

7 MASTAR プロジェクト

7 MASTAR Project

7-1 MASTAR プロジェクトにおける音声翻訳技術

7-1 Speech Translation Technology in MASTAR Project

隅田英一郎

SUMITA Eiichiro

要旨

本稿では、平成20年4月より情報通信研究機構において開始した音声・言語に関する研究プロジェクトであるMASTAR (Multi-lingual Advanced Speech and Text Research) について紹介する。このプロジェクトでは、人と人、人と機械のことばの壁を打ち破る技術開発を行うことを狙いとし、言語資源、情報分析、翻訳、音声コミュニケーションに関する研究を統合的に進めていく体制となっているが、本稿では、音声翻訳のアプリケーションであるVoiceTraに焦点をあてて説明する。

MASTAR project, Multi-lingual Advanced Speech and Text Research Project was launched on April 2008 at National Institute of Information and Communications Technology. The project includes research and developments aiming to break language barriers between different language speaking people and barriers between human and machines. While researches of language resources, language translation, and speech communication and information analysis are intensively and collaboratively conducted in the project, this paper concentrates, VoiceTra, the application of speech translation technology.

[キーワード]

言語資源, 翻訳, 音声コミュニケーション, 情報分析

Language resources, Translation, Speech communication, Information analysis

1 MASTAR プロジェクトの背景

近年、音声・言語処理は飛躍的な進歩を遂げている。

この理由の1つとして挙げられるのが、大量のデータ(コーパス)を収集し、機械学習により、自動的に処理系を構築する方法論(コーパスベース技術)が確立されたことである。このことは、これまで、研究開発フェーズでデータを収集し、システムを構築、その後、実用化に結びつけるといった従来の開発プロセスでなく、実際に使用される場面でのデータを直接収集し、それを機

械学習に用いることで、研究開発フェーズから実際の場面での性能向上を直接行える、新しい研究開発プロセスの到来と考えることができる。

さらに、WEBの普及、発達は、さらなる進歩を生むと考えられている。WEBの仕組み、WEB上の情報を利用することで、世の中にある固有名詞の取り込み、多言語辞書の構築やコーパス収集、単語の関係抽出、信頼性などの解析を行うことも可能になる。

このような背景の下で、企画されたMASTARプロジェクトの概要は次の通りである。

本プロジェクトの目標は、WEBを利用し、産

業界、社会とリンクした形で持続的に、音声・言語資源を蓄積、成長していく研究開発の仕組みを作ることである。

具体的には、以下の4つの研究開発を行う。

- ①総合科学技術会議の社会還元加速プロジェクトの1つに選定されたネットワーク音声翻訳に関する技術開発、社会実験、社会還元を行う（本稿 2 参照）。
- ②産業界とマニュアルなどを対象とした機械翻訳系を実現し、共通辞書、共通コーパスの蓄積、翻訳技術研究のポジティブ成長サイクルを確立する。
- ③ユニバーサルコミュニケーションの一環として、音声対話インタフェース技術の技術開発を行う。
- ④世界的言語資源を構築し、配信する。

本稿は、音声翻訳のアプリケーションである VoiceTra に焦点をあてて説明するが、『音声コミュニケーション技術の概要』（本特集号 3-1）、『多言語音声合成システム』（本特集号 3-3）、『多言語高精度を実現する専用翻訳システム』（本特集号 4-1）、『情報分析技術の概要』（本特集号 5-1）では、関連技術について説明している。

2 「言語の壁を乗り越える音声コミュニケーション技術の実現」プロジェクト

平成 20 年 4 月より内閣府社会還元加速プロジェクトの1つと認定された当該プロジェクトについて説明する。

言葉の壁はボーダーレス社会において大きな課題である。例えば、政府の『新成長戦略』^{*1}では「訪日外国人を 2020 年までに現在の 3 倍の 2,500 万人（経済波及効果 10 兆円、新規雇用 56 万人）にする」としているが、公共交通、宿泊施設、飲食店での外国語対応の遅れが、訪日外国人の最大の不満となっている。

