

次世代人工知能社会実装WG の検討状況について(案)

事務局

次世代人工知能社会実装WGで取り扱う分野

情報通信審議会技術戦略委員会「次世代人工知能社会実装WG」では、自然言語処理技術、脳情報通信技術等の次世代人工知能技術に関する取組の現状と課題を把握し、今後の研究開発及び社会実装に向けた推進方策等について調査、検討を行う。

氏名	所属・役職
(主任) 柳田 敏雄	国立研究開発法人 情報通信研究機構 脳情報通信融合研究センター (CiNet) センター長
東 博暢	株式会社日本総合研究所 主席研究員/融合戦略グループ長
麻生 英樹	国立研究開発法人 産業技術総合研究所 情報・人間工学領域人工知能研究センター 副センター長
荒牧 英治	国立大学法人 奈良先端科学技術大学院大学 特任准教授
池田 尚司	株式会社日立製作所 研究開発グループ テクノロジーイノベーション統括本部 システムイノベーションセンター センター長
上田 修功	国立研究開発法人 理化学研究所 革新知能統合研究センター 副センター長
宇佐見 正士	KDDI株式会社 理事 技術統括本部 新技術企画担当
臼田 裕一郎	国立研究開発法人 防災科学技術研究所 総合防災情報センター センター長
栄藤 稔	株式会社NTTドコモ 執行役員 イノベーション統括部長
大岩 和弘	国立研究開発法人 情報通信研究機構 フェロワー・未来ICT研究所 主管研究員
大竹 清敬	国立研究開発法人 情報通信研究機構 ユニバーサルコミュニケーション研究所 データ駆動知能システム研究センター 上席研究員
岡島 博司	トヨタ自動車株式会社 先進技術統括部 主査 担当部長
加納 敏行	日本電気株式会社 中央研究所 主席技術主幹
川鍋 一晃	株式会社国際電気通信基礎技術研究所 脳情報通信総合研究所 主幹研究員
栗本 雄太	株式会社三井住友銀行 成長産業クラスター 第四グループ長 (新分野・企画運営)
小林 哲則	学校法人 早稲田大学 理工学術院 教授
相良 美織	株式会社バオバブ 代表取締役社長
鳥澤 健太郎	国立研究開発法人 情報通信研究機構ユニバーサルコミュニケーション研究所 データ駆動知能システム研究センター センター長
萩原 一平	株式会社NTTデータ経営研究所 研究理事 情報未来研究センター長
原 裕貴	富士通株式会社 執行役員
春野 雅彦	国立研究開発法人 情報通信研究機構脳情報通信融合研究センター (CiNet) 脳情報通信融合研究室 主任研究員
本田 英二	富士ソフト株式会社 執行役員
前田 英作	日本電信電話株式会社 NTTコミュニケーション科学基礎研究所 所長
森川 幸治	パナソニック株式会社 先端研究本部 インタラクティブAI研究部 副主幹研究員
八木 康史	国立大学法人 大阪大学 理事・副学長
山川 宏	株式会社ドワンゴ ドワンゴ人工知能研究所 所長

※ 平成29年4月7日時点。このほか、内閣府、文部科学省、経済産業省からオブザーバ参加。

次世代人工知能社会実装WGにおける検討事項(例)及び論点

次世代人工知能社会実装WGにおける検討事項(例)

(1) 自然言語処理技術の社会実装の推進方策

- 自然言語処理技術の活用分野例(導入事例、技術活用への期待)
- 自然言語処理技術に係るデータの取扱い
- 主要分野(防災、医療、対話、翻訳等)における社会実装の推進方策(技術課題、ロードマップ)

(2) 脳情報通信技術の社会実装の推進方策

- 脳情報通信技術の活用分野例(導入事例、技術活用への期待)
- 脳情報通信技術に係るデータの取扱い
- 主要分野(マーケティング、設計生産等)における社会実装の推進方策(技術課題、ロードマップ)

