
「だれでもポータル」がもたらす新しい
インターネット社会の実現に向けて
(音声認識技術と XML 技術の融合による情報バリアフリーシステム)
株式会社 P F U

■ 執筆者 Profile ■



岡野 孝行

1999年 (株)PFU入社
2002年 現在 東海支店所属
営業業務担当



千葉 辰典

2000年 (株)PFU入社
2002年 現在 第四システム統括部所属
フィールド業務担当

■ 論文要旨 ■

近年 IT 化の発展・浸透により、インターネットで多種多様な情報を入手できるようになった。筆者もその恩恵に与っている一人だが、私生活や日頃の業務から思うことがある。それは必ずしも欲しいときに欲しい情報が得られないということと、すべての人がインターネットを使えるわけではないということである。

今やインターネットが利用できないと生活が困るほど浸透しており、環境や人により使えない状況（これをデジタルデバイド（情報格差）という）を改善できないだろうか？と考え、我々は「だれでも・いつでも・どこでも」使える「電話」に着目した。

本論文では、音声認識技術と XML 技術を活用し、利用者にインターネットを意識させず、同等情報を取り出せる仕組み【だれでもポータル】を構築する過程や工夫したポイント、これにより得られた評価を述べている。

今後はだれでもポータルを、デジタルデバイドを解消する手段だけではなく、コマース（商取引）用途の有効な手段として発展させていく。

■ 論文目次 ■

1. はじめに	《 5》
1. 1 当社概要	
2. IT化進展により出現した問題点	《 5》
2. 1 取り組みのきっかけ	
2. 2 情報サービスの現状と課題	
2. 2. 1 広がる情報格差	
2. 2. 2 本当に必要とされる情報提供の仕組みとは	
2. 3 電話を使ったシステム	
3. 構築にあたっての課題	《 9》
3. 1 従来システムとの比較	
3. 1. 1 だれが話しても大丈夫なシステム	
3. 1. 2 迷子にならないストーリー設計	
3. 2 XML 技術の適応	
3. 2. 1 電話機以外への対応	
3. 2. 2 管理業務の効率化	
3. 3 スモールスタート	
4. だれでもポータルによる情報提供	《 13》
5. 評価と今後の課題	《 13》
5. 1 だれでもポータル評価	
5. 2 だれでもポータルの提案・展開	
6. おわりに	《 15》
参考文献	《 15》

■ 図表一覧 ■

図1	インターネット世帯普及率	《 6》
図2	インターネット利用率	《 7》
図3	インターネット非利用者のインターネットを利用しない理由	《 7》
図4	インターネット非利用者のインターネットを利用する条件	《 8》
図5	音声認識の仕組み	《 10》
図6	ストーリー設計例	《 11》
図7	「だれでもポータル」概要	《 13》
図8	想定利用シーン	《 14》
表1	音声認識エンジン比較表(抜粋)	《 9》
表2	アンケート結果	《 13》

1. はじめに

1. 1 当社概要

当社は、富士通株式会社・株式会社内田洋行・松下電器産業株式会社などの合弁会社であり、以下が主な業務である。

① ソリューションの提供（ハードウェア，ソフトウェア，サービス）

② EMS (Electronics Manufacturing Services) の提供

③ サーバ，ディスクアレイ，周辺/応用機器のハード/ソフトウェアの研究開発/製造

その中で，IT 活用によるソリューション（デジタルコンテンツ配信，Linux，XML，セキュリティ，インフラ構築，マネージメント）をはじめ，サポートサービス（24H7days 体制），IT 製品の設計/開発/製造支援サービス（「ProDeS (Product Design Services)」）を得意としている。

本論分ではデジタル音声技術と XML 技術を融合した電話によるインターネット活用システムについて述べる。

2. IT 化進展により出現した問題点

2. 1 取り組みのきっかけ

近年，パソコンや携帯電話に代表される情報機器を使用し，インターネットでの情報入手が容易となっている。しかし，周囲を見渡すと必ずしもそうでないことに気付くことがある。

それは必ずしも欲しいときに欲しい情報が得られないということと，皆がインターネットを使えるわけではないということである。例えば，外出先で何か知りたいことがあった場合，パソコンは使えないし携帯電話のインターネットもどこを探せばいいのかかわからない状況に陥り，不便を感じるが多々ある。そこで電話をして誰かに調べさせるのだが，その人がインターネットを使わない人であればそこで行き詰まってしまう。筆者も日常の中で，また業務の中でこのような問題に遭遇することがある。

パソコンや携帯電話を使いこなしてインターネットを利用する人は，まだまだ一部の人々に限られていないだろうか。例えば，情報機器の操作が必ずしも容易ではないことから高齢者は使うことに抵抗を感じていたり，モバイル環境（外出先）での操作性の悪さや通信環境に影響され思うように使えなかったりと，インターネットは利用面，操作面，技術面など解決しなければならない課題はまだ多いように思える。

