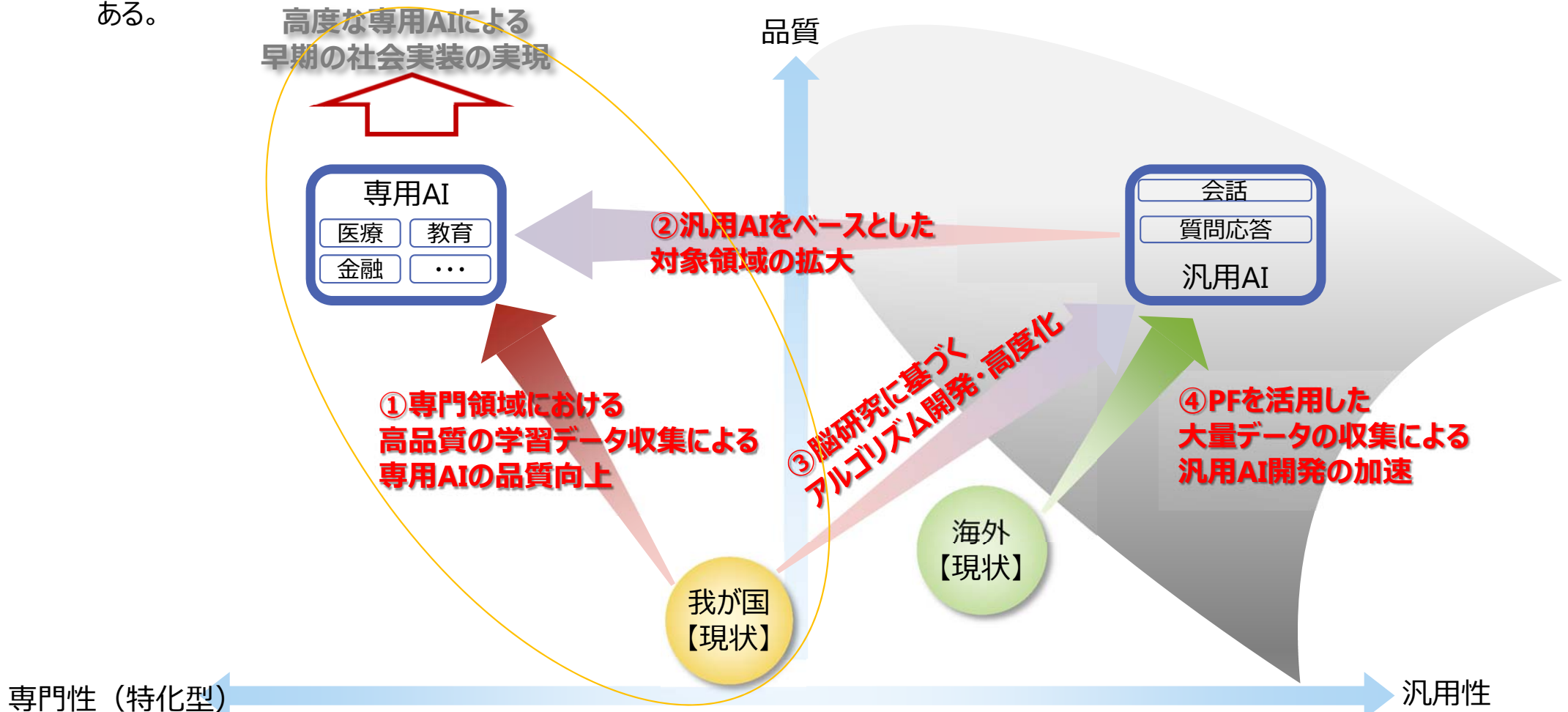
【海外事業者の動き】

- クラウド事業者によるAI技術のクラウドサービス化
- API提供による各種サービスへのAI導入、連携の促進
- 多数のユーザ、サービスが利用することによるプラットフォームでのデータの囲い込み
- 大量のデータに基づく日本語を含む多言語での自然言語処理技術の性能が向上