(3) その他の分野の推進方策

- 自然知×人工知能 等

次世代人工知能社会実装WGにおける主な論点

- 自然言語処理技術及び脳情報通信技術の国内外の導入事例を調査し、技術の実用化への期待等について検討を行い、利用分野及び研究開発・社会実装に向けた諸課題を整理する。
 - ・どのような利用分野を優先して考えるべきか。
 - ・どのような技術課題が解決されれば、実用に期待が持てるのか。
- 自然言語処理技術や脳情報通信技術に関する「データの取り扱い(作成、収集、公表(外部共有)、管理)」等について、国・研究機関と民間企業等との間のデータ流通を活発にするための方策を議論する。 このためプレゼンでは、データを持つ民間企業や、データを使いたい民間企業から国・研究機関への要望をヒアリングするほか、国・研究機関とこのような民間企業との間で意見交換を進める。
 - ・データ流通を活性化させ、研究開発を促進させる仕組みはどうあるべきか。
 - ・国・研究機関・企業の双方にメリットがあり公表(外部共有)できるデータ(今後作成すべきデータ)は何か。
 - ・データの作成、収集、公表(外部共有)、管理体制はどうあるべきか。

※ データの著作権等の知財的な扱いに関しては、他の会議(内閣官房知財事務局の「新たな情報財検討委員会」等)にて検討中であるため、本WGの主な論点とはしない。

事前会合(自然言語処理と脳情報通信へのユーザーの期待)

- 技術戦略委員会の今後の検討方向について
- 自然言語処理技術及び脳情報通信技術に関するNICTの取り組み
- ユーザー・開発企業の取組み・課題、WGへの期待
 - ・自然言語処理技術 活用事例紹介(NEC) ・トランスコスモスにおける自然言語処理や人工知能に関する取り組み
 - ・PKSHA Technologyの取組について ・バオバブにおける自然言語処理関連データ構築の取り組みと課題、次世代人工知能社会実装WGへの期待
 - ・FRONTEOの人工知能KIBITとその社会実装について ・トヨタ自動車における取組み
 - ・NTTデータグループの取組みのご紹介 ・パナソニックにおける脳情報通信関連の取組

第1回会合

- 構成員等からのプレゼンテーション
 - ・国内外における自然言語処理・脳情報通信の動向について(三菱総合研究所谷田部様) ・脳情報通信の現状と今後の展望(大岩構成員)
 - ・自然言語処理応用の現状と今後(栄藤構成員)
 - ・バオバブにおける自然言語処理関連データ構築の取り組みと課題、次世代人工知能社会実装WGへの期待(相良構成員)
- 次世代人工知能社会実装WGの主な論点について

第2回会合(自然言語処理技術に関する議論)

- 構成員等からのプレゼンテーション
 - ・自然言語処理のためのデータ整備について(鳥澤構成員) ・データ利活用とセキュリティ・プライバシー保護(NICT盛合様)
 - ・自然言語理解の社会実装に向けて(麻生構成員) ・自然言語処理の医療応用(荒牧構成員)

第3回会合(脳情報通信技術に関する議論)

- 構成員等からのプレゼンテーション
 - ・脳情報通信の応用に関するATRの取り組み(川鍋構成員) ・脳情報技術の産業利用と課題(加納構成員)
 - ・人工知能と精神医学・医療 ―社会実装に向けて―(大阪大学橋本様) ・ダイキン工業におけるIoTの取組(ダイキン工業前田様)

第4回会合

- 構成員等からのプレゼンテーション
 - ・中国における人工知能の社会実装に向けた動向について(三菱総合研究所谷田部様)
 - ・AIの社会実装・事業化に向けたオープンイノベーションの重要性について(東構成員)
 - ・消費者コミュニケーション領域におけるAI/自然言語処理に関する現状と課題(トランスコスモス北出様)
 - ・深層学習の社会へのインパクト(Preferred Networks丸山様) ・高品質データベース構築と個人情報を含むデータの利活用における課題(八木構成員)
- 次世代人工知能社会実装WGの検討状況について(案)

