

LK-012

広報へのユニバーサル・ユビキタスアクセスの実現 Universal and Ubiquitous Access to public information

立花 健二*
Kenji Tachibana

荒木 雅弘*
Masahiro Araki

概要

地方自治体等から発信される広報を代表とする公共性の高い情報を対象として、視覚障害者・高齢者などの情報弱者や、様々な情報端末への対応を配慮したアクセス手法の開発を行った。XML で記述された広報文書から、携帯端末向けの cHTML、音声のみでアクセス可能な VoiceXML、ユーザフレンドリなインタラクションを可能にするエージェント+HTML のそれぞれの記述言語に自動変換を行う手法、およびユーザの嗜好を学習し、記事の伝達量を調整する手法について説明する。

1. はじめに

近年、地方自治体や公共性の高い団体からの情報発信が Web で行われるようになり、これらの公共性の高い情報へのユニバーサルアクセスが求められている。また、様々な通信機器の発達により多様な環境 (ユビキタス) でこのような情報にアクセスしたいという要求も高まっている。

企業や自治体からの情報発信が Web で行われるようになり、それとともなってアクセシビリティガイドラインがいくつか作成されてきた。これを受け 2004 年 6 月に、高齢者・障害者等配慮設計指針 - 情報通信における機器、ソフトウェア及びサービス - 第 3 部：ウェブコンテンツを定めた JISX8341-3 が制定された。しかし、これは個々のノウハウに関する部分が多く、様々な配信形態に配慮したものではない。

このような、ユニバーサル・ユビキタスな情報取得要求に応えるためには、パーソナルコンピュータ、携帯電話、PDA など多様なデバイスに対応した情報配信が必要になる。また、視覚障害者を対象として音声による情報配信も求められる。本研究では、XML で記述された広報用コンテンツを様々な発信形態に自動変換する手法の開発を目的とする。具体的には、Web ページ用には HTML、携帯電話用には cHTML、音声対話用には VoiceXML にそれぞれ変換を行う。

しかし、小さい画面の携帯端末や音声のみによる配信は、情報の一覧性を損なうという問題点がある。そこで、ユーザの嗜好を学習し、その結果に応じて配信情報の分量を調整する技術の開発も行う。ここでは、ユーザ情報を記録する主体のメタファとしてエージェントを用い、よりユーザフレンドリなインタフェースの実現を目指す。

以下、本稿では 2 章で文書構造化の手法を説明し、3 章で多彩な配信用言語への変換について述べ、4 章で記事フィルタリングに関する説明と、実験結果について述べる。そして 5 章で考察と今後の課題を述べる。

2. 広報文書の構造化

ユニバーサルアクセスを実現するためには、多種多様な方法で情報にアクセスできる必要がある。これを可能とするために、まず広報文書を機械処理のしやすい XML へと自動的に変換することを目的とした。

今回はテキスト形式で配布されている広報文書「京都市民しんぶん」を対象とした。「京都市民しんぶん」は他の多くの広報文書と同じく、文書の内容を記事単位に分割することができる。各記事はさらに、タイトルを表す部分、記事の内容を表す部分、記事に付随した問合せなどの情報を表す部分、に分けることができる。本研究では、これに朗読に適さない非可読部分 (例えばクロスワードパズルや表など) を加え、タイトル部、内容部、付随情報部と、非可読部の四つに分類、図 1 のように構造化文書へと自動変換するプログラムを作成した。

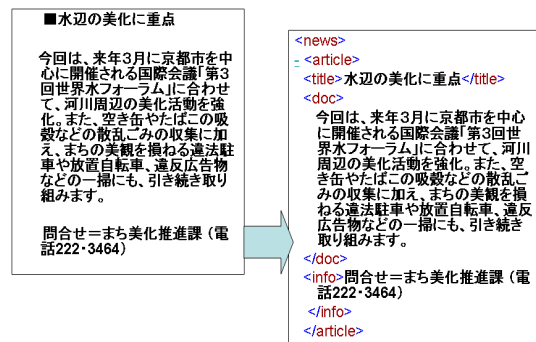


図 1: XML 変換例

例えば「京都市民しんぶん」では、例外も多いが、

- 付随情報部ではその先頭行の先頭部に「問合せ」「申込み」といった文字列が付与されている。
- 内容部には各行先頭に全角スペースか半角スペース、あるいは箇条書きを表す記号が付与されている。
- タイトル部では行頭に何も付与されない。
- 非可読部は記号や空白文字が非常に多い。

