

行政分野への音声認識ソリューションの展開

The Development of a Solution with Voice Recognition for the Administrative Field

行政分野において、IT 技術を用いた各種行政サービスは、市民生活へ大きな影響を及ぼしており、その推進にあたっては利用者と生活者の視点に立って情報化を進めていくことが重要である。(株)日立東日本ソリューションズ(以下日立 TO)では、株式会社アドバンスト・メディアの「音声認識技術 **AmiVoice**® (アミボイス)」の行政分野への新たなソリューションの検討を開始した。電子自治体で各種行政サービスの利用促進のためには、高齢者やパソコンの苦手な人にも使いやすい環境が必要という声も多くあがっており、この音声認識ソリューションはそのようなニーズに対する一つの解である。

河野 吉宏
川下 賢司

Kohno Yoshihiro
Kawashimo Kenji

音声認識技術分野の市場規模は、富士キメラ総研によると 2007 年 60 億円となっており、2012 年には、倍の 137 億円へ拡大すると予測されている。本論文では、音声認識を利用したシステムの紹介、および日立 TO としての今後の音声認識ソリューションの展開方法について、音声認識技術の ASP による提供やテキストマイニングとの連携、行政の新たなインターフェイスとしての活用方法などについて述べる。

1. はじめに

音声認識システムとは、人間の音声言語をコンピュータによって解析を行い、会話もしくは話している内容をテキストデータとして変換処理するシステムのことである。音声を処理して得られた音声データをテキストデータに変換することによる新たな入力方法として注目されている。この音声認識の応用分野の展望は大きく、公共、産業、金融などあらゆる業種の様々な業務において適用されており、ユビキタス社会におけるインターフェイスを考える上で重要な技術要素となっている。

行政分野における IT 技術を用いた各種行政サービスは、市民生活へ大きな影響を及ぼしており、その推進にあたっては利用者と生活者の視点に立って情報化を進めていくことが重要である。このためには、地域が抱える諸問題や市民ニーズを明確にした上で、行政課題に即したサービスの構築と活用のソリューションを提案していく必要がある。

今回日立 TO で推進する音声認識ソリューションは、株式会社アドバンスト・メディア社の音声認識技術 **AmiVoice**® の高度な音声認識技術、および行政分野や他分野への音声認識ソリューションの多数の適用実績を

もとに、行政分野への新たなソリューションを検討するものである。音声認識技術 **AmiVoice**® は「不特定話者認識(事前のトレーニングを必要としない)」「言語モデルを構築し改訂することで高水準の認識精度を維持」「使いやすいインターフェイスにより認識結果を短時間で編集可能」などの特徴を備えており、音声認識ソリューションとして、国内トップレベルの導入実績を誇る国内有数の音声認識ソリューション製品である。

以下、第 2 章、第 3 章で行政分野における音声認識ソリューションの適用方法の紹介を行い、第 4 章では音声認識技術を活用したソリューション展開について述べる。

2. 行政分野における音声認識技術の必要性

2.1 電子自治体での各種行政サービスの利用促進

NTT レゾナンド株式会社および株式会社三菱総合研究所が 2004 年に発表した「電子自治体に関する一般生活者の意識」に関する調査結果¹⁾によると、地方自治体の行政サービスに対する満足度について、「サービスを受けられる時間が限られている」という時間的制約への不満を挙げている人が 37.0%と最も多く、「手続きごとに申請窓口が別であるため非効率で煩わしい」27.1%、「サー

ビスを受けられる場所が少ない」20.7%などと合わせ、利便性と効率性を求める声が大きいがわかる。また、今後電子サービスを導入することに対しては、過半数が期待感を持っていることがわかり、世代別に見ると、高齢になるに従い期待している比率が増えている傾向にある。

総務省が発表した「地方公共団体における行政情報化の推進状況調査（2008年4月1日現在）等の取りまとめ結果」²⁾によると、行政サービスに関するオンライン化について、都道府県では全団体、市区町村においては約8割の団体がオンライン化を実施済みであるものの、総務省が選定したオンライン利用促進対象手続きに関しての2007年度のオンライン利用実績は、約2割と低い値に留まっているのが実情である。すなわち、オンライン化が推進され始めたころに抱いていた利便性や効率性などの期待感を満足させることのできる状態には至っていないものと推測できる。