我々は，今現在の環境下でできる限り利用する人や環境に影響されずにより簡単に情報を入手でき，「だれでも・いつでも・どこでも」インターネットの恩恵を享受できる仕組みの必要性を感じるようになった。

現在，政府主導による e-Japan 戦略の下，2005 年の高度情報社会実現に向けて取り組みが進められている。我々の生活する地方自治体で IT インフラ（情報技術）の整備が進められているのも周知のことであろう。しかし，IT インフラを利用して豊かな個性を伸ばす社会ができる一方，懸念材料も噴出している。“利用できる人とできない人の間に経済格差が生じてしまうのではないか”，“IT は我々が理想的な目的と考えるのと違い，不公平を拡大してしまうような道具になってしまうのではないか”，という懸念である。

e-Japan 戦略が成功するためには、できるだけ多くの国民や企業がネットワークに参加することが重要である。そのためには、IT の利用機会を増やすと共に IT 活用能力の格差（デジタルデバイド）が生じないような施策や、既に生じている格差を解消するための施策が必要となる。

そこで我々はこの“悩める” e-japan 戦略に応えるべく「e-japan ワーキンググループ」を設置し、デジタルデバイド解消に向けた取組みを開始したのである。

2. 2 情報サービスの現状と課題

ここ数年で IT 化が急激に進展しており、**図 1**のように、一般家庭へのパソコン・携帯電話の普及率と共にインターネットの利用率が爆発的に伸びてきている。

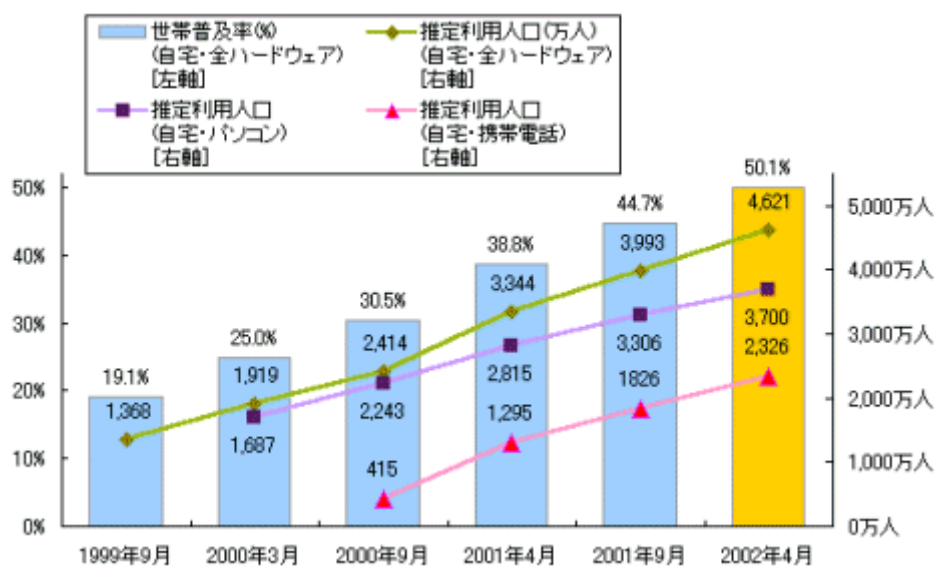


図 1 インターネット世帯普及率

地方自治体が行政の実態を住民に公開し、住民の選択と監視をもとに行政が実施される「見える自治体(ビジブル・コミュニティ)」を実現するに当たり、インターネットは情報公開の手段として最も適しているといえる。現在、各自治体のホームページ開設率（平成 13 年度）は省庁及び指定政令都市では 100%，市区では 83.6%，町村では 61.1%となっている。しかしながらそのホームページの利用率は 40%台に留まり、自治体関連者の頭を悩ます種になっているのが現状である。

その原因として挙げられるのが、“多種多様な情報は公開されているが、それを見れる人が限られている”ということではないか。ここではインターネットが使えない人を指すが、例えばそれは高齢層の方々であり、また家や会社以外の外出先では使えないという利用環境である。

2. 2. 1 広がる情報格差

デジタルデバイドは性別や年齢層、環境が影響し生じている格差である。**図 2**を見ると、20～30 代の年齢層は男女ともインターネットを利用している率が高いが、40 代以降では利用率が低くなる傾向が顕著に表れている。