内閣府社会還元加速プロジェクトとしてのゴールは、「国際化の進展の中で海外の人々と言語の壁を乗り越えて直接会話による交流を可能にすることができる自動音声翻訳システムに関し、当面の利用ニーズと今後 5 年程度で期待できる技術向上等を考慮して観光やショッピングなどの分野

における実証を企画・推進し、プロジェクト終了後短期間で産業界における事業化ベースでのサービスにつながるよう、その成果の社会還元の加速を目指す。」である。より具体的には、技術開発面では、地名などの固有名詞や幅広い話題に対応するためにネットワーク上に分散する翻訳知識を活用し翻訳端末と組み合わせるネットワークベース音声翻訳技術の確立に取り組むとともに、このシステムの円滑な実用化・事業化を図り、普及を促進するために音声翻訳コミュニケーション技術によるイノベーションの「見える化」の課題に取り組んでいる。「見える化」に際しては「開発サイド」と「利用者サイド」との密接な連携による実証を積み重ねることにより、音声翻訳技術の社会還元の加速を図るというデザインである。

3 ネットワーク音声翻訳プロジェクトの成果「VoiceTra」

3.1 VoiceTra とは

VoiceTra^{*2} はスマートフォンに話しかければ自動的に通訳するソフトウェアであり、以下その詳細を説明する。無償でダウンロードできるこのアプリケーションを使えば、例えば、図 1 と図 2 の組み合わせで示したような日本語と英語の対話ができる。電話するときのようにスマートフォンを耳元に近づけると短時間振動するので、これを合図に音声を入力すると、翻訳結果が音声で返る。図 1 の 1 番目の窓はシステムが認識した（聞き取った）結果、3 番目の窓が翻訳結果である。2 番目の窓は、「逆翻訳」（翻訳文をもとの言語に逆に翻訳する）の結果で、これを見て正しく翻訳できたかを確認できる。VoiceTra は 2010 年 8 月に公開し、2012 年 3 月時点で累計 60 万件を超えるダウンロード数を記録している。日本人の 200 人に 1 人が利用者であり、日本のスマートフォンの 33 台に 1 台にダウンロードされた計算になり、音声翻訳技術を多数の方に知っただけたといえる。さらに、後述するように、民間事業者と事業化が始まり、NICT 技術の社会還元の flagship の 1 つになっている。

*1 <http://www.kantei.go.jp/jp/sinseichousenryaku/>

*2 <http://mstar.jp/translation/index.html>

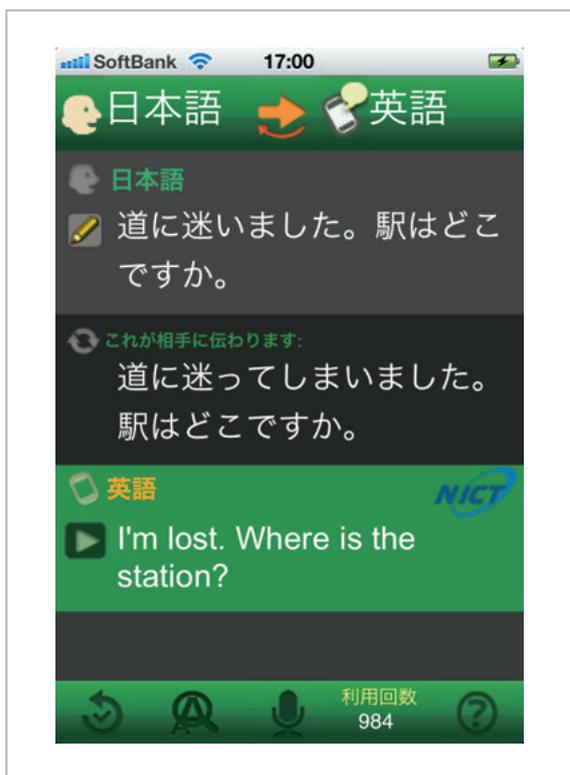


図1 VoiceTra 画面 (問い)

