

ICTを活用し コミュニケーションの壁を超える

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
研究所長

木俣 豊氏に聞く

情報通信技術が急速に進歩するなか、インターネットなどのサイバー空間と私たちが生活している現実社会を結びつけ、様々な社会的課題を解決するための技術が求められている。しかし、高速・大容量通信ネットワーク上の大量のデータから必要な情報を選別し、活用するには、人との親和性の高いコミュニケーション技術が必要だ。

災害時や異文化間などで立ちふさがるコミュニケーションの壁を超える新たな技術開発について、けいはんな学研都市の国立研究開発法人情報通信研究機構(NICT)ユニバーサルコミュニケーション研究所の木俣豊所長にうかがった。

技術は使えなければ意味がない

榎館 NICTのユニバーサルコミュニケーション研究所(UCRI)では「コミュニケーションの壁を超える研究」を掲げて、新たな情報通信・コミュニケーション技術を研究開発されているとうかがっています。まず、NICTとUCRIについて、組織の概要から教えていただけますか。

木俣 NICTは情報通信を研究する唯一の国立研究機関であり、情報通信にかかわる幅広い分野で研究開発を行っています。そのルーツは1891年に開設された旧通信省の電気試験所であり、120年以上の歴史があります。ところが、同じ国立の研究機関でも産業技術総合研究所や理化学研究所がよく知られているのに比べ、皆さんにあまり認知されていない。ちょっと悔しい思いをしています(笑)。NICTの研究分野は第4期中長期計画(平成28~32年度)でみると「観る」、「守る」、「拓く」、「繋ぐ」、「創る」に分けられ、「観る」がセンシング基盤分野、「守る」がサイバーセキュリティ分野、「拓く」がフロンティア研究分野、「繋ぐ」が統合ICT基盤分野、「創る」がデータ活用基盤



Interviewer
京都総合経済研究所
取締役調査部長
榎館 孝寿

分野となっています。このうち、UCRIの担当分野はデータ活用基盤分野です。具体的にはおっしゃるように第3期中長期計画から継続して「コミュニケーションの壁を超える研究」をターゲットとして、いくつかのプロジェクトに取り組んでいます。すでに一般に公開した技術や、実用化の一手手前まで近づきつつある技術もあります。

その一つが「言語の壁」を超える翻訳技術です。東京オリピックが開催される2020年を目標に、特に必要とされる10言語を対象に精度の高い音声翻訳技術を駆使したシステム「VoiceTia (ボイストラ)」を開発し、スマートフォン（以下スマホ）などの端末を通じて実用化するプロジェクトが進行中です。それから「情報の量と質の壁」を超えるということ、大規模Web情報分析システム「WISDOM X (ウィズダム・エックス)」の開発にも取り組んでいます。このシステムは約40億ページにおよぶWeb上の日本語テキストを利用し、単なるキーワード検索ではなく、文単位での意味解析をして、それぞれをつなぎ合わせて一種の仮説まで作成するシステムです。

国の研究というと、とかく実用に結び付かない基礎研究を、白衣を着た研究者がやっているというイメージを抱かれがちです。しかしNICITはこうした音声翻訳技術等にみられるように、技術は使えなければ意味がない、使える技術をきちんと創ろうという思いで有用な研究開発に取り組んでいます。そのうえで創った技術は積極的に企業に移管し、民生技術・製品として社会に還元していくことを目指しています。

幼いころから機械で遊び、 高校時代、コンピュータに衝撃

楢館 そもそも所長ご自身、幼いころから機械に興味をもたれていたとお聞きしますが、どんなきっかけ、体験から、情報通信分野の研究者を志されたのですか。

木俣 先日も神戸の実家に帰って、父がテレビのリモコンが動かなくなったと言うので、「かしてみい」と言って分解し始めた。「お前はいつも、それやなあ。子供の頃から修理すると言っては分解し、結局、最後は壊していたやろ」と言われました。家がプロパンガス店だったので、風呂釜、取り外したガスメーターや、配管パイプをカットする機械などがあって、それを玩具にして育ちました。自分でも分解すると壊れるのがわかっていながら、興味があつて中を調べたくなり、バラして組み立て直してみると部品が一つ余っている（笑）。小学校高学年になると今度はプラモデルにモーターを載せて電池をたくさん取りつけすぎてモーターを焼き切るなど、理科で勉強したことを実際に試していました。『科学と学習』など科学実験の雑誌を読み、付録の実験道具で遊ぶのが大好きでしたね。そして、マイコンが出てきたのがちょうど高校時代でした。テレビの『スタートレック』、『宇宙大作戦』が好きで、コンピュータという機械はどういう仕組みになっているのかと興味がありました。しかし、当時コンピュータは非常に高価で、興味はあつたが触れる機会がない。『ザ・ベーシック』、『マイコン』などという雑