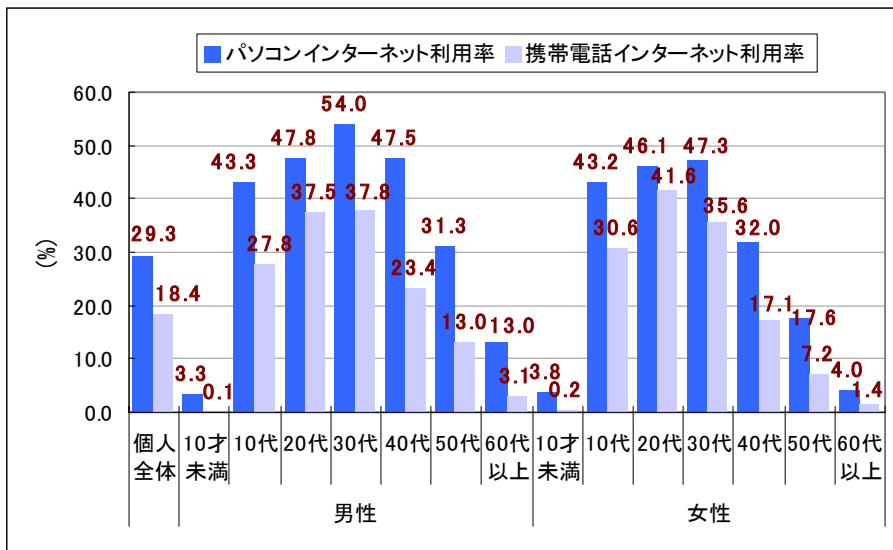


図2 インターネット利用率

日本は高齢化社会国であり、2002年10月に発表された平均寿命も世界一である。現在6人に1人が65歳以上であり、10年後には4人に1人にまで進むといわれている。将来、情報から“落ちこぼれ”ないようにするには、40代以降の年齢層をインターネット利用人口にするのが急務であると考えられる。

ではどのような手段でデジタルデバイドを解消するのが理想的なのか？

2.2.2 本当に必要とされる情報提供の仕組みとは

世界一のIT国家を目指す日本の最大課題は、デジタルデバイドの解消である。高齢者や障害者にやさしい情報提供の仕組み、「情報バリアフリー」整備が政府主導により進められてきているのもこのような背景からである。

我々は情報バリアフリーの実現を考えるに当たって、下記図3と図4の赤丸内の項目に注目した。

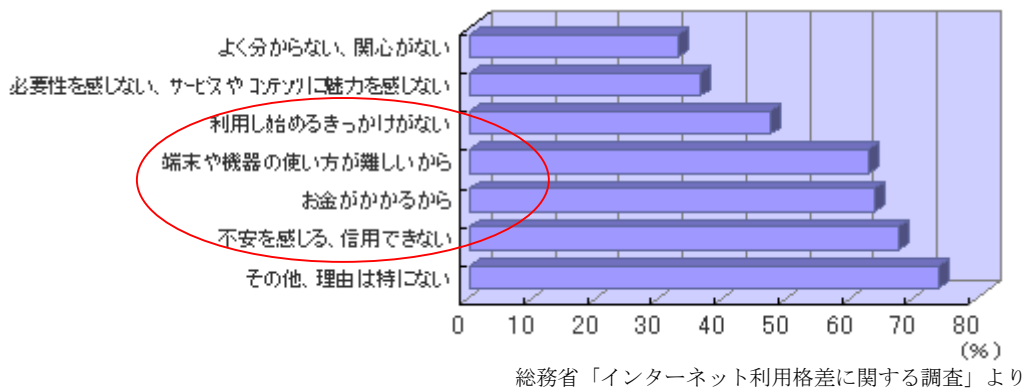


図3 インターネット非利用者のインターネットを利用しない理由

図3の「インターネットを利用しない理由」で、“不安を感じる、信用できない”や“端末や機器の使い方が難しい”が非常に多い。特に40代以降の年齢層には限らない事だが、新しい製品（技術）導入の際に感じる不安や信頼性が、消費者の購入（導入）意欲を下

げていることは確かである。ここで言う不安とは決してセキュリティ面だけではなく，“ちゃんと使いこなせるか”だと推測する。特に高齢者の方には，パソコン操作は難しく，また習得まで時間もかかるので敬遠されてしまう。