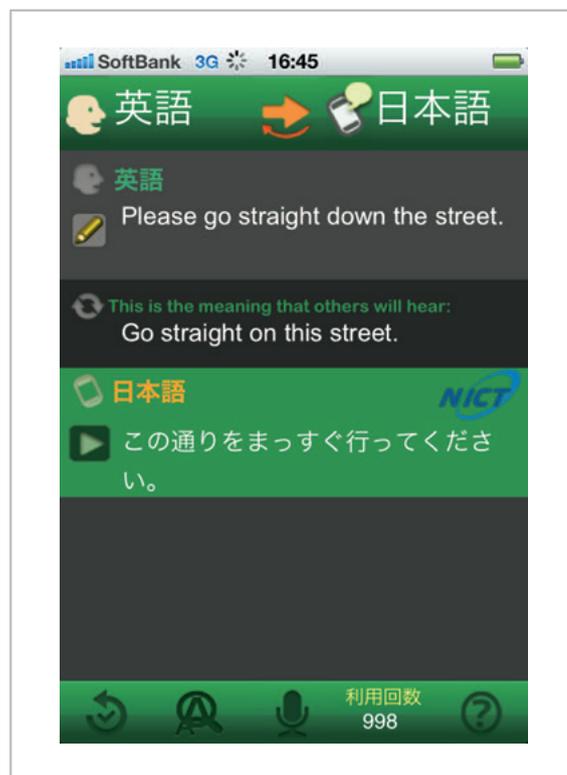


図2 VoiceTra 画面 (答え)

3.2 「基本の」音声翻訳技術

図3は、日本語音声認識されて日本語文章となり、さらに英語文章に翻訳され、英語音声に合成される例を表している。音声認識モジュールで、多くの話者の音声データから構成された音響のモデル（モデルは音声の要素である音素ごとに構成）と、入力音声との照合が行われて、音素列に変換される。次に、この音素列は、かな漢字で表記される単語列確率（言語モデルと呼ぶ）を最大化するように変換される。この変換では、日本語の大量のテキストから学習された、3つ組の単語列の生起確率をもとに、適切な単語列を求める。これをさらに翻訳モジュールで、日本語の単語列が対応する英語の適切な単語の選択、および語順の入れ替えが行われる。日本語の単語列と対応する英語の単語列を選択するために、日本語と英語の対訳文から学習された翻訳モデルを用いる。次に、語順を英語に合わせるため、大量の英語のテキストから学習される。3つ組の単語列の生起確率から英語として適切な単語列を求める。それを音声合成部へ送る。音声合成部では、まず、英語の単語列にあわせて発音、イントネー

ションを推定する。次に、それにあう波形を、大量の音声から学習された音声特徴量に合わせてフレームと呼ばれる時間単位で作り、それらを接続して音声合成を行う。

図の下方にある大規模コーパス（日本語のデータや対訳文や英語の音声のデータなど）を基盤にして、そこから自動的にシステムを構築するため、コーパスベースの技術と呼ばれる。

3.3 翻訳ソフトウェアの性能

VoiceTraは旅行会話を対象としている。その翻訳能力としては、おおよそTOEIC 600点の人に相当する。VoiceTraの特徴は、多言語対応であると同時に高品質な点にある。次のグラフ(図4)は、日本語への他の20言語からの翻訳について、広く利用されている多言語ソフトウェア（薄い灰色で表示）とNICTのソフトウェア（濃い灰色で表示）と、翻訳率（翻訳者が評価した意味が通じる率）で比較したものである。

