

ユニバーサルコミュニケーション 技術の課題と展望

— コミュニケーションの壁がない社会を目指して —

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



1. はじめに

インターネットが生活の隅々にまで情報を流通させ、老若男女がスマートフォンやタブレットを活用して、いつでもどこでも誰とでも情報を交換できる時代になって久しい。この大量の情報と大容量高速なネットワークによって、我々の社会活動は一変したと言える。もはや、インターネットのない世界は想定できない時代である。我々は高速なネットワーク技術と高性能なICT機器の恩恵を受けて遠隔地の家族や友人と音声や映像を用いたコミュニケーションを行える。しかしながら、ネットワークの高帯域化とICT機器の高性能化だけでは越えられないコミュニケーションの壁が存在しており、我々は依然として自分以外の人々や社会を十分に理解することが困難である。このように、いまだに残るコミュニケーションを阻害するいくつかの壁は、新たな技術開発によって乗り越えられる可能性がある。本稿では、IoT (Internet of Things) 技術によってさらに多様化する情報社会において、コミュニケーションの壁を越えるためのユニバーサルコミュニケーション技術の現状と将来について述べる。

2. ユニバーサルコミュニケーション技術とは

2005年7月29日に総務省情報通信審議会によって、UNS (Universal Communications, New Generation Networks, Security and Safety for the Ubiquitous Network Society) 戦略プログラムの答申が行われ、ユビキタスネットワーク社会における重要な研究開発の1つとしてユニバーサルコミュニケーション技術が示された[1]。このプログラムにおいては、ユニバーサルコミュニケーション技術が「世界最先端のユビキタスネットを活用して知的創造活動を促進するコミュニケーションの技術、年齢・身体・言語・文化などの壁を乗り越え高齢者や障碍(しょうがい)者をはじめ人に優しいコミュニケーションの技術を実現」

と定義されている。この答申が出されてから10年間、我が国において「言語の壁を越える技術」「情報の量と質の壁を越える技術」「臨場感と距離の壁を越える技術」の3つの技術がユニバーサルコミュニケーション技術のコア技術として研究開発されてきた。本章では、国立研究開発法人情報通信研究機構(以下、NICT)における研究開発を中心に現在までに研究・開発された技術について述べる。

2.1 言語の壁を越える多言語翻訳・音声処理技術

我々にとって、日本語は母国語であるが全世界的に見るとマイナーな言語[2]であり、外国人とコミュニケーションする場合には言語の壁を痛感することになる。SFの世界においても人種や民族を越えたコミュニケーションには音声で入力すると異言語に翻訳される多言語音声翻訳装置が登場する。言葉の壁を越える技術は、まさに人類にとっての夢の技術とも言える。

多言語音声翻訳技術は1980年代後半に(株)国際電気通信基礎技術研究所(ATR)にて研究開発が始まり、我が国はその研究開発の最先端を走ってきた。NICTでは、2006年からATRの研究成果を引き継いで研究開発を進めてきた。旅行会話にフォーカスして構築した大規模な音声・対訳コーパスを用いた統計翻訳技術を開発することで翻訳の高精度化を実現して、実用レベルの音声翻訳を可能とした。その成果を社会で実証するために、iPhone(注1)上で動作する世界初の多言語音声翻訳システム「VoiceTra(注2)」を開発し、2010年7月に公開した。その後の改良を経て、日英翻訳であれば、TOEIC(注3)600点レベルの翻訳機能を有するシステムとなり、ユーザーからは「実用的である」という高い評価を得て、運用停止