また，“利用し始めるきっかけがない”ということも大きな原因を占めている。

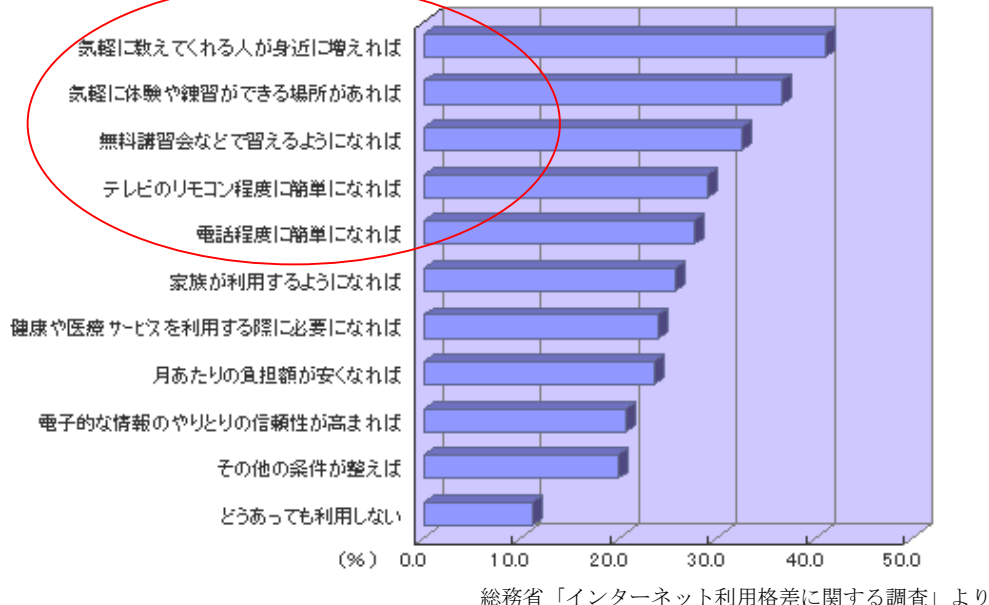


図4 インターネット非利用者のインターネットを利用する条件

そこで図4の「インターネットを利用する条件」を見てみると，具体的にテレビのリモコンや電話程度に簡単な仕組みになれば，インターネットを使ってみようと思っていることが分かる。つまり“使い方が難しいから使わない”のであって，費用もかからず，誰かに教わる必要もない“【電話】くらい操作が簡単”であり，そしてそれがいつの間にかインターネットの恩恵を受けていると感じたとき，その利用者はもっと高度な利用法を目指すのではないかと考える。

2.3 電話を使ったシステム

そこで我々は「だれでも・いつでも・どこでも」使えるであろう「電話」に着目した。操作が簡単な「電話」というデバイスを使い，さらに押しボタンを使わずに全ての操作を声で行える仕組みであれば，インターネットを意識することなく情報を受け取ることができる。

電話であれば馴染みも深く，使い方は誰でも知っている。また近年は携帯電話も爆発的に普及しているので，場所にもとらわれずどこからでも使用でき，難しい操作も要らない。現在，東西 NTT の加入電話契約数は 5,074 万件（平成 13 年末），携帯電話/PHS の契約数は 7,772 万件（平成 14 年 9 月）である。平成 13 年度の世帯数がおよそ 4,706 万世帯であるから，ほぼ一家に一台の割合で電話が各家庭にあるといえる。

使い慣れた既存のデバイスを活用することで，インターネットを全く意識せずに自然に情報に触れることができる。情報バリアフリー実現には「電話」が必要不可欠なデバイスであることは容易に想像できる。

しかし実際の構築にあたって，いくつかの大きな課題があった。

3. 構築にあたっての課題

3. 1 従来システムとの比較

現在、電話を使ったシステムは多々あるが、その多くはアナウンスに従って欲しい情報に該当する番号を押しボタンで押下するシステムである。しかしながら、この押しボタンを使ったシステムでは、“ボタンを押す”という煩わしい操作を伴うため使い勝手が悪い。今回の構築にあたり筆者も様々な押しボタンを使うシステムを利用して見たが、どれもイライラ感が拭えないものであった。また利用者の発話内容を認識するシステムも同様に試してみたが、操作性は良いが認識率が悪かったり途中で何をしているのかわからなくなったり（迷子）と、あまり利便性が良いと思えるものには出会えなかった。

欲しい情報に辿り着くまでの操作性と的確性を両立させるにはどうすればいいのかが最初の課題となった。

3. 1. 1 だれが話しても大丈夫なシステム

本システムは「コンピュータシステムを意識させず自然な会話を可能にする誰でも使える仕組み」の実現が目標である。まず我々は必要とする機能を洗い出し（下記参照）、それに合う音声認識エンジンを選定することから始めた。

- ・ 認識率が高いこと
- ・ 不特定利用者、老若男女、誰でも使えること
- ・ 自然会話の中から、必要／不必要な語句をピックアップして認識できること
- ・ ガイダンスの途中でも割り込んで話せること
- ・ 日本語以外にも対応していること
- ・ 携帯電話からも使えること
- ・ 将来的な拡張性があること（VoIP、声紋認証、など）

これら要望を満たす音声認識エンジンを数社の中から絞込み、音声認識エンジンで最も優れている米国 Nuance Communication inc. の Nuance 音声認識エンジンを採用することにした（表1参照）。世界 No. 1 導入実績、認識率 96%以上、自然会話対応、不要語排除機能、割り込み機能、世界 25ヶ国語対応可能、声紋認証など豊富な実績と拡張性に優れる音声認識エンジンである。

表1 音声認識エンジン比較表(抜粋)