3.4 ネットワーク型の「音声翻訳技術」と実用化

さらに、3.2の基本技術無線通信による

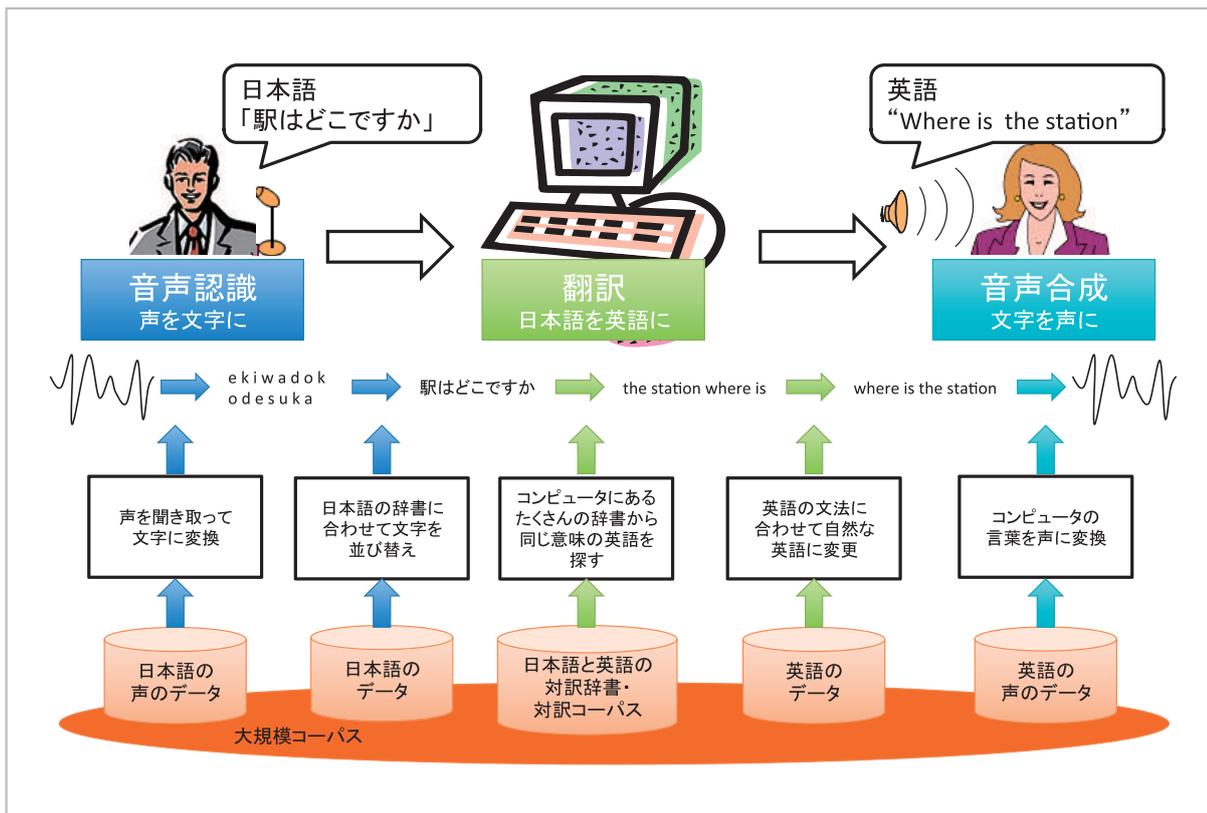


図3 音声翻訳の基本技術

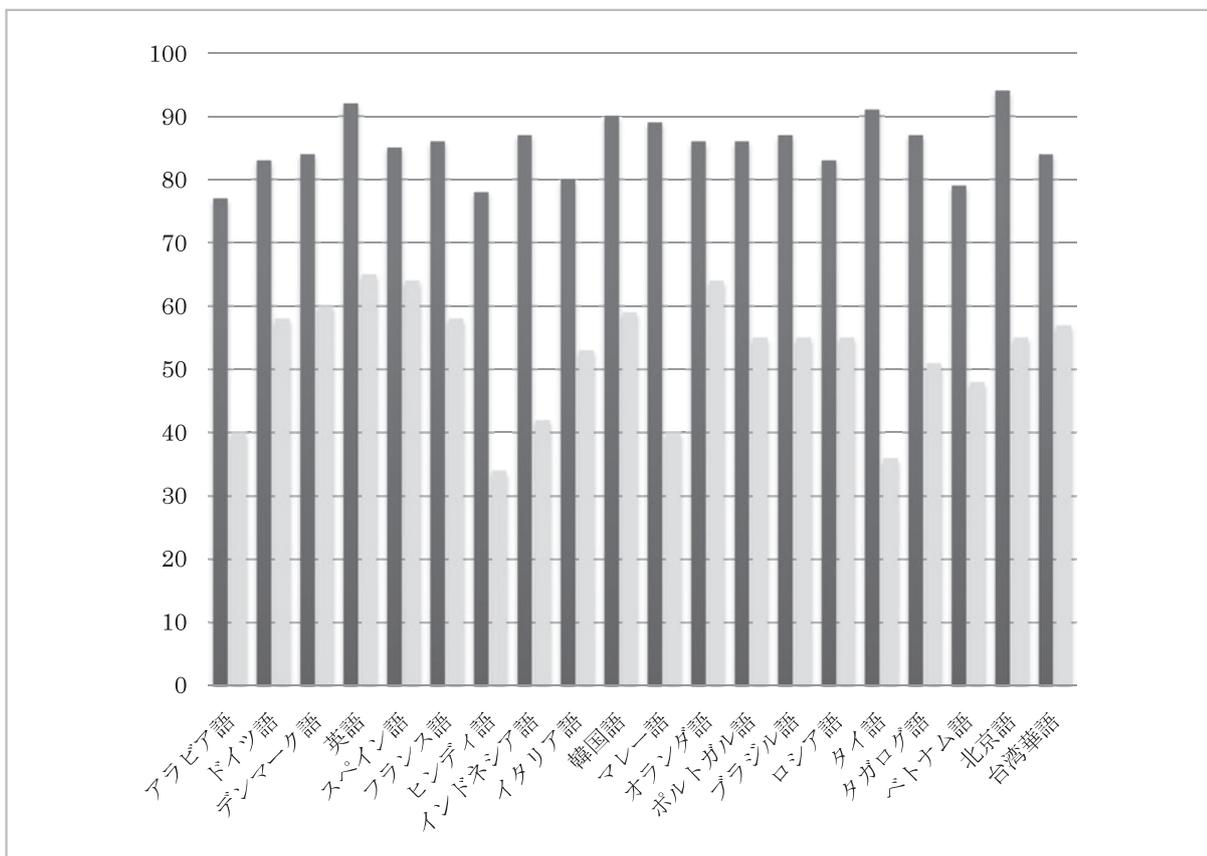


図4 翻訳率の比較（縦軸日本語への翻訳率、横軸が翻訳元の言語）

ネットワーク型にした。単一のコンピュータ内に閉じた実装では得られない可搬性、言語や語彙の拡張性、自律的な性能改善など特有の効果で、実用化を加速することができた。①クライアントを100g程度に軽くできるので可搬性が高くなり実用性が増した。②サーバはハードの制限がほぼないので、言語や語彙の拡張性は大きく、利用データに基づく自律的な性能改善が可能になった。実際に、VoiceTraの利用データの一部を利用し音声認識の改善を行ったところ、対象言語により差はあるが、5%から10%の精度向上ができた。

4 実証実験

音声翻訳技術は、処理時間が50%短くなる、誤り率が1/2になるといった基本性能だけでは実用化ができない。また、パターン認識系の他の技術と同様に誤り率をゼロにすることもできない。そのために、実際に利用者に使ってもらって、その結果をアンケートやログの形でフィードバックして、改良して実用化する必要がある。そのため、実証実験は有用である。これらの点は、『全国音声翻訳実証実験の概要』（本特集号7-2）、『VoiceTra実証実験の概要』（本特集号7-3）で詳述している。