(注1) Apple Inc.の登録商標または商標

(注2) 国立研究開発法人情報通信研究機構の日本国内における登録商標

(注3) Educational Testing Serviceの登録商標または商標

までの約2年半の間に84万ダウンロードされ、平均約1万発話/日のシステムへと成長した。VoiceTraはその後、25箇国31研究機関による多言語音声翻訳コンソーシアムU-STAR[3]の実証システム「VoiceTra4U（第1図）」へと発展している。さらに医療分野においても東京大学医学部付属病院における臨床実験によって、医療現場でも利用可能なレベルに達しつつある。また、大量に出願される他国の特許文書を日本語に翻訳するニーズも極めて高い。特に大きな市場である中国で、出願される特許が急激に増加しているが、中国語で記述されているため、それらに対応することが十分にできておらず、特許文書の中日翻訳技術の高精度化が急務となっている。特許文書は長文で、独特の言い回しが存在するが、対訳コーパスの整備によって、実用レベルの翻訳精度が得られる翻訳技術の開発が可能となる。NICTでは特許庁との共同研究によって、高精度な特許翻訳エンジンを開発[4]し、中国語特許の要約文の意味がわかる理解度を従来技術の約15%から約80%に引き上げることに成功した。



言語	入力		出力	
	音声	テキスト	音声	テキスト
日本語、中国語、韓国語、トルコ語、ヒンディ語、タイ語、インドネシア語、ベトナム語、マレー語	✓	✓	✓	✓
モンゴル語		✓	✓	✓
ウルドゥー語、シンハラ語、ゾンカ語、ネパール語、フィリピン語		✓		✓
アラビア語		✓		✓
英語、オランダ語、ドイツ語、ハンガリー語、ポーランド語、ポルトガル語、ロシア語	✓	✓	✓	✓
フランス語	✓	✓		✓
イタリア語、スペイン語、デンマーク語		✓		✓

第1図 VoiceTra4Uの概要

多言語音声認識技術においては、コンペ型の国際会議であるIWSLTにおいてTED（Technology Entertainment Design）英語講演音声に対する音声認識分野で2012年か

ら2014年まで3年連続トップスコアの音声認識精度を示しており[5]、あらかじめ音声認識することを想定していない講演の収録音声でも、高精度に認識できることを証明している。

2.2 情報の量と質の壁を越える情報分析・利活用技術

インターネットはWeb技術とともに普及し、その結果、インターネットを介して膨大な情報に容易にアクセスできるようになった。しかし、次々にWebサイト上に発信される情報によって、そのWebサイトの内容は逐次変化する。そのようなWeb情報の性質やその情報量の膨大さゆえに、もはやカテゴリ分類や情報品質などで情報を管理することは不可能となっており、現在においては検索エンジンが情報アクセスの入り口となっている。その検索エンジンは記述されている文書の意味内容までは解析しておらず、基本的には関連する単語の有無などで検索が行われている。検索結果の順位についても情報の品質によって並べられている訳ではない。しかしながら大多数の検索エンジン利用者は検索結果の上位数件を確認するに過ぎず、誤った情報などで間違ったコミュニケーションをすることも少なくない。このように「情報の量と質の壁」は情報社会において、ミスコミュニケーションの原因となっている。NICTでは2006年から情報の品質を対象とした研究開発をスタートさせて、2010年には「WISDOM」と呼ばれるまったく新しいコンセプトの情報分析エンジンを開発した。この情報分析エンジンは、収集したWeb情報6億ページからSPAM情報などを削除した約1億ページを対象として発信者情報や意見の肯定・否定判断などを行い、入力した問い合わせに関連するWebページの発信者分類や、発信者ごとの肯定・否定意見、主要なトピックスなどを抽出できるものであった。また、意味解析に基づく推論機能の研究開発成果を用いた音声質問応答システム「一休」の開発を行い、大量の情報から質の良い情報を抽出して活用させることが、コミュニケーションの壁を打ち破ることを実証した。2011年からは、両者の技術を統合し、対象とするWeb情報を40億ページまで拡大し、さらにリアルタイムで処理する技術の開発を目標として研究開発を推進している。2015年3月には、「なに」「なぜ」「どうなる」など多様な質問に回答し、また、キーワードから関連する回答可能な質問の提案までを行う新たな情報分析エンジン「WISDOM X（第2図）」を開発し、インターネット上で誰もが利用できるように公開（<http://wisdom-nict.jp/>）[6]している。