誌を立ち読みしたり、図書館で半導体、LSIなどについて調べたりしました。そんなときに家庭教師の先生がコンピュータを買ったから、見に来ないかと声をかけてくれた。当時NECのPC8001が発売されたばかり。プログラムの記憶媒体がまだカセットテープの時代で、プログラムを入れて「F1」とキーを叩くとゲームの画面が立ち上がる。あのときの衝撃は忘れられません。プログラムを入れるとどうしてこうなるのか。とても不思議で雑誌を買って、プログラムを勉強しました。家はガス店なのでからは化学系の学部に行けと言われましたが、どうしてもコンピュータの勉強がしたいとロボット開発にも関係する神戸大学の計測工学科に進みました。コンピュータが好きでしたが、やはり機械も好きだったんですね。

大学4年になると、当時、第2次AIブームが起こり、ニューラルネットワーク（人間の脳の情報処理の働きをモデルにした人工知能のシステム）やフuzzy制御（人間の認識のあいまいな部分をコンピュータで処理する技術）などがブームになりました。私は人間の脳を模した計算アルゴリズムで、学習を重ねると自分で処理できるようにするニューラルネットワークを専攻したいと考え、制御技術の研究室に入りました。しかし結局は、研究室のテーマがロボットアームの力制御の研究で、卒論でAIを取り上げることができませんでした。就職先も制御技術を活かし圧延機の制御などに取り組みみたいと（株）神戸製鋼所に入社して研究所に配属され次世代品質管理データベースの研究などをするこ



木俣 豊 (きだわら ゆたか)

1965年神戸市生まれ。1990年神戸大学大学院自然科学研究科計測工学専攻修士、同年神戸製鋼所入社。1999年神戸大学大学院自然科学研究科情報メディア科学専攻博士後期課程修士、工学博士。2001年独立行政法人通信総合研究所(現、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研究部長。2011年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研究部長。2016年国立研究開発法人情報通信研究機構先進的音声翻訳研究開発推進センター開発推進センター長、統合ビッグデータ研究センター研究センター長。日本データベース学会理事。

とドイツのBOSSCH社との合併会社に出向し、アンチロックブレーキ(自動車のABS)の品質管理をするためのリアルタイムシミュレータや自動試験装置の開発にかかりました。そして、1995年、私が本格的に情報処理分野の基礎研究を進めるきっかけとなる阪神・淡路大震災に遭遇します。当時、旧郵政省に通信・放送機構(現在のNICT)があり、そこが震災復興プロジェクトを進めていました。私はそこに神戸製鋼所から研究員として派遣され、情報通信の基礎研究を学ぶことになりました。

TOEIC600点レベルの 音声翻訳アプリを無料で公開

榎館 まさにICTの進歩の歴史とともに歩んでこられた世代ですね。それにドイツのBOSSCH社との合併会社では「言葉の壁」も体験された(笑)。ところで、その「言葉の壁」を超える音声翻訳システム「VoiceTra」について、もう少し詳しく教えていただけますか。

木俣 音声翻訳技術はNICTが長年行ってきた

た音声の研究と翻訳の研究を合体したもので、そのアプリの「VoiceTra」はもともと旅行者との会話を想定したものです。すでに2010年7月にiPhone上で動作する世界初の多言語音声翻訳システムとして公開しました。その後、改良を

経て現在ではアンドロイド系のスマホでも無料でダウンロードでき、日英翻訳であれば、2ヶ月前のシステムでTOEIC600点レベル(470点で通常会話であれば支障ない。730点でどんな状況でも適切なコミュニケーションができる素地を備えているとされる)の翻訳機能があります。現在はさらに機能アップし、2020年までには最低で650点、できれば700点をクリアしたいと思っています。それから、欧米の有名な研究機関も参加する音声翻訳技術のための国際評価ワークショップ(IWSLT)でも、我々のチームが英語のTEDtalks(各界のエキスパートによる素晴らしいプレゼンテーションが無料で視聴できる動画配信サービス)の音声データを使用し翻訳するコンペで2012年、13年、14年と連続して世界一になっています。翻訳の対象となる言語は精度にバラつきはありますが、テキスト翻訳も含めると29言語になります。先ほど述べたように、その中から英・中・韓など10言語に絞り、2020年を目標に翻訳の精度を上げて実用化するオンラインプロジェクトを現在、進めています。チ