音声認識エンジン	Nuance	A社製	B社製	C社製	D社製
不特定話者	○	○	○	○	○
自然会話対応	○	×	×	○	○
認識単語数	無制限	最大 50 万語	最大 20 万語	?	最大 100 万語
不要語排除	○	△	△	△	△
声紋認証	○	×	×	×	△
他国語認識	25ヶ国語	日本語	日本語	日本語	18ヶ国語
VoIP対応	○	×	×	×	×

PFU/2002年5月

3. 1. 2 迷子にならないストーリー設計

“だれでも”使える認識を実現することは Nuance で実現できる。しかし、**図5**で示すとおり Nuance はあくまでも利用者の声を認識するだけである。“使いやすさ”を実現するためには、エンジン性能を最大限に引き出すための足回り（ストーリー設計）が必要不可欠であり、音声認識システム構築における一番のポイントとなるものである。

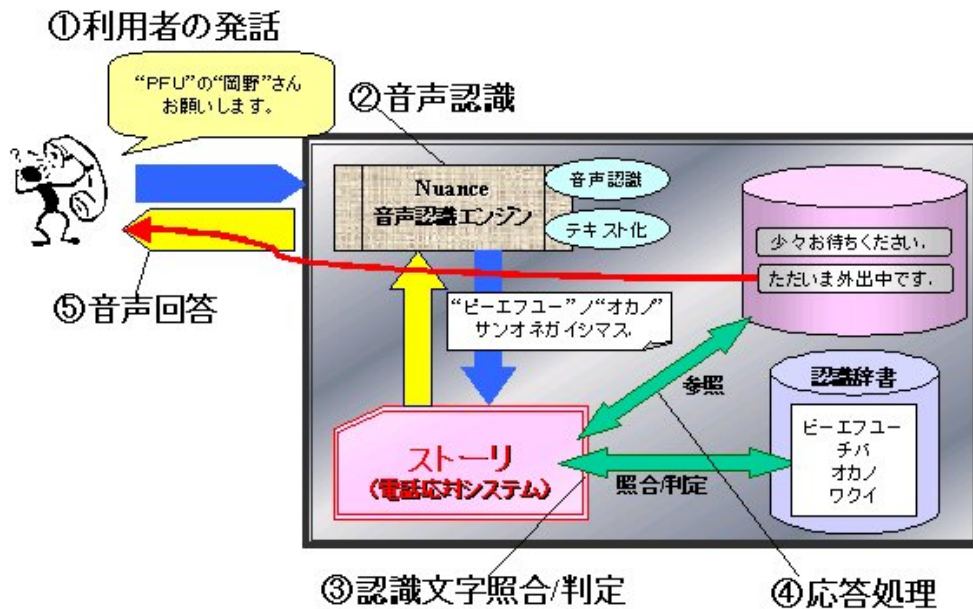


図5 音声認識の仕組み

これまでの音声認識システム（使用してみた感想）は、誤認識による度重なる同じガイダンス出力や会話のループが、“使いにくさ”の原因となっている。この課題を払拭できるかが“使える”音声認識システムを構築する上での課題である。

本システムでは従来システムの課題を、認識率だけではなく、ユーザフレンドリーなストーリー設計で克服することにした。使用中迷子にならず、安心して利用でき、また気軽に利用したいと思われる仕組みへの挑戦である。

構築に当たって注意したポイント、及びストーリー設計例（**図6**）を以下に記す。

<主なポイント>

- ・ 語句だけを認識するのではなく、会話を認識する。
 - 「〇〇」ではなく、「えーっと、〇〇お願い。」でも認識可能。
- ・ 利用者に喋って欲しい語句を、予めガイダンスとして話す。
 - ただし予想外の発話にも対応するため、予想できる語句を登録しておく。
- ・ ガイダンスの途中でも割り込んでの発話ができる。
 - ガイダンスの最後まで待つのではなく、随時移動ができる。
- ・ 発話を促しても、返事がなかった場合の対応を設定する。
 - 数秒経っても発話がない場合は再度発話を促す。
- ・ 誤認識時の再ガイダンスには同じものを流すのではなく、回数によって異なったガイダンスを流す。

- 同じところをループさせず、いらいらさせない.
- ・ 誤認識を繰り返した場合の受け皿を設定する.
 - 延々とループさせず、最後にはオペレータ（人）へ転送する.