また、2011年に発生した東日本大震災においては、SNSの情報が災害発生後の救助活動に重要な役割を果たしたが、情報の量と質が災害時においてもコミュニケーション



第2図 WISDOM Xの利用例

の大きな壁になっていることが明らかとなった。そこで、WISDOM Xの研究開発と並行して災害時のコミュニケーションの壁を打ち破るための災害情報分析システム「DISAANA」の開発に着手した。この開発プロジェクトはTwitter Data Grants^(注4)にも採択され、Twitter, Inc.からTweetデータの無償提供[7]を受けることで研究開発が効率的に進み、Twitter情報をリアルタイムに分析できるシステムとして2015年4月から誰もが利用できるように公開 (<http://disaana.jp/>) している[8]。

このように自然言語処理技術を中心とした意味解析技術の高度化によって、必要とする人に適切な情報を提供することが可能となり、Web上のテキスト情報を中心とした「情報の量と質の壁」を克服しつつある。

2.3 臨場感と距離の壁を越える超臨場感通信技術

現在では、我々は無料のソフトウェアを活用して従来の電話のように自由にテレビ電話を利用しているが、UNS戦略プログラムが発表された2005年当時にはテレビ電話は専用機器を用いる特別なものであった。音声だけの通話から映像付きの通話になることによって、遠隔地の状況がわかりやすくなりコミュニケーションの壁が低くなったと言える。しかしながら、限られた範囲の平面的な映像は実空間の広がりや、その場の臭い、物体の感触までは伝わらない。このような課題に関して裸眼立体映像表示技術や、その映像を伝送するための符号圧縮通信技術などの超臨場感通信技術の研究開発を推進している。裸眼立体表示技術は、再生光を完全に再現できるホログラフィ表示技術が究極の技術と言われている。NICTでは2011年にフルカラーの動画像が表示できる対角5 cmの電子ホログラフィ表示技術を開発しているものの、大画面での実用レベルの技術には、まだまだ多くの研究課

題がある。大画面での裸眼立体表示技術としては、200台のプロジェクタレイによる200視差を有する大画面裸眼立体表示装置を2012年に完成させて、グランフロント大阪に設置して実証実験を開始している。さらに、感触や臭いを伝達する技術を開発して、映像や音声以外の触覚や嗅覚を伝える技術の研究開発を進めている。このような超臨場感通信技術によって「臨場感と距離の壁」を克服を目指している。

2.4 ユニバーサルコミュニケーション技術の社会実装

NICTにおいては、3つのコミュニケーションの壁を対象として、その壁を越える研究開発を進めてきた。これらの成果は世界的にもトップクラスのものであり、我が国においてはNICTが中心的な役割を果たしてきたと言える。また、開発した技術はNICTから社会へと技術移転が行われている。特に、音声認識技術や多言語翻訳技術は民間企業の商用サービスのコア技術として活用されている。情報分析技術においても、DISAANAが自治体から関心をもたれており、自治体の災害訓練などにおいて試用され、一定の効果をj得ている。さらにWISDOM Xの技術は企業などからも注目を集めている。大画面裸眼立体表示装置も企業によって商用化されている。一方、NICT以外においても多言語音声翻訳技術の研究や自然言語処理に基づく情報分析技術の研究開発は加速しつつある。多言語音声翻訳はGoogle Inc.が2014年にGoogle翻訳^(注5)をリリースした。音声だけでなく、カメラで撮影した文字も多言語に翻訳する機能を有している。また、同年にMicrosoft Corp.が音声通話に多言語翻訳機能を組み込んだSkype^(注6)のβ版をリリースしている。

我が国では、これらの米国大手企業に先駆けて、多言語音声翻訳研究開発を進めてきた。しかし、大規模なコーパス構築などが必要となっている現在の多言語音声翻訳技術は、1つの組織だけで研究開発することが困難な分野となっている。従って、今後は産官学一体となってユニバーサルコミュニケーション技術の社会実装を進めていくことが、将来の我が国の情報産業において極めて重要である。このような背景の下で、2014年4月11日に新藤総務大臣(当時)が、世界の言葉の壁を越えることをミッションとした、「グローバルコミュニケーション計画[9]」を発表した。この計画に基づき、NICTは先進的音声翻訳研究開発推進センターを設立し、産学官の研究者が結集したAll Japanの体制でNICTのVoiceTra技術をベースに次世代の多言語音声翻訳技術の研究開発を進めている。

