

FAXを用いたWWW上のデータベースの構築と利用

瀬川 典久
対馬 伸行
宮崎 正俊

岩手県立大学 ソフトウェア情報学部
〒020-0193 岩手県滝沢村滝沢字巢子152-52
E-mail: sega@acm.org

概要

コンピュータを使って作られた情報が、一方ではコンピュータの中で利用され、もう一方では現実社会に存在する物に出力されることによって利用される。利用形態が異なる同じ情報を、異なる利用形態の間で変換を行う。これを、情報の再利用と考える。

本論文では、この考え方の具体例として、データベースシステムを2つの手法(WWWとFAX(紙))を利用したシステムについて述べる。

キーワード：ペーパーインタフェース、WWW、データベース、Recycle Media

Paper interface by fax for a database on WWW

Norihisa Segawa, Nobuyuki Tsushima, Masatoshi Miyazaki

Faculty of Software and Information Science, Iwate Prefectural University
152-52, Sugo, Takizawa, Iwate, Japan 020-0193
E-mail: sega@acm.org

Abstract:

An information which is constructed by computers is used by themselves. The other translates into objects in real world. The same information which is used in different forms is translated between one form to another. We consider this translation as 'recycle information.'

This paper reports on our trial of the development of a paper interface by fax for a database on the World-Wide Web (WWW). The interface can treat database queries. User can choose between WWW interface or the paper interface by fax.

keyword: paper interface, WWW, database, Recycle Media

1 はじめに

現在の情報社会では、事務作業において何らかの形で電子的な方法がとられている。例えば、文章作成ではワープロ（と言う電子的なアプリケーションソフトウェア）を利用し、作成した文書の配布には、WWWやメール（と言うメディア）等の電子的な方法や、紙（と言うメディア）に印刷し配布する非電子的な方法を用いたりしている。このように、現在の事務作業は、電子的な方法と非電子的方法を人間が相互に使い分けながら利用している。

しかしながら、それら2つの方法（電子的な方法と非電子的な方法）の間で相互に情報を変換し、利用することはできない。例えば、計算機の情報を紙に出力した後、その紙にペンで情報を書き加えても、紙上で行った処理（訂正の二重線を引く、赤ペンで計算式を書き加える、など）と計算機上の処理（文字の追加、削除、処理の変更等）を関連付けて処理できない。よって、それぞれで行われた処理は、ユーザの手で変換を行わなければならない。

本研究テーマは、コンピュータ利用における2つの方法やコンピュータで利用するアプリケーションソフトウェアをメディアとして一般化し、それらで作成した電子的な情報や電子的な配布方法を関連付けて相互に変換できる環境（シームレスな情報環境）の構築を目指している（図1）[1][2]。この環境の確立により、各ユーザにとってその利用する状況で一番適したメディアを利用する

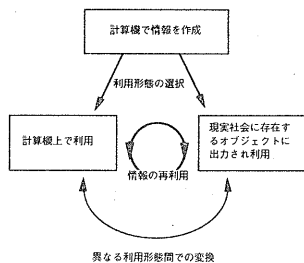


図1 異なる形態間の情報の変換

事が可能となり、従来のメディアにとは、比較できないほどの幅広くかつ様々な場面で利用されるメディアの構築が可能となる。

本論文では、この研究テーマの具体例として、FAXを用いたWorld Wide Web(WWW)上のデータベースの構築とその利用について述べる。現在、WWWクライアントを利用してデータベースを構築、検索するシステムが幅広く利用されている。本システムでは、一般ユーザが、紙を用いた手書き情報を利用しFAXによる送受信を行うことで、WWWクライアントを用いたときと同じ動作が可能になっている。この事は、先ほど述べた、電子的な手法と非電子的な手法をシームレスにつなぐ具体的な例としてあげることができる。

以下、本システムの考え方、システム構成、システム動作、考察について述べる。

2. システムの目的と意義

近年のネットワーク環境の向上により、WWWを用いたさまざまなシステムが構築されている。その中で、データベースとWWWを連携させて利用するシステムは、電子商取引等幅広い分野で利用されている。

WWWとデータベースを組み合わせるシステムは、いろいろな手法が提案されている。図2は、WWWとデータベースを組み合わせるときの一般的な一例である。WWWクライアントは、WWWサーバにHTTPを用いデータベースへ要求（データ登録・削除など）をだす。WWWサーバ