これらの事実から電子自治体における各種行政サービスの利用促進のためには、ユーザ視点での利便性と効率性の向上を図り、さらには、高齢者も利用することを念頭に置いたオンラインサービスの提供を行う必要があるとのニーズが読み取れる。

上記のニーズに対し、現在各地方自治体では、オンライン利用の促進に向けて利便性向上、提供手段の改善、メリットの拡大、広報、普及の強化など各種措置を講じている。我々は、これらの措置に加えて、各地方自治体が急務として講じるべき措置として、窓口業務の改善が必要であると考えている。オンライン利用促進のため、利便性の向上や提供手段の向上など、オンラインサービスの使い勝手が良くなれば利用率も高まっていくと考えられるが、一方で窓口に行かなければならない手続きや、対面で職員に相談をしながら申請する必要のある手続もあることも事実である。また、インターネット上でオンラインサービスを提供した場合も、パソコンを利用できない高齢者には利便性を訴えることが出来ない。オンラインサービスの利便性、利用率向上という問題として矮小化するのではなく、従来からの窓口対応を含めた住民視点の全体最適化を考える必要がある³⁾。

窓口対応の改善として、音声認識技術を活用した問い合わせ窓口システムの導入をすることで、高い認識精度を活かした通話内容のテキスト化によりコンプライアンス

の強化や問い合わせ窓口の適切かつ有効な運営をサポートし、住民視点での全体最適化につなげることができるものと考えている。

2.2 首長及び議会での意思決定の迅速化

議会は地方自治体の意思決定機関である。その意思決定には地域住民の意向が反映される必要があり、意思決定過程の透明化から情報公開のニーズが強まっている。そのため、これまで自治体では迅速な情報公開のために、議事録作成の効率化、迅速化と経費の削減が課題となっていた。音声認識技術を導入することで、議事録作成事務の効率化、意思決定の迅速化など「生産性」を向上させることが出来る。住民からのニーズを客観的にかつ効率的に整理することは、首長及び議会の意思決定にとって、また、これからの行政にとって重要なことといえる。

3. 音声認識ソリューション事例

3.1 音声認識問い合わせシステム

3.1.1 機能概要

本システムは、市民から問い合わせ窓口を受けた問い合わせに対応する通話内容を、音声認識サーバで処理することで、担当者の端末上に問い合わせに必要な情報を表示する仕組みを実現している。

図1に問い合わせ窓口システムの概要を示す。市民からの問い合わせ電話は、まず問い合わせ窓口システムのPBX装置および音声認識サーバで処理され、オペレーター（問い合わせ担当者）の多機能電話（MFT）へ接続される。その際に、音声認識サーバでは、通話内容をリアルタイムにテキスト化し、オペレーター（問い合わせ担当者）の端末画面上に抽出したキーワードを表示する。また、担当者端末からは、各キーワードに関して、FAQシステムを通じて過去の問い合わせ内容と回答内容や、NGワード、問い合わせ回答に必要な資料、補足情報等を検索した結果を画面に表示することができる。なお、この際に通話した通話内容は全て通話録音サーバで録音される仕組みとなっており、あとから会話内容をテキスト化することも可能である。市民からの問い合わせから、問い合わせ担当者の回答、管理者によるチェックの流れを図2に示す。

3.1.2 導入効果

(1) 対応効率・品質向上・コンプライアンス強化

本システムでは、音声認識の技術を活用することで、問い合わせ担当者の対応効率と品質の向上を実現することが可能である。オペレーター（問い合わせ担当者）は、通話内容から FAQ システムを通してリアルタイムに表示される情報や資料をもとにして問い合わせ対応をすることが可能となる。そのため、担当者に専門的な知識、スキルがない場合でも、担当者育成に要する教育時間を削減することができる。また回答の保留、折り返し、転送などを減らすことができるため、住民の求める情報を迅速に回答することができるとともに、呼応回数の増加により、住民の満足度向上にも寄与することが可能となる。

また、本システムでは、スーパーバイザー（管理職）向けの機能として、問い合わせに対応する複数のオペレーター（問い合わせ担当者）の通話内容を音声と文字による同時モニタリングする機能を備えている（図 2 の④）。会話中の NG ワードによるスーパーバイザーへの注意、警告機能、変換履歴をクリックして通話の一部を聞き返す機能なども備えており、説明責任項目のチェックや、通話監視の強化、情報漏洩の防止など、コンプライアンス強化の一翼を担うことが可能である。