(注4) TwitterはTwitter, Inc.の登録商標または商標

(注5) GoogleはGoogle Inc.の登録商標または商標

(注6) Microsoft Corp.の登録商標または商標

さらに、旅行会話を対象として実用的に使える多言語音声翻訳技術を2020年までに医療現場やショッピングの現場、鉄道やタクシーなどの交通機関の現場などで利用可能にするために、グローバルコミュニケーション開発推進協議会[10]において、活発な議論が行われ、研究開発と社会実装の計画が作られている。

3. 次世代のユニバーサルコミュニケーション技術

これまで、3つのコミュニケーションの壁を克服する技術の研究開発を行ってきたが、克服しなければならない新たな壁として「人と環境のコミュニケーションの壁」が大きな課題となっている。我々は自分の周りの環境をほんの少ししか理解できておらず、少し離れた環境の情報ですら、ほとんど理解できていないと言っても過言ではない。その結果、よく知っている場所で活動する場合でも、その日によって状況が異なり、想像以上の混雑や天候によって求めていたサービスを受けられず、疲弊することも少なくない。この1つの原因は、インターネットからWeb情報を探すような手軽さで実世界の状況を理解できる情報を見つけられないことである。しかしながら、その環境もIoT技術によって大きく変化しつつある。これは「モノのインターネット」と呼ばれ、ユビキタスコンピューティング技術の次のフェーズとして、コンピュータや携帯端末だけでなく、さらに粒度の細かなセンサや、物体そのものをネットワークに接続可能にする技術である。この技術によって、インターネット上から実空間の情報にアクセスできる環境が整いつつある。

これまで、インターネット上で流通されていた情報は、テキスト、音声、映像など、人の手によって編集された人の意図を含んだ「デジタルコンテンツ」が主流であった。IoT技術の普及とともに、人の手を介さずに「モノ」が「センサデータ」として直接インターネットに情報を流すようになると、そのデータの種別や数は「デジタルコンテンツ」の比でなくなり、さらに情報が多様かつ膨大となる。IoTによって生み出される情報はWeb情報とは異なり、ある特定の環境や社会活動の状態を人の主観を含まずに表現できる性質をもっているため、「異種情報の量と質の壁を越える技術」の研究開発が新たな課題となる。また、このように大規模・多様化する情報の特異点を見つけるためには、従来の可視化技術では限界があり、人の五感を活用する「情報の可感覚化技術」が重要となる。

これらの2つの技術は、新たなコミュニケーションの壁である「人と環境のコミュニケーションの壁」の克服につながる。次節にそれぞれの研究課題を述べる。

3.1 異種情報の量と質の壁を越える横断的情報分析技術

実空間でセンシングされた環境情報や社会・経済活動の情報は、一般的には時系列の数値データであり、専門家以外にはその情報の価値が見いだせない。従って、一般の人々にその情報価値を提供するためには、他の情報との相関・因果関係を明らかにする必要がある。例えば、あるイベント会場の入場者の遷移情報は、単なる数字の変化にすぎないが、その会場の収容者数やWeb上での評判情報などを組み合わせて分析することで入場者数の変化の理由などがわかるようになる。つまり、実世界の情報と人の主観情報などの異種・異分野の情報を横断的に分析することで、「人と環境のコミュニケーションの壁」が克服できるようになる。