また、この技術は1機関で多言語化することは、そのコストから現実的でない。この点を克服するために、音声翻訳システムに必要なプロトコルを標準化し、各国の機関が連携することが求められる。これについては『ネットワーク型音声翻訳の国際標準化と国際研究協力による標準化技術普及への取り組み』（本特集号7-4）で詳しく述べている。

5 事業化

NICTと成田国際空港株式会社(NAA)は2010年10月4日～翌年2月25日、商用化検証実験を実施した。成田国際空港に関連する固有名詞(エアライン名、観光地名、駅名、商品名等)1,600件を追加し、従来、語彙の不足から「穴のカウンターは何処ですか?」と誤認識されていた音声も「ANAのカウンターは何処ですか?」と正しく認識・翻訳が可能となった。NAAは、ネット

ワーク型の「音声翻訳技術」が外国人との「言葉の壁」解消のソリューションとなると判断し事業化に着手し、2011年12月末にアプリケーション(NariTraと呼ぶ、図5)を旅行者のスマートフォンにダウンロードするサービスを開始した。

VoiceTraは本件を含め5社に技術移転された。

社会還元加速プロジェクト「言語の壁を乗り越える音声コミュニケーション技術の実現」を5年計画(平成24年度末終了予定)で実施していたところ、同研究計画を上回る成果を出したので1年前倒しで平成23年度末に成功裏に終了できた。これは、社会還元加速プロジェクト6件の内唯一のことであり、非常に高く評価されたと言える。以下に、関連する内閣府/大臣・総合科学技術会議有識者議員会合資料の3.3節『今後の進め方』を引用する。「これらのことを総合的に考慮すると、普通の旅行者が、日本、英語、中国語圏でほとんど支障なく海外旅行を楽しめる環境の実現を加速するというプロジェクトの終了時の目標を、概ね達成したと考えられる。したがって、本プロジェクトは、当初のプロジェクト終了時期である平成24年度末を1年前倒して、平成23年度末で終了することが適当と考える。」



図5 NariTraのポスター

6 【音声翻訳研究の今後】

音声翻訳技術は1986年に基礎研究が開始されたが、VoiceTraは、同技術がついに実用化に至った、そのマイルストーンである。

現在の音声翻訳技術には、長い文は対応できないこと、文脈を理解できないこと、などの課題がある。NICTは、これらの困難な研究課題に取り組み、ニュースや会議の同時通訳という次の大きな夢の実現を目指している。

謝辞

本研究について、実証実験が不可欠であり、多数の方にご支援いただいた。個々に記して感謝する。

●実証実験に参加いただいたボランティアの方々

●実証実験を実施していただいた事業体:

ドーコン

北海道日本電気ソフトウェア(株)

(株)北海道観光振興機構

(株)ツーリズム・マーケティング研究所

(株)JTB 北海道

(社)ふらの観光協会

富良野市役所

知床斜里町観光協会

ラジオふらの

旅の宿ステラ (上富良野町)

ふる郷荘 (富良野)

(株)JTB グローバルマーケティング&トラベル

日本電気(株)

吉田 均 (山梨県立大学国際政策学部国際コ

ミュニケーション学科)

矢野 久 (山梨県)

石田幸司 (やまなし観光推進機構)

(株)駅探

(株)JTB 中部

(株)ATR-Trek

関西経済連合会

大阪市

大阪府

財団法人大阪観光コンベンション協会

奈良県

奈良市

奈良市観光協会

全日空

新阪急ホテル

上海フェリー

学生情報センター

熊本ソフトウェア

九州産業交通ホールディングス(株)

(株)JTB九州

NPO 法人産学官ながさき

(財)京都産業 21

国際電気通信基礎技術研究所

(株)東映京都スタジオ

日本電気(株)

(株)ウイルコム

(株)シルクラボラトリ

(株)インテージ

(株)UQ コミュニケーションズ

(株)彌榮自動車

神戸市

公益財団法人神戸国際医療交流財

みずほ情報総研(株)

徳島県庁 (商工労働部)

徳島大学病院

京都府

熊本聾学校

●商用化検証実験:

(株)オリエンタルランド

成田国際空港(株)

●商用化:

成田国際空港(株)

(株)FEAT

(株)ATR-Trek

●ITU-T:

内藤悠史 (三菱)

後藤義則 (NTT)

大久保 榮 早稲田大学 教授

Andrea Saks

●U-STAR 実証実験:

Dr. Seung Ku HWANG (Senior Vice President)

Pansak SIRIRUCHATAPONG, D.Ing.