といったものが挙げられる。今回はこれらの文書特有の情報を元に図 1 のように構造化を行った。

作成したプログラムによる「京都市民しんぶん」の構造化文書変換精度を調べた。実験の対象は、規則を作成したものと別の号からの全 78 記事であった。構造化によって分類される前の各記事単位で、構造化後の記事を

優：記事が正しく区切られており、「タイトル」、「内容」、「付随情報」、「非可読」部のそれぞれが適切にマークアップされている。

*京都工芸繊維大学, Kyoto Institute of Technology

表 1: 構造化精度

| 分類結果 | 記事数 | 比率 |
|------|-----|-------|
| 優 | 36 | 46.2% |
| 良 | 20 | 25.6% |
| 可 | 1 | 1.3% |
| 不可 | 21 | 26.9% |

良：記事は正しく区切られているが、「タイトル」、「内容」等のマークアップが正確に行なえていない。

可：記事が正しく区切られていないが、その周辺記事の分類失敗のためである。

不可：記事が正しく区切られていない。

の4種類に分類、集計したところ、表1のような結果が得られた。

このうち、良の結果であった記事はほとんど全てが手作業でも分類の難しいものであり、この程度の誤りであれば、その他の作業には支障はきたさない。すなわち良の結果のものも構造化に成功していると考えることができ、71.8%の記事が構造化に成功したといえる。

なお実際の使用時には、予め著者が十分に構造化された文書を執筆し、それを構造化するか、自動的に構造化したものをボランティア作業などにより、手動で修正するか、の作業が望まれる。これらの処理には人間の手作業が必要であるが、現在のボランティアによる朗読作業に比べれば、人間にかかる負担は少なくなっている。

3. 多彩なモードでの配信

多様な方法で広報文書にアクセスできれば、結果として様々なユーザが利用でき、ユニバーサルアクセスへとつながる。ここでは前章で生成したXML文書(図1)を多彩な配信用言語へと変換することで、実際に携帯情報端末、電話、エージェントを用いてアクセスする手法について説明する。

システムの概要を図2に示す。各JSPは朗読対象の記事番号を保持し、XMLから対象記事内容を取得、これを様々な配信用言語へと変換、読み飛ばしなどの機能を付与することで、多彩な手法によるアクセスを実現する。

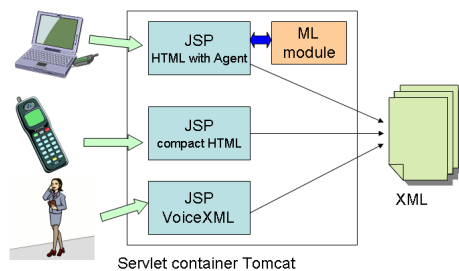


図 2: システム概要

3.1 携帯情報端末からのアクセス

携帯情報端末からのアクセスを可能とすることで、ユーザはどこからでも情報を引き出すことができるようになる。しかし現在、携帯電話に代表される携帯情報端末は、通信速度が低速である、メモリ容量が少量である、といった問題を持ち、単純に文書の内容全てを携帯情報端末により受信することは現実的ではない。そこで今回は構造化された記事から、対象記事のタイトルを抽出し、携帯情報端末用のHTMLであるcHTMLへと変換するシステムをJSPにより実装した。

本システムでは、対象となる記事内容をXML文書から抽出し、その前後の記事へのリンクを付与してcHTMLへと変換する。これによりユーザは好む記事に対してリンクからアクセスすることができる。実装したシステムの出力を図3に示す。

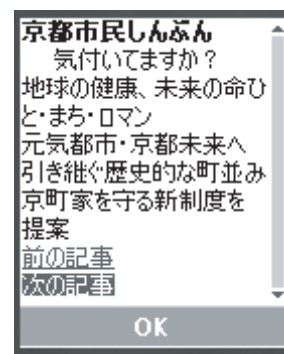


図 3: 携帯情報端末からのアクセス

3.2 電話からのアクセス

電話からのアクセスにより、ユーザは視覚情報に頼らず情報を得ることができ、視覚障害者にも利用可能なシステムとなる。そこで音声対話記述言語VoiceXML[1]を用いて、電話からアクセス可能なシステムを実現した。具体的にはJSPを用い、XML文書からタイトル部、内容部、付随情報部を抽出し、VoiceXMLへと変換した。

携帯情報端末と同様に、ここでも電話からのアクセス専用のJSPを作成した。このJSPの出力は図4のようなVoiceXMLであり、この場合デフォルトではタイトル部分のみが朗読されるが、後述する記事フィルタリングの結果に応じて、12行目のコメントアウトをはずすことで、内容部、付随情報部の朗読も可能である。