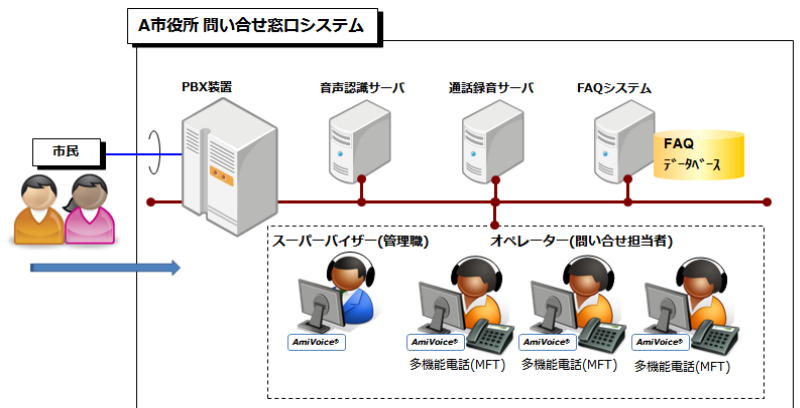


図 1 音声認識問い合わせシステムの概要

(2) 行政マーケティング

テキスト化された住民との通話内容や問い合わせ情報はサーバに保存されデータベース化される。録音サーバでの録音データを、あとから一括してテキスト化することも可能である。データベース化された問い合わせ情報は、自治体にとって重要な資産となる。テキスト情報からテキストマイニングツールと BI ツールを用いてキーワード分析やクレーム分析を行うことで、住民の視点に立ったニーズ分析を行うことが可能となり、新たなサービスの提供や、既存サービスの改善を行うことが可能となる。