(1) 自然言語処理技術及び脳情報通信技術の国内外の導入事例

自然言語処理技術の国内外の導入事例

カテゴリ	概要	導入事例
質問応答	<p>➤ Web上のデータや、顧客の保有するFAQ、問合せ履歴等のデータを活用し、音声認識技術と組み合わせることによるコールセンター業務の高度化、情報の抽出・集約等の効率化を実現するシステムが企業で実運用されるケースが増えつつある。</p>	<ul style="list-style-type: none"> ● NEC：声の高さや声色などから相手の感情を認識する「会話解析」の技術によりコールセンター等の高度化を実現。さらに、テキスト含意認識技術によって、要管理案件等を抽出することで業務を効率化。 ● トランスコスモス：コールセンターへのユーザーの問い合わせ内容を分類、さらに現状のFAQや注文書などのカバー率を分析し、FAQや注文書、カタログの改善に活用。 ● PKSHA Technology：汎用型対話エンジン「BEDORE」 コールセンター等の一次受付をAIが担うことによって効率化を図る。AIでは対応しきれないケースでは、有人オペレータへ繋ぐことも可能。 ● NICT：大規模Web情報分析システム「WISDOM X」、対災害SNS情報分析システム「DISAANA」 WISDOM Xは、Webの40億ページの情報を基に様々な回答を整理して表示。DISAANAは被災状況に関する質問にツイッターをベースに回答。
翻訳	<p>➤ 多言語音声翻訳の精度が向上。訪日観光客の増加に対応すべく、観光や防災の分野で様々な事業者と提携し、実用化が進められている。</p>	<ul style="list-style-type: none"> ● NICT：多言語音声翻訳アプリ「VoiceTra」 実サービスとして、成田国際空港が公開する多言語音声翻訳アプリ「NariTra」やKDDIが提供する“おはなしアシスタント”の音声翻訳機能として採用されている。 ● Google：翻訳サービス「Google翻訳」 100以上の言語に対応。2016年9月、新たにGNMTと呼ばれるディープラーニングを用いた技術を導入し、精度を向上。
チャットボット	<p>➤ 2016年以降、FacebookやMicrosoft、LINE等がチャット開発環境の提供を始めたことにより、様々な企業や開発者がチャットボット開発に取り組み、急速に拡大。</p>	<ul style="list-style-type: none"> ● Facebook：「Facebook Messenger Platform」 チャットボットの開発が可能なMessengerプラットフォーム。独自の技術に加え、買収した自然言語処理技術の開発企業であるWit.aiの技術を活用。 ● Microsoft：チャットボット「女子高生りんな」 2015年7月にLINEアカウントとして登場した人工知能。2016年9月には、この技術とMicrosoft Azureをベースにローソンが独自のLINEアカウント「あきこちゃん」を提供開始。
パーソナルアシスタント	<p>➤ スマホや家電への搭載等により、ユーザーの日常行動データを収集し、嗜好を学習。近年では特にサードパーティによる採用が進むAmazonのAlexaが注目されている。</p>	<ul style="list-style-type: none"> ● Amazon：音声アシスタント内蔵スピーカー「Echo」 クラウドベースの音声認識サービスAlexaを搭載した、スピーカー型スマートデバイス。音楽の再生や質問回答など、様々な機能を果たすことができ、2016年12月時点で既に500万個が販売されたと推定されている。 ● Google：対話型アシスタント機能「Googleアシスタント」 自然言語で会話可能なアシスタント機能。2016年12月、メッセージアプリ「Allo」にて日本語の利用が可能となった。本機能を搭載した家庭用高音質スピーカー「Google Home」も2016年11月に発売開始。
医療	<p>➤ 症例データや論文等の医療関連データの蓄積を診断支援等へ活用できる可能性があり、社会実装に向けて研究開発が進められている。</p>	<ul style="list-style-type: none"> ● NAIST：学会や大学病院等の保有する症例データを活用し、自然言語処理によって診断支援システムを開発。さらに、SNSのデータと医療データを組み合わせ、感染症の発症状況を検出するシステムも検討。 ● FRONTEO：精神疾患患者の支援記録を解析し、症状悪化の予兆を発見するシステムを開発。 ● IBM：2000万件以上の癌に関する論文を学習した人工知能「Watson」が、特殊な白血病患者の病名を診断し、適切な治療法を提示したことで患者の命を救うことに成功。