VoiceXMLを利用すれば記事の朗読だけでなく、読み飛ばしなどの朗読制御も可能であり、今回はユーザの「次へ」という発話により、現在朗読中の記事を読み飛ばすことができ、「詳しく」という発話により、朗読を省略された部分も朗読できるように実装した。これらにより、ユーザは対象記事の朗読を聞き、興味がなければ読み飛ばしを行え、興味があればより詳しい情報を取得できる、健常者が実際に広報を読む場合に近いシステムを実現した。

3.3 エージェントによるアクセス

ここでいうエージェントは、ユーザと計算機とのコミュニケーションを支援するインタフェースエージェントの

```

<?xml version="2.0">
<link next="#document">
  <grammar> kuwashiku (in detail) </grammar>
</link>
<link next="http://localhost:8080/vxml.jsp">
  <grammar> tsugi (next) </grammar>
</link>

<form name="title">
  <block>
    Title element
  </block>
  <!-- <goto next="#document"/> -->
  <goto next="http://localhost:8080/vxml.jsp"/>
</form>
<form name="document">
  <block>
    Document element
    Information element
  </block>
  <goto next="http://localhost:8080/vxml.jsp"/>
</form>
</vxml>

```

図 4: VoiceXML テンプレート

ことである。文書の内容を表す Web ページ上にエージェントが現れ、それにより記事が朗読されると、ユーザは視覚と聴覚の両方の情報を利用することができるようになる。これにより視覚障害者だけでなく、健常者にとってもわかりやすく、また健常者には記事の朗読を聞きながら、それとは全く関係のない作業を行えるなど、様々な人に使いやすいシステムとなる。エージェントと HTML による表示が、テレビニュースのメタファとなり、様々な人に扱いやすくなることが期待できる。

本研究では、Microsoft Agent が記事を朗読するシステムを実装した。具体的には、まず JSP により対象記事の内容を XML 文書から抽出、抽出された情報はタイトル部、内容部、などの分類により、それぞれ HTML の適切な部位に埋め込まれ、様々な Web ページデザインを用いて表示される。また後述する記事フィルタリングの結果に応じて、記事の適切な部分を朗読する Microsoft Agent を VBScript により出力する。実装したシステムの出力を図 5 に示す。

このシステムでは VBScript により、「次へ」という発話かボタン入力による読み飛ばしや、「詳しく」という発話またはボタン入力により、フィルタリングの結果、朗読が省略された部分の朗読指示を実現している。

4. 記事フィルタリング

4.1 フィルタリング手法

合成音声により単純な全文朗読を行えば、ユーザは多くの時間を浪費することとなり、ユーザの不快感へも繋がる恐れがある。とはいえ、タイトルだけを朗読するようなシステムでは、全文を朗読させたいユーザにとっては、記事内容を知ることができないか、さもなくば記事ごとに朗読を指示することとなり使いにくい。

そこで本研究では、ユーザの嗜好をシステムが判断し、ユーザが読みたいと判断することが予測される記事はあらかじめ全て読むように、読みたくないと判断することが予測される記事はタイトルのみ、あるいは付随情報以



図 5: エージェントからのアクセス

外を読むように、システム側であらかじめ設定し、ユーザに記事を提供するシステムを実装することとした。

人間による朗読を模倣するならば、システムは何らかのユーザモデルを想定し、ユーザの嗜好に応じた朗読を行うべきであろう。このようなユーザモデルを用いた研究として [2] があるが、この研究で必要とされる事前のアンケートは、広報文書の性質を考えると実用的ではないため、本研究ではユーザが使用することにより、ユーザの嗜好を機械学習により学習することができるシステムを考えることとした。

そのため、読み飛ばしや、詳しく読むといったユーザの反応から、ユーザの嗜好を学習することで、

1. タイトル部のみ
2. 付随情報部以外全て
3. 記事全て

の三つから適切な朗読量を選択し、出力するシステムを考える。

今回対象とする広報文書「京都市民しんぶん」の特性として、前月分など、過去の記事読み飛ばし情報を得ることは難しい、記事は比較的少数である、などが挙げられ、これらの特性を考えると、この学習には、

- テキストマイニングの能力が優れていること
- 少数のデータから利用可能であること
- 高速な処理が可能であること

の三つが求められる。本研究ではこれらを満たす学習アルゴリズムとして、機械学習 RIPPER [3] を選んだ。

RIPPER は文書分類が可能な規則学習アルゴリズムの一種であり、分類対象となるの特徴量、訓練用データ(訓練用要素の特徴量とその分類結果)、を入力とし、分類のための規則を出力とする。RIPPER はデータマイニング能力が優れており [4]、C4.5 [5] と異なり、文書の分類に適している。