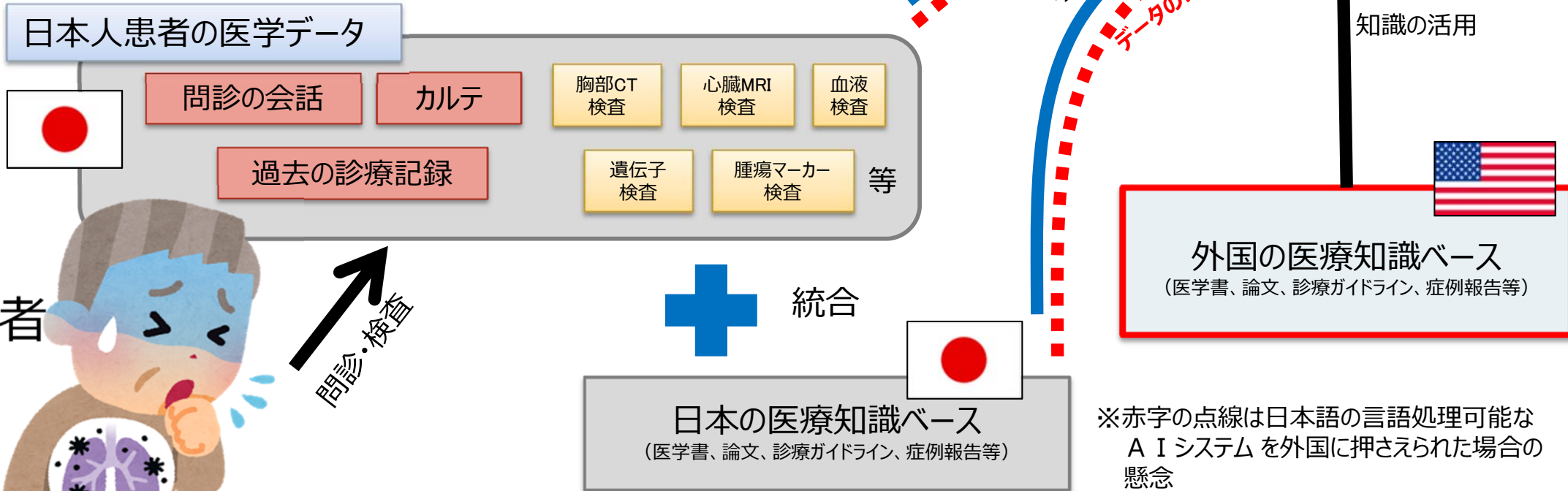
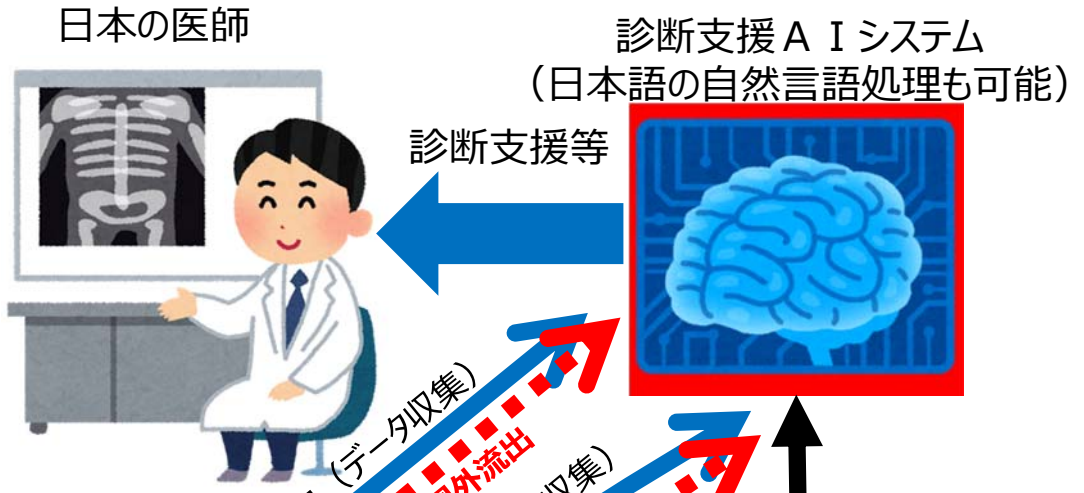
人工知能技術の社会実装に向けた方策