脳情報通信技術の国内外の導入事例

カテゴリ	概要	導入事例
医療 (治療)	<p>➤ 脳活動の解明とニューロフィードバックの活用により、脳の状態をコントロールし、主に精神・神経疾患の治療に活用する取組が進められている。</p>	<ul style="list-style-type: none"> ● CiNet他：2016年、被験者にづらい経験を思い出させることなく、記憶によって引き起こされる恐怖反応を弱める技術を開発したことを発表。被験者の視覚野に恐怖記憶の対象を表す脳活動パターンを検出したときに、被験者に報酬を与えることで恐怖記憶を消去することに成功した。 ● 大阪大学：2016年10月27日、BMI技術を活用した義手を開発し、幻肢痛の患者がBMI義手を使うことで、痛みをコントロールできることを発見したと発表。 ● ATR他：精神疾患患者と健常者の脳の差異を示すバイオマーカーの開発、及びその情報をフィードバックしトレーニングすることによる治療への応用に取り組んでいる。なお、NF治療は安全面・倫理面に十分配慮して進めている。
医療 (リハビリ)	<p>➤ ニューロフィードバックのリハビリへの効果を活用したシステムが開発され、実用化に向かいつつある。</p>	<ul style="list-style-type: none"> ● パナソニック：慶応大学との連携により、ワイヤレス脳波ヘッドセットとニューロフィードバックによる、脳卒中のリハビリ機器を開発。現在臨床評価の準備中。 ● ATR：上肢麻痺のリハビリテーションのための簡易外骨格ロボットを開発。脳活動に対してロボットが動作することでリハビリ効果向上を確認、BMI療法として実用化。また、簡易脳波計と、動作を意図する脳活動に反応して動く外骨格ロボットによる運動支援・リハビリのシステムも開発。
マーケティング	<p>➤ 企業が自社広告や製品に対する消費者の印象を分析する際、アンケートでは取得できない深層心理の把握手段として脳情報の活用が期待されている。</p>	<ul style="list-style-type: none"> ● NTTデータ他：TVCMなどの動画広告・コンテンツに対する視聴者の評価を、視聴中の複雑な脳活動のパターンから可視化する脳情報解読技術を開発。2016年8月、マーケティング支援サービスとして実用化。 ● ATR：fMRIのデータから視覚情報を解読する研究を進めている。ニューロ・マーケティング、ニューロデザインへの応用が期待されている。
生活支援	<p>➤ 主に日常生活が困難な患者を対象とした、意思を把握するためのシステムが開発されている。</p>	<ul style="list-style-type: none"> ● 産総研：患者が画面上に表示された意思伝達メニューの中から選んだ項目を、簡易脳波計から取得した脳波を解析することで推測可能な「ニューロコミュニケーター」を開発。
脳機能の再現	<p>➤ 脳機能を解明し、コンピュータ上で再現することによって、新たなICT技術の創出を目指している。</p>	<ul style="list-style-type: none"> ● NEC：大阪大学と協働し、脳の振る舞いから人間の理解・認知の構造を解明してコンピュータ上で再現することによって、新たなコンピューティングの基盤を目指す。
計測機器	<p>➤ fMRIは高解像度の脳情報が取得できるが、計測機器が大規模・高価であり、日常生活での利用（大量の脳情報収集）は難しいことから、最近では簡易かつ負担の少ない脳情報計測技術の開発が活発化している。</p>	<ul style="list-style-type: none"> ● CiNet：ドライ型電極を装着した携帯型脳波計を開発、市販済み。日常生活空間での脳活動計測を実現。 ● パナソニック：低消費電力化・ノイズの耐性強化等を実現し、日常生活で計測可能なワイヤレス脳波計を開発。 ● 大阪大学産業科学研究所：2016年8月、冷却シートを額に貼るような感覚で装着できるパッチ式脳波センサの開発を発表。リアルタイムに脳状態を可視化し、手軽に睡眠中の脳波を計測する事に成功した。