<ストーリー設計例>

デモストーリー例(休日夜間診療案内)

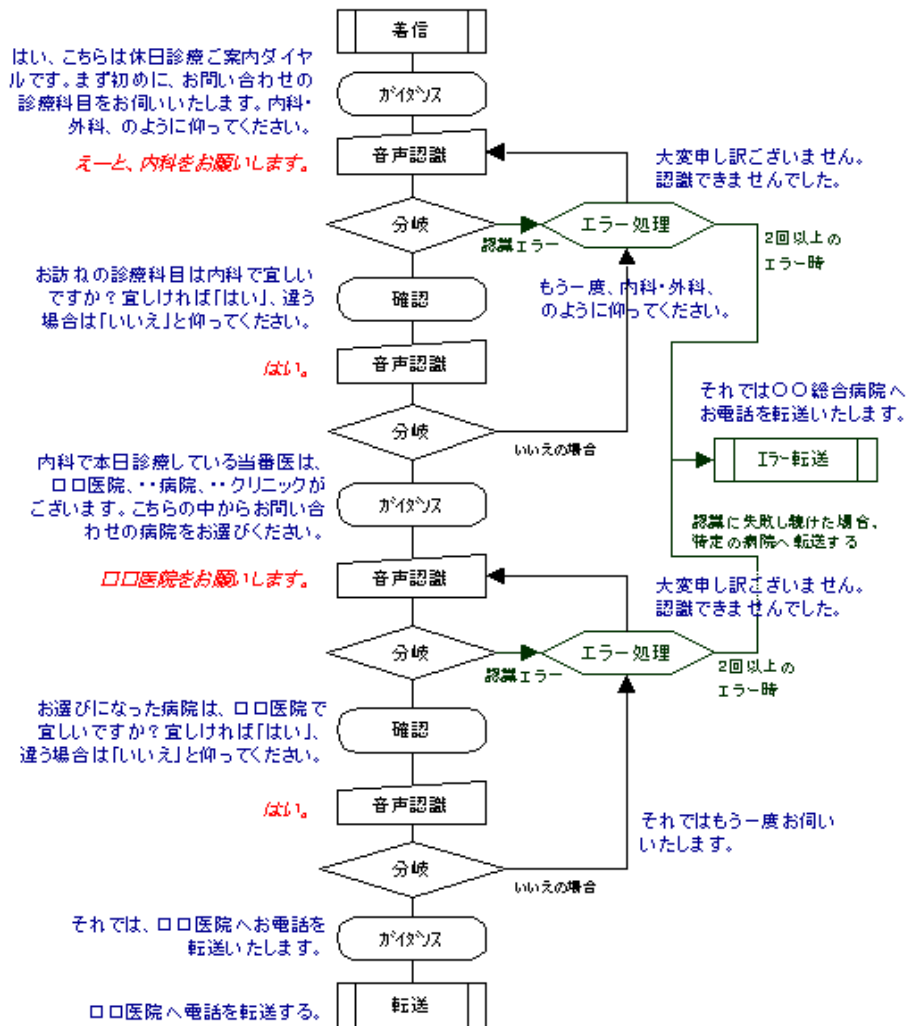


図6 ストーリー設計例

3.2 XML 技術の適応

3.2.1 電話機以外への対応

3.1 項で述べた手段で音声認識の仕組みは整ったが、電話を使っただけのシステムではバリアフリーを実現したとはいえない。本システムは、インターネット利用を不得意とする層の人達だけではなく、幅広い方々に利用していただくために、電話以外にも、パソコンや携帯電話インターネット、FAX や L-mode などの多種デバイスに対応させる必要がある。

そこで、“ワンソース・マルチユース”が可能なデータベース“XML”を採用することにした。XMLの採用決定には確固たる理由がある。

当社ではXML技術を使ったマルチデバイス対応システム導入事例がある。石川県金沢市に導入した子育て総合コミュニティサイト「エンゼルネット」（参照：<http://kodomo.city.kanazawa.ishikawa.jp/angel/servlet/WhatsNewTop>）である。これは、パソコンと携帯電話（i-modeなど）に対応する情報提供システムで、2000年4月から運用されており、利用している若い世代のお母さんやお父さんから好評を頂いている。この仕組みを拡張し、FAXやL-modeなどの新デバイスに対応することから始めた。

この拡張により、利用者は多様な選択肢の中から希望する方法で情報を受け取ることが可能となる。つまり、声で検索した情報を音声情報で聞けることはもちろん、FAXや電子メールによる視覚情報として手に入れることも可能となるのである。例えば行政機関：各種申請書の交付手順を知りたい場合、電話で申請書名や何々手続きに関する件と検索し、必要な書類をFAXで受け取ることができるということである。

3. 2. 2 管理業務の効率化

XML技術を使うメリットとして、データベースの一元管理が挙げられる。これまではデータに更新があった場合、パソコンのインターネット用、携帯電話のインターネット用、FAX用…と個々にデータベース更新する必要があった。本システムでは基本的に一つのデータベースを更新すれば全てに対応ができる。サービス提供者は、対応デバイス増加時、データ更新処理をそのまま利用でき、構築費用削減や管理効率化を図れるわけである。

更に、データ更新は、比較的簡単に使用できる汎用的アプリケーションで更新できることが望ましいと考え、Microsoft® Excel®を選定した。このように利用者にとって使いやすだけでなく、管理者も使いやすいシステムとなるように配慮したことも重要なポイントであるといえる。