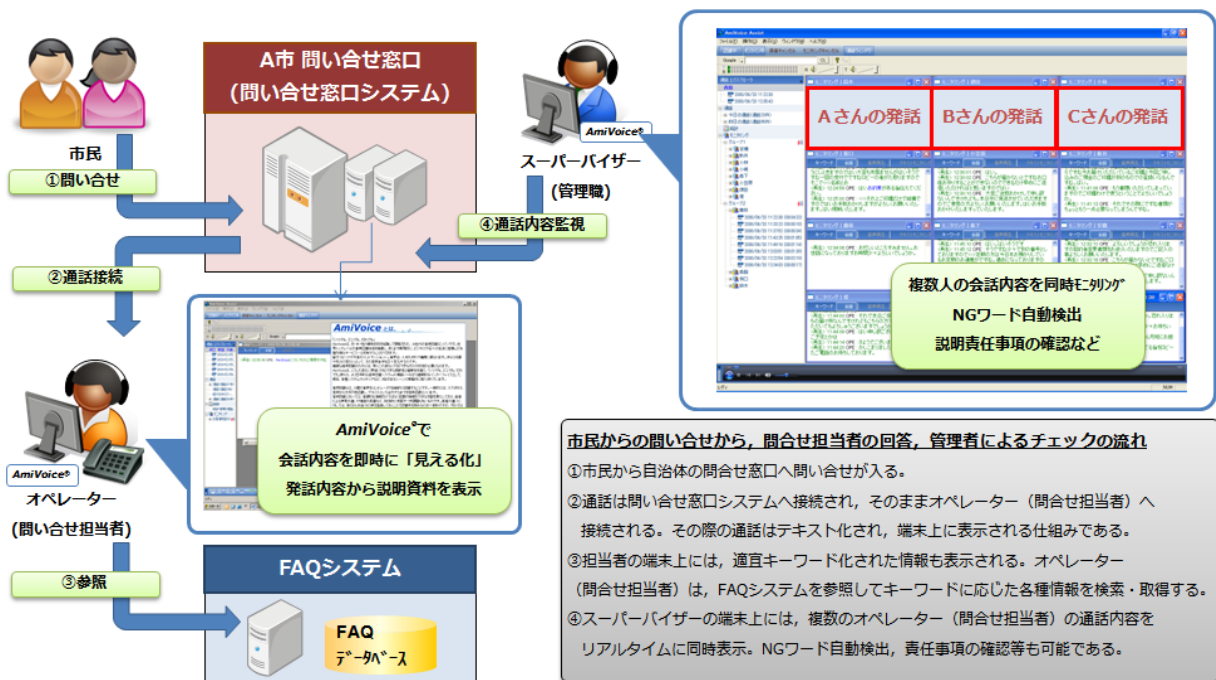


図 2 問い合わせから回答、チェックの流れ

3.2 議事録作成支援システム

3.2.1 機能概要

議事録作成支援システムは、議会の議事録作成を支援するシステムであり、議会での発言をテキスト化するという主に自治体を対象としたシステムである。

図 3 に議事録作成支援システムの概要を示す。自治体の議会での音声を、ミキサー・コンプレッサーを利用して認識しやすい音声信号として調整する。調整した音声信号は音声認識技術を利用し、録音および文字データとして認識する。さらに、編集機能を利用し誤認識部分や言い回しを修正し、議事録として編集する。音声データについては上記のようなリアルタイムに認識する方法の他、録音された音源の品質によっては、あらかじめ録音された音源を利用して会議後に音声認識および編集をする方法もある。なお、複数人による同時編集が可能であり、即時性を求められる場合においても編集速度を上げることが可能である。

3.2.2 導入効果

(1) 議事録作成作業の効率向上

議事録作成支援システムを導入することで、議事録作成に関わる作業量を低減することが可能である。ある事例においては、テープレコーダを使用した従来の書き起こしに約 6 時間程度要していた。(最初からキーボード入力した場合)しかし、音声認識技術を利用した場合は、認識率に応じて約 1.5 時間(認識率 90%の場合)～約 3 時間(認識率 62%の場合)の作業量となる。作業量としては、同 75%～50%の削減効果となる。これは、認識結果を修正するのみとなるためである。

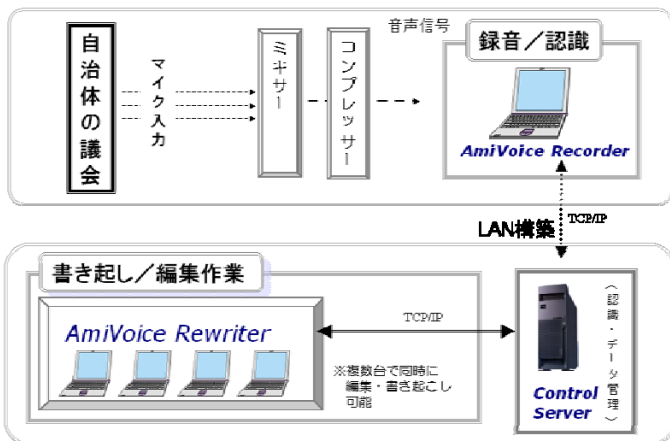


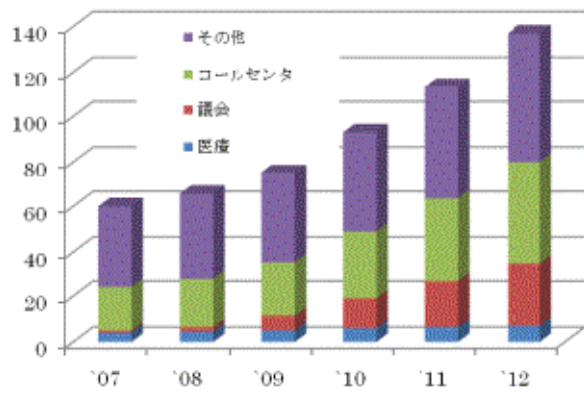
図 3 議事録作成支援システムの流れ

4. 音声認識ソリューション展開

4.1 分野の展開

音声認識技術分野の市場規模は、富士キメラ総研⁴⁾によると 2007 年 60 億円(実測)となっており、2012 年には、2 倍以上の 137 億円へ拡大すると予測されている。導入するために必要なハード、周辺ソフトウェア、システムインテグレート及び周辺ユーザプログラム開発など音声認識技術により誘発される周辺需要を考慮すると、その数倍の市場規模が存在すると予想できる。

次に 2007 年度に音声認識技術が比較的活用されていた「コールセンタ」、「自治体向け議会議事録」分野の今後の伸びは、「コールセンタ」は 2007 年度 19 億円に対し 2012 年度 45 億円と 3.8 倍の伸び、「自治体向け議事録システム」は 1 億円に対し 28 億円と大幅な 28 倍の伸びと予想している。自治体の説明責任の拡大と財政縮小を背景とした自治体の事務効率化が課題となっている状況を鑑みると、議会の音声認識システムの市場は今後急速に拡大すると思われる。なお「医療」分野は 4 億円に対し 7 億円と 1.8 倍の伸び、「その他」分野は主に民間企業(コマンド制御など)、文教(e-learning など)を示すが、36 億円から 57 億円と約 1.5 倍の伸びとなっている。