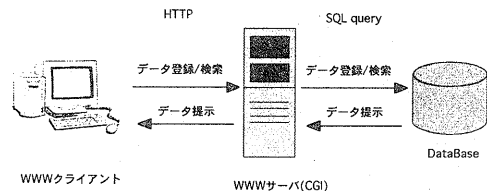


図2 データベースとWWWのCGIを組み合わせたシステム

バは、クライアントの要求を理解し、データベースに対し要求に従ったSQL queryをだす。データベースは、WWWサーバからのqueryに対して、適切な結果を、WWWサーバに返す。最終的に、WWWサーバは、データベースからの結果をHTMLをもちいてWWWクライアントに渡す。この一連の流れで、ユーザは、WWWクライアントを用いてデータベースにアクセスすることが可能となる。

WWWとデータベースの連携システムの具体例として、岩手県再利用事業協同組合が構築した建設リサイクルの広場（リサイクル情報システム）(<http://www.sairiyou.or.jp/>)をあげる（図3）。建設現場から出た廃材などのリサイクル利用を行うために、組合参加者間で情報を交換するために作られたシステムである。図3のページを利用し、各組合参加者は、建設現場で発生した廃材の情報を入力し、その情報はデータベースに蓄積され、各利用者によって検索されることによって情報が交換される。

しかし、このシステムは現在まで頻繁に利用されていない（2000年6月末で、6件しかデータが入力されていない）。それは、建設現場という特殊性に起因する。建設現場では、(1)WWWクライアントを導入する事がなかなか難しい（回線、コン

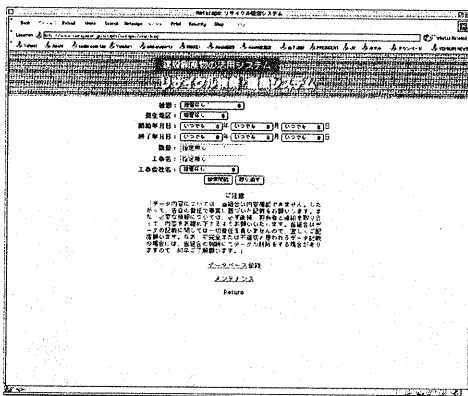


図3 リサイクル情報システム

ピュータの現場への導入まで手が回らない）、(2)WWWを利用できる人が少ない等が現状である。

そこで、この状況を変えるために、紙に情報を手書きで書き込みFAXを利用して、データベースをアクセスするシステムを考えた（図4）。このシステムは、WWWクライアントの代わりにFAXを利用し、WWWクライアントからWWWサーバへの情報は、手書きのイメージによって、WWWサーバからWWWクライアントへの情報は、コンピュータ出力のイメージとして扱う手法である。これは、(1)建設現場にはFAXが導入されることが多い、(2)手書きのインターフェースなら使い方を理解するのが容易であると考えたからである。また、既存のWWWクライアントのシステムも利用できるので、利用者自身がシステムを利用するインターフェースを選択できる。

このシステムが実現することにより、データベースというメディアが、2種類以上の利用方法（WWWを利用する手法と紙を利用する手法）によって利用され、また、それぞれの利用によってもたらされた結果は、それぞれの利用方法でシームレスに利用できる。また、メディアをいろいろな交換することによって、いろいろな応用を考えることが可能である。

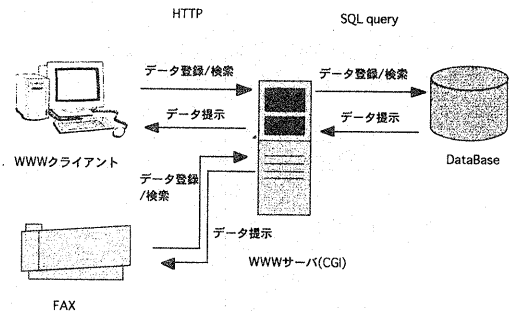


図4 FAXとWWWを利用したデータベースシステム

3. システムの構成と構築

本システムは、インターネット上に存在するデータベースにWWWとFAXを利用して、情報を登録、検索するシステムである。この研究で実装されるのは、FAXを利用して、情報を登録、検索するシステムである(図5)。

利用者は、紙にデータベースへの要求(情報の登録、検索)を書き込み、それをFAXを用いてシステムに送ることによって利用する。システムは、利用者から受け取ったFAXの画像情報を処理し、データベースに対して適切な処理を行う。また、データベースから受け取った情報を、FAXの画像情報に変換し利用者へ送信する。この事によって、利用者はWWWを用いなくても、FAXだけを利用してWWWを利用したデータベースが利用可能となる。