3. 3 スモールスタート

音声認識の仕組みは、富士通株式会社を始め、株式会社NTT-DATAやオムロン株式会社が販売している。しかし、そのほとんどが音声認識エンジン（Nuanceなど）やストーリー設計を行うミドルウェア、音声認識辞書などの部品販売であり、導入する側はこれらを組み合わせてシステムを構築する必要がある。したがって、導入する側に構築技術が要求されたり、開発委託による費用増加もあり、音声認識システムはハードルが高いと思われる。

そこで当社は、このような“作ってもらう”仕組みではなく、“使ってもらう”仕組みをソリューションすることとした。要点を以下に挙げる。

- ・ 情報提供シーンに見合ったベストなストーリー設計。
- ・ コンサルティング体制。
- ・ 電話回線は1回線から導入/運用でき、拡張も可能。
- ・ ASPサービスの提供。

大規模な投資を必要とせず、まずは小規模な1回線システムから始め、状況に応じて回線やコンテンツを増強できる。我々が考案したシステムの強みである。

4. だれでもポータルによる情報提供

我々はこのような過程を経て【だれでもポータル】を構築した。音声認識技術と XML 技術を活用し「だれでも、いつでも、どこでも」電話をかければ様々な情報が取得できる（ポータル）仕組みである。下記図7は、だれでもポータルの「休日/夜間診療ご案内システム」としての運用例を挙げている。

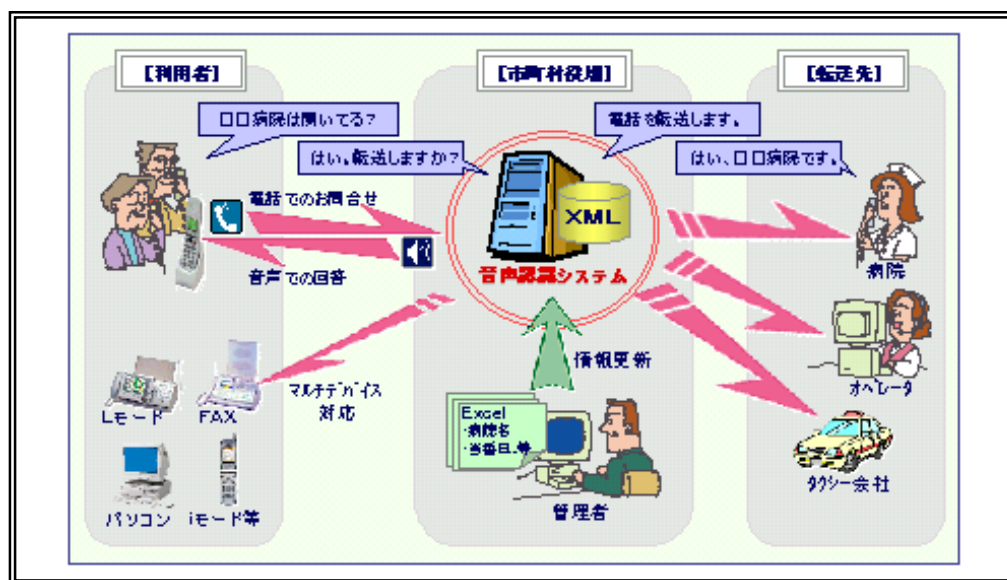


図7 「だれでもポータル」概要（例）

利用者は音声認識システムへ電話を掛け、コンピュータの応答に従って声で欲しい情報（病院の診察科目や診察日など）を指定する。コンピュータは XML データベースを検索して、該当する情報を音声や FAX などで回答する。利用者が希望する場合、またはコンピュータで解決しない場合は、最寄の施設やオペレータへ電話を転送する。

5. 評価と今後の課題

5. 1 だれでもポータル評価と苦労した点

これまで、デモシステムとして「休日/夜間診療ご案内システム」を社内外 8 展示会へ出展した。そこで、来場者から頂いた意見や評価を以下に記す。なお、対象人数は 70 人である。（重複回答あり）

表2 アンケート結果

利用者の立場での意見・評価	人数
音声認識の技術も凄いね。	63
インターネットを使わない利用者、及び、苦手なユーザでも利用が可能。	59
話すだけで操作できるので使い勝手が良い。	68
自然な発話でストレスなく操作ができる。	46
電話さえあれば 24 時間いつでも利用できるので便利。	64
欲しい情報を、欲しい時に、欲しい方法で入手できると思う。	44
PC や携帯端末と比べ、情報へのアクセスが早い。	37
まだまだコンピュータとの会話に違和感を感じる。	13

サービス提供者の立場での意見・評価	人数
インターネットを閲覧をしないユーザ、及び、苦手なユーザの取込みが可能.	48
コールセンタへ適用（人員管理やコール振り分けなど、運用の効率化）が出来る.	64
音声認識システムにより、サービスの24時間化が可能となる.	56
人件費削減によるコスト効果が期待できる.	35
デジタルデバイドの解消に繋がる.	12
クレーム対応には不向き.	4