ームには民間企業からも14人の研究者が参加し、まさにオールジャパンの研究開発体制で臨んでいます。

観光旅行の会話に加え、外国人が日本で具合が悪くなったときに、病院でも使えるように、東大病院と協同で医療専門用語の翻訳なども強化しています。たとえば、具合の悪い外国人が病院の受付で「I have an irregular pulse」と音声入力すると「私は不整脈の持病があります」と日本語の音声に翻訳される。これは入力された音声スマホやタブレットを通じてサーバーに送られ、サーバーが認識して文字として表示。これを翻訳して英語のテキストに翻訳し、さらにそれを英語の音声として再生する。それを瞬時に行うわけです。音声が入ると認識されたかどうか、端末に表示された文字で確認できる検証機能もあります。日英翻訳だけではなく、たとえば日本語・ミャンマー語でも試してみると「このあたりに日本食のレストランはありませんか」と日本語音声を読み込むと、ミャンマー語の音声に翻訳される。だから、このアプリは逆に我々日本人が海外に行ったときに、大変心強い。英語が通用しない中国や東南アジアで想定外のことが起きたときなどにも、事態を把握するのに役に立ちます。私もアジア等の海外出張の際には積極的に使うようにしています。たとえばベトナムの街中にある雑貨店でお土産を買うときにも、「もう少し安くありませんか」、「30%引きでどうですか」などとコミュニケーションがとれました。

榎館 2020年には多くの国からたくさんの方が日本を訪れることになりましたが、このアプリ

リがあれば、おもてなしの役に立ちそうですね。さきほどプロジェクトチームに企業の研究者も参加されているということでしたが、音声翻訳技術を使った製品開発などは行われていないのですか。

木俣 一緒に研究している企業の方たちもこの翻訳エンジンを活用しているいろいろな商用システムを創って下さっています。おもしろいところでは、パナソニック(株)様が開発しているメガホン型の翻訳システムがあります。今年の東京マラソンで実証実験を行いました。マラソンを観戦している人たちに警察官が「白線の後ろに下がってください」と日本語を入力した後にボタンを押して4か国語に翻訳した音声で注意を促すという実験をしました。

我々の翻訳システムを使った、こうしたグローバルコミュニケーション・プロジェクトの実証実験が日本各地で実施されています。たとえば鳥取県ではKDDI(株)様がタクシーで翻訳システムを使う実証実験を行い、好評だと聞いています。また、京都府、京都市ではヤマハ(株)様が電車アナウンス情報をSound U.D.化する技術と組み合わせた多言語に翻訳するプロジェクトを実施しています。我々の翻訳エンジンはサーバーの翻訳する部分に使われています。

音声翻訳技術を活用して、聴障者と健聴者のコミュニケーションを図る

植館 実用化が楽しみですですね。ところで、開発にあたってのエピソードがあればお話しただけですか。

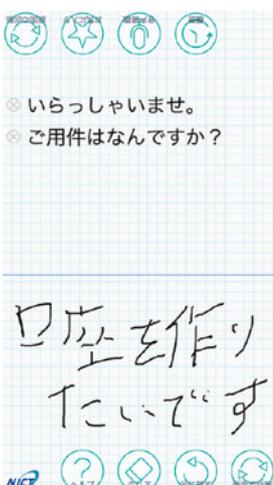
木俣 音声認識のシステムに関してぜひお伝えしたいエピソードがあります。2012年に県立熊本聾学校の山田先生という方からメールをいただいた。内容は「いままでいろいろな音声認識システムを試してみたが、使いにくいものが多い。ところが「VoiceTra」をダウンロードして試してみたところ、こんなに精度の高い音声認識は初めてでびっくりした。ついてはこの音声翻訳技術を活用して、聴覚障がい者(以下聴障者)と健聴者のコミュニケーションを支援するためのシステム、健聴者が話した音声を文字に変換して聴障者が読めるようなシステムを創れないか」というものでした。話をいただいて、これは国の研究機関として当然取り組むべき事業だということになり、学校と1年半かけて協同開発したのがスマホアプリの「こえとら」です。スマホで持ち歩ける非常にシンプルなシステムで、聴障者は文字で「私は耳が聴こえないので、これを使ってお話しします。ちょっと、よろしいでしょうか」と入力します。健聴者はその答えをスマホに向かって話しかける。音声で入力すると画面に「はい、いいですよ」と表示される。聴障者は文字や定形文で質問を入力します。「交番はどこですか」等の場所を聞く質問では健聴者は音声入力や地図で場所を示して答える。という具合にスマホを介して会話が成立する。文字入力も音声入力もできます。このアプリは聴障者の方たちに大変、喜んでいただいているようです。ただし、我々は研究機関であって、いつまで運用できるか約束できない。そこで総務省とともにある事業者へ技術を移管し、費用はキャリア各社に負担してもらうとい