このシステムは、具体的には、次の4つの機能からなる。

(1) FAX画像処理

利用者から送られてくるFAXの画像情報を処理する機能である。一般に、FAXの画像情報は、G3規格と呼ばれるFAX専用のフォーマットで処理されているので、これをPCで利用できるTIFFフォーマットに変換する。

リサイクル検索・登録システム
http://www.ginga.or.jp/cg-bin/kumiai-1/saizyousearch.cgi

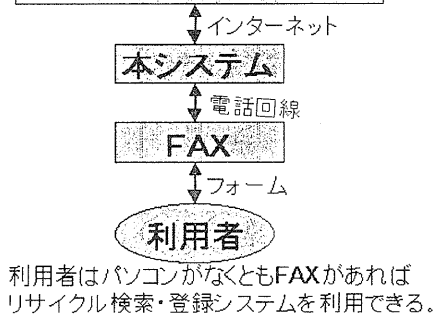


図5 本システムの概要

(2) 文字・図形認識

文字認識機能を利用して、TIFFフォーマットに変換された画像情報から文字情報を等を抽出する。利用者は、データベースへの処理内容を手書きで記述しているため、それらの情報をこの機能によって取り出す。取り出す情報は、(1)データベースへの処理内容(文字、選択記号等)と(2)識別マーク(データベースへのどのような処理化を識別)の2種類である。

(3) WWWサーバへの処理

(2)の処理によって取り出された、データベースへの処理内容をWWWサーバに渡し、WWWサーバによって処理を行わせる。直接データベースに対してアクセスを行ってもよいが、WWWを利用したシステムとの利用方法との一貫性を持たせるためにこの手法を取っている。

また、WWWサーバから見た場合、WWWクライアントとして、FAXを利用したシステムが追加されただけと認識される。よって、WWWサーバとデータベースの間に新たなシステムを構築する必要がないのもメリットの一つである。

(4) 画像出力

データベースからの出力を、利用者に対してFAXを用いて送信する必要がある場合、データベースの出力をFAXの画像情報に変換する必要がある。

これらの流れを、図6にまとめる。

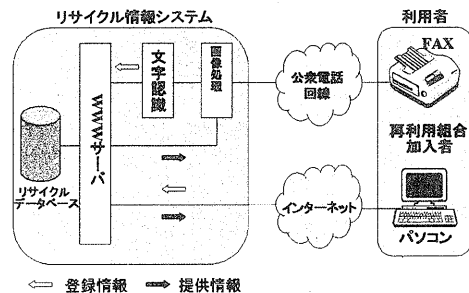


図6 本システムの構成

4つの機能を次の手法を用いて実装を行った。

(1)FAX画像処理

(株) インターコムのみと一くFAX2001を利用して実現した。みと一くFAX2001の、(1)画像変換機能(2)プログラム起動機能を利用している。

(2)文字・図形認識

(株) 富士通のDynaEyeライブラリ[3]を利用して、文字・図形認識プログラムを構築した。DynaEyeライブラリは、画像情報から手書きの文字、図形を取り出しOCR処理を行う機能をもつVisual Basicライブラリである。この、ライブラリを利用しVisual Basicによって実現した。

(3)WWWサーバへの処理

(2)によって取り出された情報を、WWWサーバに渡すための通信プログラムを作成した。これは、一種のWWWクライアントとして動作している。

WWWサーバは、本来なら図3のようなWWWクライアントからの情報を、HTTPを利用して受け取る。この、WWWクライアントと同じ動作をするプログラムを作成し、なおかつ(2)によってFAXから得られた情報をWWWサーバに伝える動作をするように作成した。

WWWサーバからは、ただの一つのWWWクライアントとして接続要求を受け付け、処理を行う。

(4)画像出力

データベースからの結果を、FAXの画像情報に変換するプログラムを作成した。仮想WWWクライアントを作成し、そのクライアントがWWWサーバに通信をし、データベースからの結果を受け取る。その際に、WWWクライアントとして実行されたときの結果の画面のイメージをFAXの画像情報に変換する。その変換した画像情報をみと一くFAX2001に出力し、最終的に利用者にFAXを利用してデータベース結果が渡される。