実際に来場者に体験していただいた上でのアンケート結果から、だれでもポータルの実用性を確認できたことは非常に大きな収穫であったが、好評価があった反面、使用する人や場所により誤認識や認識率の低下が発生する場面が多々あり、その都度ノイズ対策強化やストーリー設計の見直し、認識語句の追加などのブラッシュアップ作業が一番苦心した点でもあった。

このように、導入して終わりではなく、導入後のサポートを強化しより“使いやすい”システムを顧客と共に構築していくことが、音声認識システムには必要不可欠であると考える。

5.2 だれでもポータルの提案・展開

だれでもポータルは一般的に使われている情報公開手段に、音声認識という方法でデバイス（電話）を追加したものである。弊社はそのさまざまな利用用途を図8のように考える。そのほとんどは、私たち一般市民が普段の生活の中で「こんな仕組みがあれば便利だな」と望んでいるものではないだろうか。

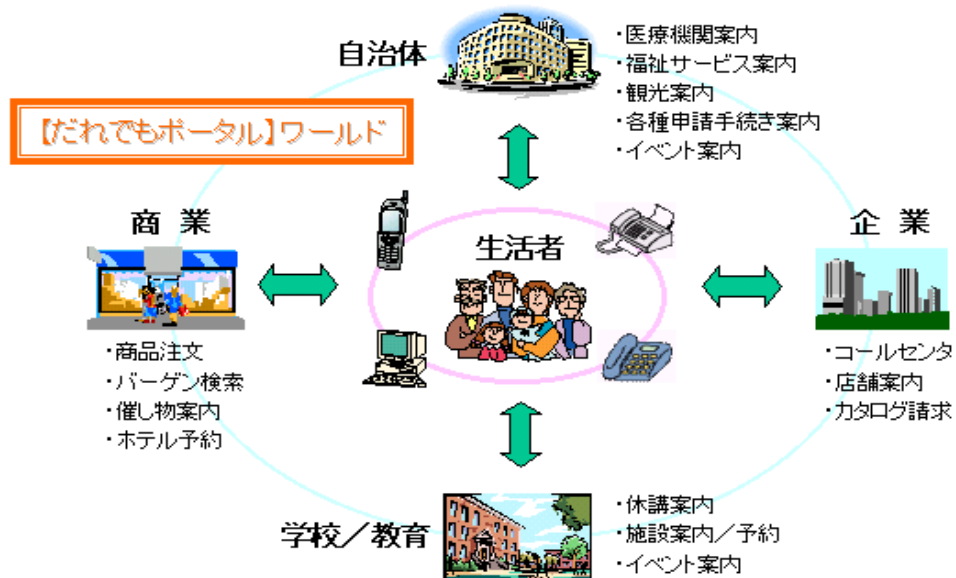


図8 想定利用シーン

今後はさまざまな利用シーンへシステム提案を行っていく中で、いかに利用者にやさしく使いやすいシステムを作るかが最大の課題であり、当社が注力するところである。だれでもポータルがデジタルデバイドを解消し、情報バリアフリー社会実現の礎となることを望む。

6. おわりに

近年音声分野の技術は日々進歩している。音声認識技術は、より自然な会話の認識を可能とし、音声合成技術はより人間らしい発声へ近づき、また、声で発話者を特定できる声紋認証技術も実用レベルにまで到達している。

現状の音声認識システムは、スイッチング（切り分け転送）や Q&A のような応答システムが主であるが、究極の利用目的はコマースとされている。声紋認証技術で利用者を特定し、その利用者固有の情報を即座に提供（インターネットの“マイページ”機能）、更には、商品の売買をすべて電話（声）で行えるような仕組みが今後主流になると予測する。

当社も今回の電話を使った情報バリアフリーシステムの構築を第一段階とし、次の段階では RosettaNet や Ariba などの電子商取引（GtoG, inG, GtoB, BtoB, inB, BtoC）との連携強化によるコマースシステムに取り組む予定である。これら電子商取引には XML 技術が必要不可欠なものとされており、だれでもポータルがデータベースに XML を採用している理由はここにも隠されているのである。

だれでもポータルが、情報バリアフリー社会を実現できる確かなものとなるよう、当社は今後もだれでもポータルを改良/改版していく。

参考文献

- [1] 総務省：「情報通信白書平成14年版」2002年7月2日
- [2] 総務省：「通信利用動向調査」2002年5月21日
- [3] ビデオリサーチネットコム：「インターネット普及状況調査結果」2002年6月26日

※Microsoft・Excel は米国，Microsoft Corporation の米国およびその他の国における登録商品です。

その他，記載されている会社名，製品名は，各社の商標および登録商標です。