※2007 年度は実績、2008 年度以降は見込み。
※単位：百万円

図 4 音声認識技術分野の市場規模

4.2 機能の展開

4.2.1 ASP サービス展開

「総務省における電子自治体推進の主な取込」(総務省自治行政局 平成 20 年 4 月)⁵⁾によると、2010 年度までに「利便、効率、活力を実感できる電子自治体を実現する」とした上で、「共同アウトソーシングの推進」が挙げられている。音声認識システムの共同アウトソーシ

グによる提供は、低コストで高いセキュリティを提供するものであり、同時に、運用は地元企業が受託するという地域 IT 産業の活性化を図る地域貢献も視野に入れている。規模の大きい自治体であれば、システム導入とそれに見合う投資効果を得ることができるが、中小規模の自治体の場合、予算に伴い導入の壁が高いと予想される。このため、「音声認識問合せシステム」「議事録作成支援システム」を ASP サービスとして提供することを検討している。

(1) ASP サービスの構成

ASP サービスの構成イメージを図 5 に示す。構成としては大きく音声認識処理部とフロントエンド処理部に分かれる。音声認識処理部は音声認識技術本体を持ち、複数自治体の参加を想定し拡張性と性能を有する構成としている。また、機能毎にサーバを分割することによりスケラビリティを確保する。フロントエンド処理部には、認証処理、課金処理、データ送受信機能に対応する。認証処理においては、ID とパスワードを顧客側に配布し利用する運用を想定しているが、公的個人認証カードのようにセキュリティが高い認証方法にも対応するなど、ユーザの要望に応じた認証方法を可能とする。また、課金処理については、基本利用量プラス使用量に応じた従量制を検討している。今後、各サービス用のオプションメニューを準備し展開して行く予定である。

(2) ASP サービスの機能

ASP サービスの機能として、拡張性、性能、セキュリティに分類しまとめる (表 1)。

拡張性および性能については、仮想化ソフトを利用した仮想 OS 上に構築し、物理的にサーバ増設した場合で

も、仮想 OS イメージを配布するだけで、増設可能とする。上記対応により、開催時期が集中する議会においても性能を確保し、また、サーバ増設作業を可能な限り簡素化する事により、一時的なサーバ増設を可能とする構成としている。

セキュリティに関しては、ASP サービスとして提供しているため、一律の水準を確保することが可能である。サーバールームをデータセンタ内に設置することにより物理的なセキュリティを確保し、構成する OS とソフトウェアを一元管理することにより各種脆弱性にも対応することが可能である。また、送受信データのセキュリティについては、次のとおり対応する。例えば、公開用の議事録においては、非 SSL を使用しデータ通信速度を優先する。秘匿性が高い社外秘の議事録の場合は、SSL 通信を使用してセキュリティを確保する。

表 1 ASP サービスの機能

分類	概要
拡張性	・ 仮想化技術を利用した拡張性
性能	・ 処理負荷に応じサーバ追加による負荷分散
セキュリティ	・ 一律のセキュリティレベルを確保 ・ データセンタによる物理セキュリティ確保 ・ 通信セキュリティの確保