これまではコミュニケーションの壁を越えるために、知識としての「情報のつながり」に主眼を置いてきたと言える。いわば、知識空間上で言語間や文単位で情報をつなげることによって、コミュニケーションの壁を越える技術を研究・開発してきた。次世代のユニバーサルコミュニケーションにおいては、「知識空間上の情報と実空間情報の融合」が重要となる。知識空間と実空間情報を融合する鍵となるのが、その知識や情報をもつ時空間情報である。知識の情報源となり得るSNS情報は、スマートフォンなどで発信されることが多くなっており、その発信情報には発信場所の時空間情報が付与されている。また、社会で発生した事件、事故、イベントに関するWebやSNS情報の記述には、文章として空間と時間情報が付与されている。その情報を分析することで「いつ」、「どこで」、「何を」したのかという情報が抽出できる。これらの情報とIoT情報で取得された環境情報などを横断的に検索して、融合的に分析することによって「異種情報の量と質の壁を越える技術」を実現できる。NICTでは、既にその研究開発を進めており、その成果の一部をDISAANAやビッグデータ解析ツール「Event Data Warehouse」に実装している。

3.2 異種情報の量と質の壁を越える情報可感覚化技術

近年、計算機の進化によってビッグデータ処理技術としての機械学習技術が注目されている。DNN (Deep Neural Network)は大規模データで学習することによって分類精度を飛躍的に向上させる効果を示しており、その期待も大きい。しかし、最新の機械学習技術を用いても100%の精度で分類することは不可能であるため、条件を変更して何度も繰り返して分析することが必須である。実空間にいる分析者が大規模情報を操作しながら最適な知識空間を作り上げていくために障害となる「人と環境のコミュニケーションの壁」を克服するためには、人の

知覚能力を強化させて大規模情報の特異点を発見するためのインターフェースをもつ情報分析技術の研究開発が必要不可欠である。これまでは、情報の関係や性質を理解するために可視化技術を活用していたが、IoT時代に分析対象となる情報量は、視覚による情報識別能力を超える可能性がある。そこで、目に見えない情報を可視化するだけでなく、音声や触覚、嗅覚など、人間の感覚を活用する「情報の可感覚化技術」が大きな研究課題となることが想定できる。

4. ユニバーサルコミュニケーション社会の実現に向けて

前章までにユニバーサルコミュニケーション技術の「これまで」と、「これから」について述べた。本章では、5年後と15年後のユニバーサルコミュニケーション社会について予想する。

4.1 2020年のユニバーサルコミュニケーション社会

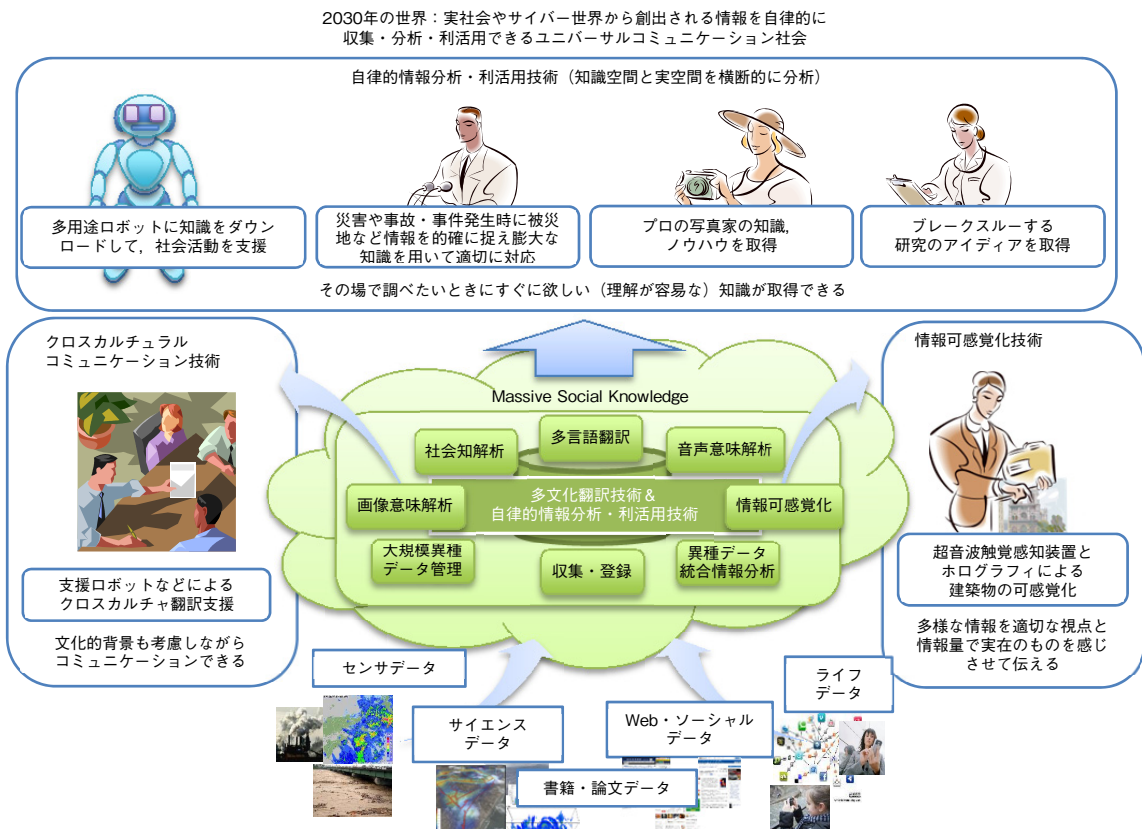
現在、NICTが中心となってパナソニック（株）を含めたAll Japan体制で研究開発を進めているグローバルコミュニケーション計画で、音声翻訳技術の社会実装が一気に進み、スマートフォンやペンダント型音声入出力装置、