4. システムの利用

この章では、具体的なシステムの動作について説明する。

4.1 データベースにデータの登録

まず、利用者が図7の用紙に手書きで情報を記入する。この用紙には、このデータベースで利用される情報（この実装では、廃材についての情報）を記入する。

利用者は、記入した用紙をFAXを用いて送信する。システムは、FAX画像処理を行いTIFFフォーマットのデータが出力される。

文字・図形認識部は、画像処理によって得られたTIFFフォーマットの画像データから、次のデータを取り出す。

(1)紙の隅に書かれている識別マーク

識別マークには、(1)紙の向きの情報(2)印刷の位置合わせの情報(3)この用紙がデータベースに対して行う処理の情報が含まれている。この用紙の隅のマークを読み取ることによって、この用紙がデータベースへのデータの登録を行う処理を行う用紙ということがシステムで識別できる。

(2)文字・マークの取り出し

識別マークを取り出した後、その紙の処理内容にしたがって、文字・マークを取り出す。DynaEyeのOCR技術によって処理され、テキストデータとして取り出される。

図7 データベースデータ登録用用紙

取り出されたテキストデータは、WWWサーバに渡され、データベースに登録される。

4.2 データベースを利用した検索

利用者が、検索を行うための用紙に情報を書き込みFAXを送ると、登録時と同様に、システムによって処理されデータベースに処理が渡される。データベースからの結果は、画像出力部に渡され、FAXの画像情報となり利用者へ送信される(図8)。

5 考察

現在このシステムは、岩手県再利用事業協同組合で運用を行っている。運用した際に気づいた問題点を次にあげる。

5.1 文字認識の問題

紙から情報を取り出すときに、DynaEyeのOCR技術を利用している。OCR技術は、近年進歩したが100%確実に文字が読み取れるわけではない。今回のシステムでも、文字の読み間違いが発生した。

この問題について、次の2つを行うことによって対処した。

(a)OCRのための特殊辞書の作成

データベースの対象が決まると、そこでよく使われる単語の集合を決めることができる。例えば、今回のリサイクル情報システムでは、廃材名(コンクリート・木材)・氏名(組合参加者の集合)等である。その情報をOCRに入れておくと、

認識率が上がる。手書きシステムの場合は、この辞書の作成がポイントであることがわかり、現在も辞書の作成のノウハウを収集しているところである。

(b)FAX画像のデータベース管理者への送付

(a)の辞書作成を行ったとしても、100%の確率で文字認識を行うことは不可能である。そこで、このシステムでは、利用者が送付したFAX画像をデータベース管理者にメールで送付する付加機能を設けている。この機能によって、もし文字認識がうまく働かず、データベースに処理が渡らない場合には、データベース管理者がFAX画像を利用することによって、処理を行うことが可能となっている。

6 まとめ

本論文では、シームレスな情報環境の具体例として、FAXを用いたWorld Wide Web(WWW)上のデータベースの構築とその利用について述べた。

今後は、(1)様々なデータベースへの応用(2)専門用語の開発(3)新しいインタラクションの枠組みとそのモデル化について行なっていきたい。

追記

この研究は、岩手県中小企業団体中央会からの補助金を利用した、岩手県再利用事業協同組合からの委託研究の一部である。

参考文献

- [1]瀬川典久 佐藤究 布川博士 宮崎正俊:Recycled Media-概念と可能性-,情報処理学会シンポジウム「インタラクション'98」,pp137-144 (1998)
- [2]山下大輔、高木啓伸、萩谷昌己:紙に埋め込まれたプログラムによる紙インタフェース,情報処理学会研究報告,99-HI-83,pp 25-30(1999)
- [3]DynaEye:<http://www.fujitsu.co.jp/hypertext/ocr/dyna-feature.html>

HTTP://0.200.0K Server: Netscape-Enterprise/2.04 Date: Mon, 22 Sep 1999 15:05:02 GMT Content-type: text/html

建設産物の活用システム

・リサイクル情報検索システム

検索結果

種類	名称	所在地	開始年月日	終了年月日	数量	工事名	会社名	詳細	備考
木材	建設行	湯沢市	1998/11/20	1999/11/20	1000kg	(仮)建設 工事		01976611 2832 2514	
木材	清沢村		1999/05/01	2001/12/01	200kg	建設産物		01976611 2288 2514	備考に収録してあり 木片、木屑
コンクリートガラ	建設行	湯沢市	1999/01/01	1999/01/01	kg		サンプル	018-2000 0000	

図8 データベースからの結果 (FAX画像)