- プラットフォームを取れていない我が国の人工知能技術の推進方策は、①専門領域における高品質の学習データ収集による専用AIの品質向上 と、これに続けて②汎用AIをベースとした開発、提供を通じたノウハウを生かして、対象領域を拡大していくことで、専用AIを実現することが、方向性として考えられる。ただし、技術がコモディティ化しないように③脳研究の成果を活用したアルゴリズム開発やプラットフォームに入っていないデータの確保にも力を割くべきである。
- 一方、良質なデータのみを学習データとした場合、ノイズに弱かったり、模範解答しかできない（例えば、医療分野において革新的な治療法を提案できない）AIになる危険性があるため、汎用型AIの特徴もうまく踏まえたAIに昇華させることが必要である。



自然言語処理技術への期待（例：医療分野）

- 日本の医療知識ベースをA Iシステムに学習させ、個々の患者の医学データを入力することで医師の診断を支援
⇒ 効率的・効果的な医療の実現
- 一方で、日本語の言語処理可能なA Iシステムを外国に押さえられると、日本人の医療知識ベース・医学データが国外に流出する懸念
⇒ 将来の高度医療や創薬のシーズをとる情報を外国に押さえられる懸念
- したがって、欧州でも欧州共通の最重要プロジェクトとして自然言語処理の研究開発プロジェクトを推進



データの戦略的な確保（データビリティ）の重要性（例：医療分野）

将来（20XX年）における高度医療
～肝臓がんの病院ランキング～

順位	病院名	手術数	手術後の 5年生存率	X社の AI診断支援 システムの有無
1	国立〇〇病院	400	50%	○
2	国立△△センター	350	48%	○
3	□□大学病院	300	45%	○
4	赤十字◇◇センター	280	43%	○
5	⋮	⋮	⋮	○
6	⋮	⋮	⋮	○
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	×
⋮	⋮	⋮	⋮	×
⋮	⋮	⋮	⋮	×

- ・日本を含め全世界の最新の症例、医学論文を瞬時に解析して、最適な治療をしてくれるから安心！！
- ・やっぱりX社のAIシステムが入っていない病院は不安ね！

高度医療を行う病院は
全てX社の
AIシステムを
利用
⇒最新データは
X社が囲い込み



★日本人の医療DB・医学データ等を外国企業に独占（国外流出）された場合の懸念

- 医療の効率化・高度化を図るために不可欠な医療DBが特定の外国企業に囲い込まれる
- 日本全体の高度医療が特定の外国企業にロックインされる
- 創薬や予防医療等の日本の医療産業の発展の道が閉ざされる 等

以下、参考事例

クラウドAPI事例：Project Oxford (Microsoft)

- 顔認識、音声処理などの機械学習アルゴリズムを使ったアプリケーションを開発者が構築できるようにAPIを提供。
- 多様な応用が可能なコア技術を提供し、利用者（開発者）増加を目指す。

Project Oxfordの提供API一覧

Project Oxford		概要
Vision	Computer Vision APIs	カテゴリ分類、サムネイル生成、OCRなどを行う画像解析APIの1つ。画像のURLからクラウドで処理を行い、画像内の何が主体かを判別する機能に特徴がある。
	Face APIs	人間の顔が含まれた画像から顔部分を抽出し、その人の年齢や性別を計算するAPI。顔の位置や数にかかわらず高精度な顔認識が可能。同一人物判定は顔認証にも適用可能。
	Emotion APIs	人間の顔画像から、その顔をしている人がどのような感情をしているかを判別。
	Video APIs	動画の手ブレ補正や、物体認識、動き検知などを行うAPI。認識だけでなく、認識結果に基づいた動画編集も可能。
Speech	Speech APIs	音声ファイルから人間の話し声だけを抽出し、抽出した話し声から内容をテキスト文書化するAPI。逆の処理であるテキストを読み上げる機能もある。
	Speaker Recognition APIs	話者認識APIで、複数人での会話音声から、誰が話しているかを特定したり、音声データが本人の声によるものかを判別することが可能。
	Custom Recognition Intelligent Service	ユーザや環境に合わせた自然言語処理によって音声認識を行うAPI。
Language	Spell Check APIs	スペルチェックを行うAPI。タイプミス正しい単語に修正する機能だけでなく、同音異義語などの人間がし易いミスの修正も可能。
	Language Understanding Intelligent Service	自然言語処理によってどのような話題かを判別。専門用語に強い処理モデルなど、文章を理解するモデルを変化させることが可能。
	Web Language Model APIs	単語間のつながりやすさを用いて、熟語が一般的に使われているかの計算や、次に来る確率が高い単語を提案するAPI。