(Executive Director)

Dr. Ir. Marzan A. ISKANDAR (Chairman)

Prof. Bo XU (Deputy President)

Dr. George VARKEY (Executive Director)

- Assoc. Prof. Thai Quang VINH (Director)
Prof. Kin Mun LYE
(Acting Executive Director)
Phuntsho TOBGAY (Director)
Dr. Waqar MAHMOOD (Director)
Dr. Sanat Kumar BISTA
(Chief Executive Officer)
Dr. Damsinsuren BAYANDUUREN
(President)
Dr. Tumur-Ochir SANJBEGZ (President)
Dr. Gihan WIKRAMANAYAKE (Director)
Dr. Sergio S. CAO
(Chancellor of UP Diliman)
Mrs. Michele SAUMON (Director, CNRS)
Dr. Leonel Sousa
(President of INESC-ID Lisboa)
- 社会還元加速プロジェクト「言語の壁を乗り越える音声コミュニケーション技術の実現」の関係者:
奥村直樹 総合科学技術会議 議員
清水慎一 立教大学観光学部 特任教授
辻井潤一 マイクロソフトリサーチアジア
首席研究員
匂坂芳典 早稲田大学大学院
国際情報通信研究科 教授
- 畑下 潔 経済産業省 産業技術環境局
産業技術政策課 課長補佐
柏木隆久 観光庁 国際観光政策課 課長
文化庁
中村 哲 奈良先端科学技術大学院大学
情報科学研究科 教授
栄藤 稔 (株)NTT ドコモ
サービス&ソリューション開発部
部長
渡辺文夫 KDDI(株) 理事 技術統括本部
技術開発本部長
三浦 貢 日本電気(株) マネージャ
木村晋太 (株)アニモ 取締役 副社長
潮田 明 (株)富士通研究所
メディア処理システム研究所
主管研究員
梶田直揮 内閣府 大臣官房審議官
大石善啓 内閣府 大臣官房審議官
伊丹俊八 内閣府 政策統括官(科学技術政
策・イノベーション担当)付参事官
(共通基盤技術(情報通信)担当)
大竹清敬 内閣府 政策統括官(科学技術政
策・イノベーション担当)付、参事官
(共通基盤技術(情報通信)担当)付

参考文献

- 1 柏岡秀紀, “音声コミュニケーション技術の概要,” 情報通信研究機構季報, 本特集号, 3-1, 2012.
- 2 志賀芳則, 河井恒, “多言語音声合成システム,” 情報通信研究機構季報, 本特集号, 3-3, 2012.
- 3 隅田英一郎, “多言語高精度を実現する専用翻訳システム,” 情報通信研究機構季報, 本特集号, 4-1, 2012.
- 4 安田圭志, 松田繁樹, “全国音声翻訳実証実験の概要,” 情報通信研究機構季報, 本特集号, 7-2, 2012.
- 5 松田繁樹, 安田圭志, 河井恒, “VoiceTra 実証実験の概要,” 情報通信研究機構季報, 本特集号, 7-3, 2012.
- 6 堀智織, “ネットワーク型音声翻訳の国際標準化と国際研究協力による標準化技術普及の取り組み,” 情報通信研究機構季報, 本特集号, 7-4, 2012.

(平成24年6月14日採録)

すみ た えい ちろう
隅田英一郎

ユニバーサルコミュニケーション研究所

多言語翻訳研究室室長

博士（工学）

自然言語処理、機械翻訳

eiichiro.sumita@nict.go.jp