1. 自然言語処理技術

<現状・課題>

- 自然言語処理精度向上には大量のラベル付き学習データが必要となるが、現状の学習データの規模では不十分。
- 学習データの設計は、実現するサービス、アプリケーションに大きく依存するため、サービス、アプリケーションの開発を行う研究者等が学習データの構築法、設計まで担当。
- 意味を文章の単語の出現頻度で捉えるような手法でもそれなりに質問応答が成立しているように見える。

<主な意見>

- 翻訳にしても音声認識にしても対話にしても、質のよいデータを国の研究機関で集めてそれをコアにして、半教師あり学習等に広げていく研究が有効ではないか。
- 必要とされる自然言語理解のレベルを明確にしてターゲットを設定する必要がある。ターゲットの設定では普通の人考える課題の難しさと、技術としての実現の難しさは必ずしも一致しないことに留意した方がよい。
- 目標とする実サービスやアプリケーション、アーキテクチャを踏まえて学習データを設計しなければ、実運用・社会実装には結びつかない。データ作成者からサービスプロバイダーまで、垂直な連携が必要。
- 特定のシステム向けであればデータも作りやすく、活用しやすいが、汎用的なデータというのは難しく、かえって使いにくいデータになることがあるので、配慮が必要。
- 診断データなど、既にラベル付きの教師データが蓄積されている分野も存在するのではないか。
- 災害データなどデータ数は少ないが、強化すべき分野については、新しくデータを作るためのモデルを構築するなどして、そのデータを標準化・共有化していくことが重要。



<今後の方向性>

- ◆ 社会的な利活用ニーズに応じたアプリケーション・サービスを想定した新たな学習データの整備を推進する。
- ◆ わが国の社会課題解決を目指した医療、防災等の分野を優先し、まずは少量でも質の良いデータを作成・収集した上で半教師あり学習等の人工知能によってデータを有効活用する取組を推進する。ただし、処理の内容(例えば構文解析など)によっては適用範囲の広い汎用のデータを作成することも有効なので留意が必要。

2. 脳情報通信技術

<現状・課題>

- 脳情報の学習データの蓄積量に関しては、まだ海外に大きく差をつけられている状況ではないものの、計測設備等の理由で、学習に必要な数が十分集まっている状況ではない。また、他国にぬきんで、データ取得するには、大きなコストがかかる。
- 既を取得された脳情報の学習データに対して、事後に別のアノテーションを付与することが困難。
- 脳情報データと、比較的簡単に計測可能である生理学データのような一般データとの組合せが有効であると想定されるが、個人情報への壁が存在。
- 脳情報通信を社会実装しようと思うと、人とのインタラクションも重要だが、間接的に取らないといけない内面や意識といったデータを取るところが難しい。
- 従来のBMIは重症患者用として期待されてはいるものの、リハビリテーションなど、より広い市場が期待できる。

<主な意見>

- 実環境での脳情報通信技術の活用には、脳情報データの取得の障壁を下げるためセンサ装置の簡易化が重要。
- 少量データでも学習できる脳型AIの解析技術開発が必要。さらに、脳情報のデータ収集の人材不足も解消すべき。
- 安静時脳活動はデータ取得方法等に対するコンセンサスが取りやすく、データも蓄積されつつある。一方、脳にフィードバックをかけてその影響を見るような場合はライブのデータが必要になるが、取得条件等のコンセンサスが取りにくい。
- どのようなデータを、どのような環境・条件・目的を設定して、体系的にデータベース化していくかの検討が必要。
- 研究側から積極的に利活用を提案・推進したり、研究者と利用者の興味のマッチングが必要。これらを全体最適化するために、研究と市場をタイムリーにつなぐ双方向のシステムが必要。