HMD（Head Mounted Display）などを活用した多言語音声翻訳装置を利用して、我々は、言葉の壁のない社会の恩恵を得ている[11]。外国人を迎え入れる我が国の販売店やホテル、交通機関などでは、それぞれの業態に合わせた多言語音声翻訳装置で外国人旅行者への「おもてなし」を行っている。京都のみならず日本全国の都市が「世界で訪問したい都市」に挙げられ、我が国は、多数の外国人が来日する観光立国となっている。

社会・経済活動においては、従来は検索による情報検索が主流であったが、「異種情報の量と質の壁を越える横断的情報分析技術」によって、分析者が大量の情報を複数の視点で分類・整理を容易に行えるようになっている。

4.2 2030年のユニバーサルコミュニケーション社会

多言語翻訳技術は、ごく普通のあたりまえの技術となっており、文化的な背景も考慮した翻訳が可能となっている。音声認識技術はさらに進化しており、遠隔地とのテレビ会議では音声認識による自動議事録作成と同時翻訳が実現している。屋外で利用する情報端末も音声入力がスタンダードになっており、ウェアラブル端末やナビゲーションロボットによって社会活動が支援されている。プライバシー保護に配慮されながら、社会の隅々まで張



第3図 2030年のユニバーサルコミュニケーション社会

り巡らされたセンサネットワークによって、実世界の情報が取得できるようになっている。人々が発信するデジタルコンテンツと実世界のセンシング情報は、インターネットを介して大規模なデータとして集積された後に、社会における人類共通の知識として活用され、さまざまなコミュニケーションの壁を克服する情報基盤として活用されている（第3図）。

5. おわりに

コミュニケーションの壁を越える技術の研究開発を目標に掲げたUNS戦略プログラムが発表された2005年から10年が経過している。その間に研究室で生まれたさまざまな技術は実証実験によって得られたデータを用いて研究開発が行われることで、さらに実用的な技術へと進化しており、さまざまな技術が社会に実装されつつある。

ユニバーサルコミュニケーション技術は人類が言葉や文化を越えてわかり合うことに貢献できる極めて重要な技術である。昨今の不安定な世界において、人類がよりよい世界を作っていくためには、ユニバーサルコミュニケーション技術が必要不可欠であることは疑う余地もない。研究開発された技術は社会で使われなくては価値がない。例え研究室で極めて精度の高い翻訳や分析技術が作られて論文として高い学術的な評価を得たとしても社会の発展のためには十分ではなく、必要としている人々がその技術の恩恵を得られてはじめてその技術の価値が見いだせる。