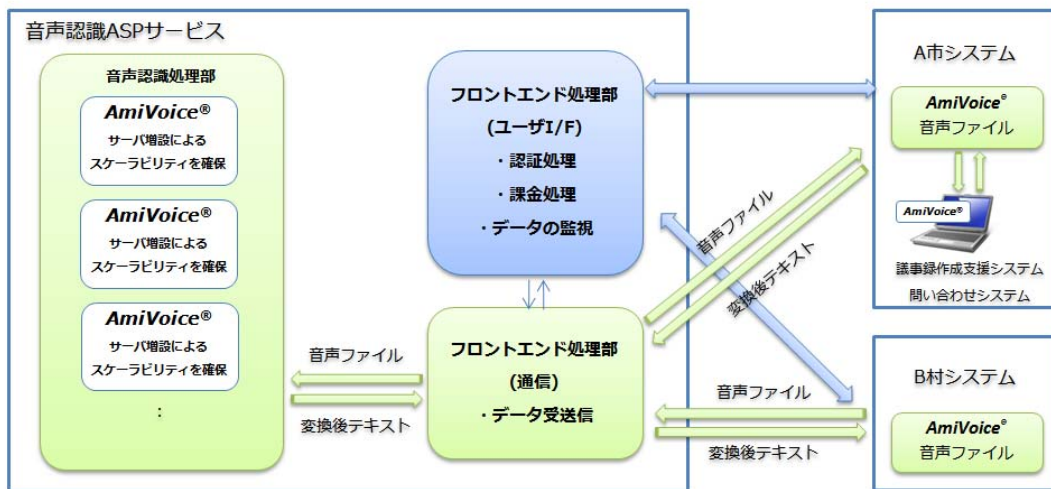


図 5 ASP サービスの構成イメージ

4.2.2 テキストマイニングと BI ツールとの連携

問い合わせシステムにおいて、問い合わせ時に照会を行う FAQ データベースからの検索時にテキストマイニングツールを活用することで、キーワードに関連する問い合わせの自動抽出などを行うことが可能となる。精度の高い検索結果を得ることで、対応品質を向上することが可能である。また、問い合わせ履歴をテキストマイニングツールで分析することで、よくある問い合わせや問題点の抽出を行い、オペレータ毎の対応品質を「見える化」しトレーニングに生かしたり、問い合わせ頻度に基づく FAQ の策定などに利用し、問い合わせ業務の改善につなげることができる。

行政の意思決定においてベースとなるものは住民のニーズであり、住民のニーズをもとに意思決定するのは首長及び議会であり、意思決定を執行するのが行政の基本的役割である。住民ニーズは問い合わせ窓口への問い合わせや住民アンケート、パブリックコメント、行政評価、行政監査さらには住民との各種会合など、様々な媒体を通じて行政は吸い上げることができる。様々な場面、様々な媒体により得られた住民ニーズを音声認識技術によりテキスト化したものを BI ツールにより整理し分析することにより、住民ニーズの抽出や苦情分析などを実施することで意思決定の迅速化、また、人的な分析による主観を極力排除することが可能となる。

5. おわりに

日本の 2050 年の全人口に対する 65 歳以上の老年人口の比率は 35.7%にまで上昇する見通しである⁶⁾。このように高齢化が進展し、労働力人口が不足してくる中では、高齢者への生活支援や、高齢者の介助業務への支援のニーズが増加してくる。音声認識技術は、今回紹介した「生産性」を向上させるツールとしての活用方法から、「聴覚から視覚へインターフェイスを変換する」という捉え方をすることにより、福祉的な観点での活用方法を見出すことができる。この考え方は“社会的な弱者”といわれる層に対し、ハンデの極小化が重要な責務のひとつといえる行政には有効な考え方である。

今後は、継続的な市場ニーズの把握とお客様の声を元に、より多くのお客様にご利用いただけるように、音声認識技術を利用したソリューションメニューの充実を図っていく。そして、これら幅広いニーズにこたえるソリューションメニューの充実に加え、テキストマイニングツール、BI ソリューションなどの日立 TO の独自技術と

連携することで、音声認識市場における日立 TO のソリューションの優位性をさらに高め、来るべきユビキタス社会のヒューマンインターフェイスとしての音声認識技術の発展に貢献していく所存である。

参考文献

- 1)NTT レゾナント株式会社および株式会社三菱総合研究所,「電子自治体に関する一般生活者の意識」調査結果
<http://research.goo.ne.jp/database/data/000141/>
- 2)総務省, 地方自治情報管理概要「地方公共団体における行政情報化の推進状況調査(平成 20 年 4 月 1 日現在)等の取りまとめ結果」
http://www.soumu.go.jp/menu_news/s-news/2008/081031_1.html#bs
- 3)日本ユニシス 森山勉, 成熟期を迎えた自治体コールセンターの今後の在り方(第 2 回)
<http://japan.internet.com/column/public/technology/20080416/5.html>
- 4)富士キメラ総研, 音声認識/合成システム市場の現状と将来展望
- 5)総務省, 総務省における電子自治体推進の主な取組(概要)
http://www.soumu.go.jp/denshijiti/pdf/index_070719_01.pdf
- 6)文部科学省, 平成 18 年版 科学技術白書
<http://japan.internet.com/column/public/technology/20080416/5.html>



河野 吉宏 1999 年入社
公共ソリューション本部
社会基盤ソリューション部 第 4G
官公庁向けソリューション提供
ykohno@hitachi-to.co.jp



川下 賢治 2000 年入社
公共ソリューション本部
社会基盤ソリューション部 第 4G
官公庁向けソリューション提供
kenjik@hitachi-to.co.jp