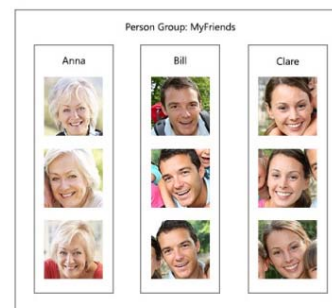
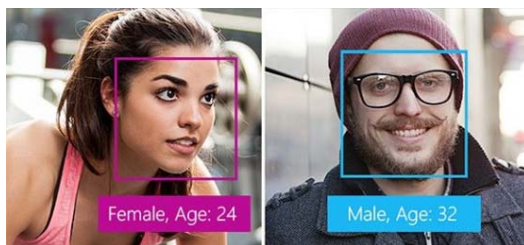
事例 : Project Oxford (Microsoft)

■ 顔認識、音声処理などの機械学習アルゴリズムを使ったアプリケーションを開発者が構築

提供サービス例

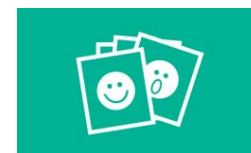
① Face APIs

- 画像内の人間の顔を検出・認識する。
- 顔認識、顔認証、類似顔検索、顔分類、顔識別の機能を提供。



② Emotion APIs

- 画像内の顔の表情から感情を判別する。
- 8つの主要な感情要素である「怒り」「軽蔑」「恐れ」「嫌悪」「幸福」「中間」「悲しみ」「驚き」を数値化。



③ Speech APIs

- 音声認識、会話の意図の認識・理解、文書の音声変換を行う。
- 意図の理解には、Language Understanding Intelligent Service (知的言語理解サービス) が活用されている。



【出典】

<https://www.projectoxford.ai/>

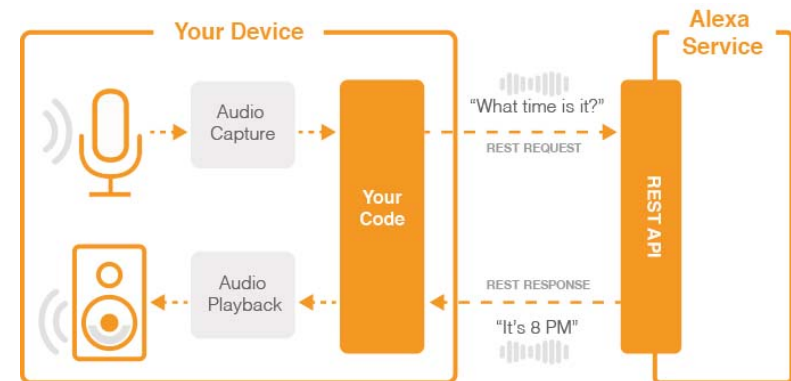
<http://jp.techcrunch.com/2015/05/01/20150430microsofts-project-oxford-gives-developers-access-to-facial-image-and-speech-recognition-apis/>

パーソナルアシスタント事例： Alexa (Amazon.com)

- 家庭用音声アシスタント端末「Echo」に採用しているクラウドベースの音声認識
- Alexaは大量の遠距離スピーチデータによって学習されたニューラルネットワーク（深層学習）に基づいた音声認識システムを使用。遠くの発声でも認識することができる。ノイズキャンセラの強化により、音楽再生中でも質問を聞き取ることが可能。
- Alexaはクラウド上にあり、利用すればするほど、スピーチのパターンや語彙、個人の好みを学習する。
- デバイス（Echo）は常にインターネットに接続しているため、学習した結果は自動的にクラウドへ送られる。