う仕組みをつくり、サービスを継続しています。今年の4月に熊本地震が発生しましたが、熊本聾学校はいちばん被害が大きかった地域にあり、聴障者の方たちも避難所に避難され、「こえとら」を搭載したタブレットが足りないかと連絡をいただきました。そこで、研究所でお貸しできるタブレット端末を持って行って、使っていたいただきました。それから、熊本地震では、被災された外国人の方が「VoiceTra」を利用されたというログデータが上がってきています。我々の開発したシステムが、災害時、皆さんの助けとなることで開発にも力が入ります。

また「こえとら」の別バージョンとして、聴障者のお客さんが商店等に来店されたときに店員側が使用するための「SpeechCanvas」(スピーチキャンバス)も開発しました。「何か御用



【VoiceTraの活用例】
スマホに向かって話しかけるか、文字入力を行うと選択した外国語に翻訳され音声流れます



【SpeechCanvasの活用例】
話した言葉が次々と画面上で文字になり、画面を指でなぞれば絵や字がかけます



京都銀行でのタブレットを活用した接客イメージ

件がありますか」とタブレットに話しかけると声がそのまま文字になり画面に表示されます。聴障者はタブレット画面を指でなぞれば絵や文字が書けます。そうやってコミュニケーションを図ります。

植館 実は京都銀行も7月下旬からこの「VoiceTra」と「SpeechCanvas」と「VoiceTra」を搭載したタブレット端末を金融機関で初めて導入させていただくことになり、17か店の窓口を設置しました。

「VoiceTra」は音声翻訳を目的とした技術でしたが、山田先生のアイデアによって、新しい用途が開発されたということで、大変素晴らしいお話ですね。ところで、「VoiceTra」は旅行客の会話をターゲットに開発されたということですが、グローバル化の中でビジネスにも言葉の壁が立ちふさがっています。ビジネス会話には対応できないのでしょうか。

木俣 いまの「VoiceTra」はビジネス会話のデータが入っていません。残念ながら人間同様、システムも知らない言葉は理解できません。しかし、基盤技術は確立されているので、ビジネスの現場で使えるデータを集めて、学習させればビジネス会話も可能になります。現にライセンスを供与しているある会社ではこの技術をベ

ー스에 中国の工場で使うための商用システムを開発しています。精度を上げるためにはデータが必要なのです。「対訳コーパス」といいますが日本語と他の言語の対訳をたくさん集めて、それを機械に学習させ確率を計算させる。この言葉が来たときには、こういう並びの確率が高いという具合に。たとえば日本語の「どこですか？」ならば、「Where is...」の確率が60%、「Could you direct me to...」が20%、「Where am I...」が20%といった具合です。よく考えてみると我々は中高6年間、英語を習っても英語が話せない。ところがアメリカで生活すればしばらくすると話せるようになる。それは、こんなときにはこう話せばいいと脳が学習しているからです。だから、データをたくさん集めることが重要になります。集めた音声データをコンピュータに入れたり、それをテキストに変換し、データをコンピュータに読ませたりするのに、人手が要るので膨大なコストがかかります。一方、音声認識に関しては最新のAI技術である深層学習（ディープラーニング）などが取り入れられるようになり、開発手法も進化してきています。

速報性の高いTwitter情報を災害救助活動に活かす

植館 なるほど、そうするとこれから、「VoiceTra」の応用範囲は大きく広がっていくということですね。大変楽しみな技術ですね。最後に冒頭でお話しいただいた「WISDOM X」についてもう少し詳しく教えてくださいませんか。