今回は RIPPER を利用し、ユーザの行動から、リアルタイムで記事のフィルタリングを行う以下のような手法を考案し、実現した。このシステムでは

1. タイトル部のみ
2. タイトル部と内容部
3. 全ての部分

のいずれの朗読を行うべきかを出力する。

RIPPER への入力としては記事の特徴量を用いた。ここでは特徴量は各単語を次元とし、単語の $tf \cdot idf$ 値を各次元での値としたものを用いた。

実際には次のように学習を行なった。

1. 最初の記事がフィルタリングの初期値に応じて朗読される。
2. 読み飛ばし情報を次のように分類、取得。
 - (a) タイトル部のみが読まれた。
 - (b) 内容部まで朗読された。
 - (c) 付随情報部まで、または記事全てが朗読された。
3. 読み飛ばし情報を入力として、RIPPER によりフィルタリング用の規則を学習する。
4. 次の記事を対象とし、フィルタリングを行い、2へ。

フィルタリング結果はユーザの情報としてサーバ上に保持される。

4.2 記事フィルタリング実験

提案したフィルタリング情報学習手法を、エージェントとのインタラクションをシミュレートしたデータを用いて評価した。印刷した京都市民しんぶん 1号分(53記事)を5人の被験者に読ませ、被験者の嗜好に応じて、各記事に対して読み飛ばしを行ってもらい、上記 RIPPER アルゴリズム step2 の (a),(b),(c) に対応するマークを付けさせた。

このインタラクションをシミュレートしたデータを用いて、学習アルゴリズムの精度を評価した。例えば、もし10番目の記事の判定結果が(c)であり、被験者のマークが(a)であれば、ユーザからの読み飛ばしコマンドが入力されたものとして、フィルタリング規則の再学習を行い、その新しい規則を用いて11番目の記事の判定を行う。実験ではn番目の記事までで学習を行い、その直後の5記事において、学習結果によるマーク(a)~(c)がユーザの読み飛ばし位置と一致しているかどうかを評価した。

実験結果を図6に示す。横軸は学習に用いた記事数、縦軸はその時点での、被験者5人の5記事におけるテスト正解率を表す。本実験では最終的には70%以上の高い正解率が得られている。また正解率は徐々に上昇を続けており、さらに多くの記事を読むことによってさらに高い性能を示すことが期待できる。

5. おわりに

本紙では、ユニバーサル・ユビキタスなアクセスを、広報文書をXMLへと変換し、これをさらに、compact HTML、VoiceXML、エージェントを表示するHTMLに変換する、という手法により実現した。さらにユーザ

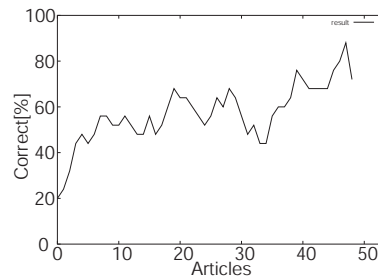


図 6: 記事フィルタリング結果

の反応から機械学習を用いてユーザの嗜好を学習し、記事フィルタリングを行うシステムも実現した。

今後の課題として、(1) ユーザの嗜好により早く適合することができる能動学習 (例えば [6]) を用いること、(2) キーワードによる検索に対応した VoiceXML の出力を実現すること、(2) 様々な広報文書に対応すること、などが挙げられる。

謝辞

本研究を支援していただいた熊本ソフトウェア株式会社 に深く感謝いたします。また、京都市民しんぶんのデータを提供して下さった京都市広報課に感謝いたします。

参考文献

- [1] McGlashan, S.et.al., "Voice Extensible Markup Language (VoiceXML) Version 2.0", W3C Recommendation, 2003. <http://www.w3.org/TR/voicexml20/>
- [2] Whittaker,S., Walker,M. and Maloor,P., "Should I Tell All?:An Experiment On Conciseness in Spoken Dialogue", In Proc. of Eurospeech2003, pp.1685-1688,2003.
- [3] Cohen, W. W., "Learning trees and rules with set-valued features", In Proc. of AAAI, pp.709-716,1996
- [4] Walker, M., Langkilde, I., Wright, J., Gorin, A. and Litman, D., "Learning to Predict Problematic Situations in a Spoken Dialogue System: Experiments with How May I Help You?", In Proc. of NAACL, pp.210-217,2000.
- [5] Quinlan,J. R., "C4.5: programs for machine learning", Morgan Kaufman, San Francisco, CA, USA, 1993.
- [6] Tur, G., Schapire, R. E. and Hakkani-Tur, D., "Active Learning for Spoken Language Understanding", In Proc. of ICASSP, 2003.