AlexaのAPI「Alexa Voice Service」をハードウェアベンダーに公開。スピーカーとマイクを備えたインターネット対応端末をAlexaと統合できるようにする。

- ソフトウェア開発者には、SDK「Alexa Skills Kit（ASK）」を提供する。既存のWebサービスとAlexaを統合し、Alexa対応の新しい音声機能を構築可能。
- 他社が自社の機器にアレクサのソフトを導入する取り組みを資金面で支援するため、1億ドルを拠出して「Alexa Fund」を設立。
- フォード（米）が第3世代車載情報システム「SYNC3」と、EchoおよびAlexaを連携。車内からEchoに接続し、自宅の照明や、室温を調整するサーモスタット、セキュリティーシステム、ガレージドアなどを音声で操作したり、これらの機器の状態も確認できるようにする。



【出典】

<https://developer.amazon.com/public/solutions/alexa/alexa-voice-service/getting-started-with-the-alexa-voice-service>

家庭用音声アシスタント端末「Echo」

- 無線LANやBluetooth機能を搭載した、円筒型デバイス。
- 上面に7基のマイクを搭載し、常時オン状態となっている。
- 音声で「Alexa」と話しかけることにより、Amazon MusicやPrime Musicなどと連携した音楽の再生や、オーディオブックの読み上げ、Googleカレンダーに登録したスケジュールの確認、Amazonの買い物履歴を参考にしたAmazonの再注文等が行える。

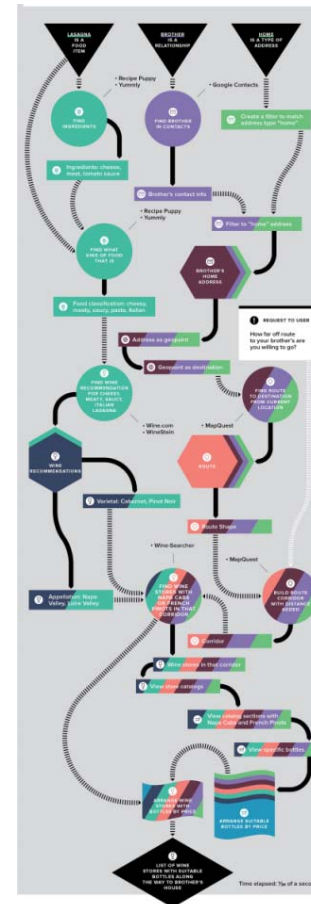
質問応答事例：Viv

- 2012年に米Appleの「Siri」開発メンバーDag Kittlaus氏が立ち上げたViv Labsによる、自然言語の音声認識をベースとしたオープンなAIアシスタントのプラットフォーム。Siriは連携できるアプリに制限があるのに対し、**内容を理解してWeb上の情報を返す**だけでなく、**サードパーティーのサービスとも連携し**、複合的な問いかけに対応することが可能。
- Viv Labsは、開発者やハードウェアベンダーに対し、Vivとの連携を呼びかけているが、これまでに製品化はされていない。
- 2016年10月にSamsungが買収

Vivに
「兄弟の家に行く途中に、
ラザニアとそれに合う安めのワインを
買ってきたい」と話しかけると、

**On the way
to my
brother's
house,
I need to
pick up
some cheap
wine
that goes
well
with lasagna.***

*VIV CAN HELP



「ラザニアの材料を調べてどんな料理なのか認識」
→「チーズ・肉・ソースが入ったイタリア風ラザニアに合うワインを照会」
→「ラザニアとワインが買えるお店を検索」

というプロセスと同時に、

「連絡先から兄弟の住所をピックアップ」
→「ルートを検索」
→「目的地までの道中にラザニアとワイン購入地点を組み込む」

といった複雑な処理を行う

<https://www.wired.com/2014/08/viv/>