木俣 我々が通常、インターネット上の情報を探す場合、インターネットの検索エンジンにキーワードを入力します。そして、検索結果のせいで最初の1〜2ページ程度に目を通すことが多い。ところが、実際は広告ページも多く、情報の質が問われないまま大量のデータが埋め込まれていて、大事な情報を見落とされている可能性もあります。たとえば、ある健康食品について、検索すると芸能人が愛好しているなど、優良な食品であるかのような情報が並んでいる。ところが、目を通さなかった10ページ目にその食品による副作用や事故の情報が載っているかもしれません。こうした「情報の量と質の壁を超える」ためには情報を自動的に判別する技術が必要になる。この技術開発については第2期中期計画から始めて10年たつてようやく実用化の目処が立ってきました。たとえば、「地球温暖化」と打ち込んで、「何をもちたらずか」と質問すると「異常気象」とか「海水温の上昇」が上がってくる。そこでそのまま「海水温の上昇」の因果関係をたどると、システムが関連テキストを引用し「腸炎ビブリオ菌が増える」という情報が出てくる。ここでまた、「腸炎ビブリオ菌が増える」とどうなるか」と聞くと「食中毒が増える」と。そうすると「地球温暖化が進むとどうなるか」→「海水温が上昇する」→「腸炎ビブリオ菌が増える」→「食中毒が増える」という仮説が出てくるわけです。どこでどうつながるのか？そこで本当にそうなるのかと調べてみると、世界で特に権威のある学術雑誌のひとつと評価されている「Nature」に、バルト海沿岸で地球温暖化と思われる人為的な原因のため

に海水温が上昇し食中毒が増えているというレポートがあったのです。このように「WISDOM」の核となる技術は自然言語処理といってコンピュータがWeb上に書かれた意味を解析し、そのつながりを計算し、つながりをたどることができるのです。

さらに東日本大震災を機に、この「WISDOM」の技術を活用し、並行して開発を進めてきたのが対災害「Twitter 情報分析システム [DISAANA (ディサーナ)]」です。

東日本大震災ではSNSの情報が救援活動に重要な役割を果たしましたが、情報の質と量が災害時でもコミュニケーションの大きな壁となったからです。「DISAANA」はツイートのデータをリアルタイムで分析し、災害状況を把握するシステムで、2015年4月から誰もが利用できるように公開しています。そのため今回の熊本地震では多くの自治体やNPOの方たちに使っていただきました。たとえば内閣官房でも「DISAANA」を使って被災状況を把握しながら支援計画の作成をしたという報道がありました。

システムには質問回答モードと災害関連抽出モードの二つのモードがあり、質問回答モードの方は「熊本県ではどこで火災がおきていますか」、「熊本県では何が不足していますか」というように質問



熊本地震時の「DISAANA」の画面。(右)質問回答モード、(左)災害関連抽出モード

を入力します。災害関連抽出モードの方は状況がわかっていない場合、とりあえずエリアで何か困ったことが起きていないかどうか、都道府県、市区町村名を入力します。また、アカウントの範囲を絞ったり、ツイートの表示条件を絞り込んだりするなど検索結果を絞り込むことも可能です。たとえば関東東北豪雨による鬼怒川決壊の折には「どこに避難していますか」と質問すると、通常のキーワード検索では1万2800件も出てくるのが、我々のシステムでやると27件に絞れました。中身を見て同じようなものチェックし絞り込むわけです。「孤立」というキーワード検索でも1900件に対し12件です。それから熊本地震では「ライオンが動物園から逃げた」というツイートと「デマです」というツイートがあった。その場合もきちんと意味的に矛盾がないか抽出して、デマだと判断できました。意味的につながるかどうかシステムが判断し、たとえばこの文章とこの文章はつ

ながるとシステムが探し出してくるのです。ツイートのいちばんの特長は速報性です。どこかで竜巻が発生したときにもNHKの速報よりも早く「Twitter」に情報が発信されたという事実があります。東日本大震災でも30分後には多くのツイートで現地状況についての情報が発信されました。「DISAANA」開発プロジェクトはこのツイートの速報性を活かさないかと、Twitter Data Grants (Twitter社がさまざまな研究機関に対し、自社の持つパブリックデータや履歴データに研究目的の為に無料でアクセスできるようにするデータ提供プロジェクト)に応募し採択されました。Twitter社から東日本大震災時1か月間の「Twitterデータ(100%)」を無償提供され、研究開発を進めました。DISAANAは通常は日本語「Twitter」の10%を購入してリアルタイム分析を行っています。熊本地震が起き、代理店を介して「Twitter」と交渉し、今年の5月、1か月間に限り残り90%のデータを無償供与していただきました。我々のシステムはその膨大なツイート情報にも十分に対応することができました。現在、自治体、官公庁、消防庁など防災関連機関から多くの問い合わせをいただき、今後、どのようにこのシステムを活用できるか議論しているところです。こうしたビジネスに結びつきにくい、民間が手を出しにくい分野の研究開発こそ、我々のミッションではないかと考えています。

植館 本日は東京オリピック、さらにその先の未来に向けた素晴らしいユニバーサルコミュニケーション技術をご紹介いただき、ありがとうございました。