我が国において、ユニバーサルコミュニケーション社会を実現させるには、産学官連携によって2020年に言葉の壁を越えることを目的としたグローバルコミュニケーション計画を成功させることが必須である。そこで得られた成功へのプロセスや実績が、次世代の技術や産業、ビジネスモデルを生み出すことになる。本稿によって、ユニバーサルコミュニケーション技術の現状と課題、そして今後の展望への理解が進み、新たな社会実装やビジネスへの発展に役立てば幸いである。

参考文献

- [1] 総務省, “報道資料 ユビキタスネット社会に向けた研究開発の在り方について UNS戦略プログラム,” 国立国会図書館インターネット資料収集保存事業, http://warp.ndl.go.jp/info:ndljp/pid/283520/www.soumu.go.jp/s-news/2005/05/0729_7.html, 参照 Oct.23, 2015.
- [2] “世界の母語人口（上位20言語）,” 文部科学省, http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo3/015/siryo/06032708/003/001.htm, 参照 Oct.23, 2015.
- [3] “U-STAR (Universal Speech Translation Advanced Research),” The Universal Speech Translation Advanced Research Consortium, <http://www.ustar-consortium.com/index.html>, 参照 Oct.23, 2015.
- [4] “プレスリリース NICTと特許庁が多言語特許文献の高精度自動翻訳の実現に向けて協力合意,” 国立研究開発法人情報通信研究機構, <http://www.nict.go.jp/press/2014/07/28-1.html>, 参照 Oct. 23, 2015.
- [5] “IWSLT2014,” IWSLT2014, <http://workshop2014.iwslt.org>, 参照 Oct. 23, 2015.
- [6] “大規模Web情報分析システム「WISDOM X（ウィズダム エックス）」をWeb上に試験公開(プレスリリース),” 国立研究開発法人情報通信研究機構, <http://www.nict.go.jp/press/2015/03/31-1.html>, 参照 Oct. 23, 2015.
- [7] “Twitter #DataGrants selections,” Twitter, Inc., <https://blog.twitter.com/2014/twitter-datagrants-selections>, 参照 Oct. 23, 2015.
- [8] “刻々と変わるSNSの情報をリアルタイムに平易な質問で分析・検索(プレスリリース),” 国立研究開発法人情報通信研究機構, <http://www.nict.go.jp/press/2015/04/08-1.html>, 参照 Oct. 23, 2015.
- [9] “グローバルコミュニケーション計画,” 総務省, http://www.soumu.go.jp/main_content/000285578.pdf, 参照 Oct. 23, 2015.
- [10] “グローバルコミュニケーション開発推進協議会,” グローバルコミュニケーション開発推進協議会, <http://gcp.nict.go.jp>, 参照 Oct. 23, 2015.
- [11] “Panasonic TVCM「ネイマール この世界をワンダーに篇」60秒,” YouTube, <https://www.youtube.com/watch?v=wyWWQzMRRH5>, 参照 Oct. 23, 2015.

《プロフィール》

木俣 豊（きだわら ゆたか）

1988	神戸大学工学部 計測工学科
1990	神戸大学自然科学研究科計測工学専攻修了（修士）
1990-1996	（株）神戸製鋼所 電子技術研究所
1996-1999	通信・放送機構 神戸リサーチセンター出向
1997-1999	神戸大学自然科学研究科 情報メディア科学専攻 博士後期課程 博士（工学）取得
1999-2001	（株）神戸製鋼所 電子情報部（2001年退職）
2001-2004	（独）通信総合研究所 次世代インターネットグループ 主任研究員
2006-2007	内閣府 政策統括官（科学技術政策担当）付参事官付 （総合科学技術会議事務局）情報通信担当出向
2007-2011	（独）情報通信研究機構 知識創成コミュニケーション研究センター 知識処理グループリーダー
2011-	（独）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長
2015-	独立行政法人通則法により国立研究開発法人情報通信研究機構に変更

専門技術分野：

データベース, ユビキタスコンピューティング