<今後の方向性>

- ◆ 実環境での脳情報を活用するため、簡易に脳情報を収集できる低コストの計測機器を開発し、質、量ともに十分なデータを体系的に備えたデータベース構築を推進する。
- ◆ 利活用を円滑に進めるため、研究と市場をタイムリーにつなぐ双方向のシステムの構築を推進する。

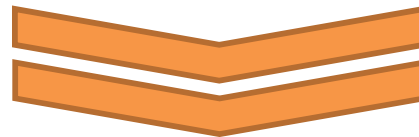
3. その他(最新の共通AI技術のプラットフォーム化)

<現状・課題>

- Amazon Alexaに見られるように、音声認識やチャットボットなどの複数のAI技術をプラットフォーム化し、再利用性を高めてオープン化することによって、急速な社会実装がサードパーティーによって進んでいる。
- 一方、日本ではシステムが個別に作られているため、社会実装としてのスケーラビリティが弱く、技術も陳腐化。

<主な意見>

- 次世代人工知能という観点では、大きく分けて、情報を収集・解析するハードウェア、アルゴリズムの開発を含むソフトウェア、データの量と質を揃え、それを集める仕組みの3点について、それぞれ検討が必要。
- 言語と他のセンサー情報との組合せなど、マルチモーダルなソースやサービスと結びつけた利用分野の検討や、データを収集する仕組みが必要。
- 大規模計算資源(GPU、専用ハード)、大量データ(学習コーパス)の重要性が増大する。
- 最新の共通AI技術のプラットフォーム化により、上位のアプリの実装に左右されずに最新技術を適用することが可能になる仕組みが必要。当該プラットフォームの共用によって、スケールメリットの享受も期待。また、上位の実装を標準化し、再利用性を高めてオープン化したり、柔軟に外部のシステムと接続可能な機能実装の仕組みに取り組むことが重要。



<今後の方向性>

- ◆ 複数分野の事業者が利用することのできる、上位のアプリの実装に左右されずに柔軟に外部のシステムと接続可能な機能実装の仕組みの開発を推進する。
- ◆ 大規模計算機や大規模データを扱うことのできる環境の整備を推進する。

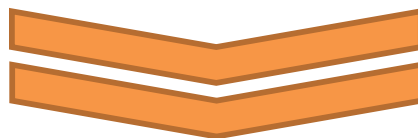
3. その他(セキュリティ)

<現状・課題>

- 利用価値が高いとされているパーソナルデータを利活用するためには、プライバシー保護が不可欠だが、多くの企業は対処方針に関する検討が不十分なため、挑戦的・スピーディーな活動がしにくい。
- 個人情報をも本人の同意を取らなくても利活用可能とするためには匿名加工情報に加工する必要がある。いかに再識別のリスクを低減し、データの有用性を保ちつつ匿名加工情報とするかが課題。分野によって匿名加工の方法は異なり、手法や基準が議論されている。
- 現状、Webから抽出した表現などを学習データとして第三者に研究用として提供する場合、著作権法上の課題がある。

<主な意見>

- 暗号化処理を用いたプライバシー保護によるデータ解析技術が確立すれば、データの共有において有効。
- 顧客の情報を解析し、その結果を個人にフィードバックすることによってさらに良いデータが取得できる可能性がある。しかし、そうすると個人とデータが紐づいてしまうため、新たな仕組みが重要。
- なかなかデータが出てこない問題を克服するためには、暗号化処理を用いた プライバシー保護によるデータ解析技術のような研究は非常に重要。特に医療系とか金融系、政府の情報をうまく扱える仕組みが出てくると期待。
- 医療分野のデータをいきなり広く公開することは難しいので、少しずつ広げていくべき。



<今後の方向性>

- ◆ 利用価値の高いパーソナルデータ等を、データの有用性を保ちつつ匿名加工情報に加工する技術の研究開発を推進する。
- ◆ 第三者提供の同意を得やすくするための暗号化処理等の研究開発を推進する。