

## WWWを利用したアンケート処理

瀬 島 紀 夫・筒 本 和 広・三 谷 康 夫

アンケート調査におけるペーパーレス化と効率的なデータ入力、自動的な集計処理と迅速な結果の公開を目的として、WWWを利用したアンケート処理システムを開発した。本システムは、第87回薬剤師国家試験問題評価のアンケート処理システムとして開発したものである。本論文では、システムの開発に必要なWWWの基本的な仕組みとWWWを利用したシステム固有のデータ処理や基本的なセキュリティ対策などについて報告する。

[キーワード：WWW, CGI, Perl, Internet, Security]

### 1. はじめに

従来アンケート調査には紙を利用したデータ収集法が行われていたが、Internet技術の普及によりWWW(World Wide Web)がしばしば利用されるようになってきた。アンケート調査にWWWを利用することは、作業の大部分をコンピュータで処理することができるため多くのメリットがある。しかし、ネットワークインフラにInternetを利用するため、セキュリティ面での大きな問題が存在する。

Internetは余りにも急速に普及したため、そのサービス提供者・利用者共にInternet技術の本質をよく理解せず、安易なサービスの提供・利用を行っているケースが見受けられる。企業などのWWWサイトで会員情報が自由に閲覧できるようになっていたという事件が今年だけでも数件マスコミで報道された。この中には、WWWサービスの基本的な仕組みを理解せずに運用しているのではないかと考えられるものもある。

このような現状を踏まえ、以下にアンケート調査にWWWを利用するメリットやシステムの設計・開発・運用、また各々の段階での基本的なセキュリティ対策などについて提案する。

### 2. データ回収と集計

Internet技術の普及によりアンケート調査のデータ回収や集計、その結果の提供にWWWを利用するケースがしばしば見受けられるようになった。これは、次のような利点があるためと考えられる。

- a) 用紙の配布・回収が不要になる。
- b) 回収データの入力作業が不要になる。
- c) 回答者に馴染みのあるユーザインタフェースが利用できる。
- d) 集計結果の公開が迅速になる。

a)の用紙の配布・回収が不要になるとは、Internet技術を利用することでアンケート用紙の配布や回収を一切せず、アンケートをペーパーレスに実施することである。その手順は、E-Mail (Electronic Mail)でアンケート案内とWWWページのURL (Uniform Resource Locator)を回答者に知らせ、アンケートに協力してもらう形で実施する。これにより、アンケート実施者は用紙の配布・回収にかかる手間と費用（印刷代、通信費等）を軽減できる。

b)の回収データの入力作業は、アンケート調査のデータ処理でもっとも負担の大きい部分である。アンケート調査においてデータ集計処理が、早い段階でコンピュータ処理されるようになった。その理由は、集計処理が最もコンピュータが得意とする部分であったことによる。しかし、このコンピュータ処理を行うためには、回収したデータをコンピュータに入力する作業が必要になる。この作業は、ほとんどの場合、人海戦術で行われ大きな負担となっている。調査規模が大きい場合は、マークシートや専用ソフトを開発して利用するケースもあるが、マークシートには記入の仕方によっては読み取れない、あるいはマークシートの取り扱いに注意が必要などの問題がある。また専用ソフトには開発費、動作保証、操作方法などの点で問題もある。開発費はアンケート内容によって大きく変動し、動作保証は回答者のコンピュータ環境の限定が不可能なため膨大な検証作業が必要になる。このためどちらも小規模な調査では採用しにくく、採用したとしても費用の増大を招く。

データ入力においてWWWを利用することは、回答者自身がデータを入力することができるため、アンケート実施者側での入力作業が不要になる。また、入力内容を回答者自身にチェックさせることができ、不正な回答データなどを排除することも容易になる。

c)のユーザインタフェースについては、b)にも一部関連している。データ入力部分に利用するWWWブラウザは、Internetを利用できるコンピュータ環境であれば必ず利用できるため、専用ソフトの配布が不要になる。また、誰もが簡単に操作できるソフトウェアであるため専用ソフトのような特別な操作マニュアルの作成も不要となる。これにより利用者教育などにかかるサポート費用を軽減できる。

d)の集計結果の公開とは、従来のアンケート調査のように集計結果を印刷物の形で行っていたものを、WWWを利用することにより任意時点での集計情報の参照・回答の訂正、集計結

果の提供などをリアルタイムに、また、低コストで行うことである。

アンケート調査のデータ回収や集計結果の提供にWWWを利用することは以上のような利点があり、有効な手段であるといえる。

### 3. システム設計

#### 3.1 システムの選定

Internet技術には様々なものがあるが、アンケート調査のデータ処理を行うにはWWWを用いたクライアント・サーバシステムで構築するのが開発・運用の面で最も効率が良い。

WWWはHTTP(Hyper Text Transfer Protocol)と言うシンプルなプロトコルを用いたクライアント・サーバシステムである。HTTPはクライアントから情報の転送を依頼するリクエストをサーバに送信し、サーバはそれに対するレスポンスをクライアントに返信するという方法で通信を行うプロトコルである。

図1にHTTPプロトコルの例を示す。これは、WWWサーバにトップページの情報を問い合わせた時のリクエストとそのレスポンスである。

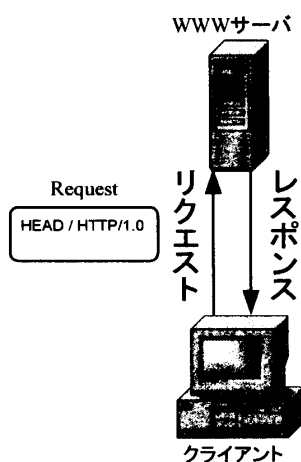


図1 HTTP プロトコル

Response

```

HTTP/1.1 200 OK
Date: Sat, 26 Oct 2002 03:33:03 GMT
Server: Apache/1.3.23 (Unix) (Vine/Linux)
Last-Modified: Sat, 06 Apr 2002 01:18:50 GMT
ETag: "fe02-883-3cae4cfa"
Accept-Ranges: bytes
Content-Length: 2179
Connection: close
Content-Type: text/html
    
```

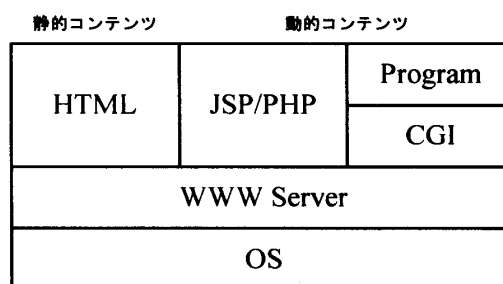


図2 コンテンツの位置付け

WWWは静的コンテンツの提供を行うために開発され、後に動的コンテンツを提供できるように拡張されている。動的コンテンツの提供機能としては、CGI(Common Gateway Interface)/JSP(Java Server Page)/PHP(PHP: Hypertext Preprocessor)などがある。このCGI/JSP/PHPは、サーバ側で処理が実行されるという特徴がある。

図2にWWWの静的コンテンツと動的コンテンツ(CGI/JSP/PHP)の位置付けを示す。大企業のWWWサイトではJSPが良く利用されるようになってきた。これはJavaのObject指向が大規模なWebアプリケーションの作成に向いているためと考えられる。PHPは個人ベースのWWWサ

イトでよく利用されるが、大規模なWebアプリケーションの作成事例はまだ見られない。

開発したアンケート調査のデータ処理システムとして、CGI + Perl (Practical Extraction and Report Language) を採用した。この組み合わせは最も古くから動的コンテンツの提供手段として利用されてきたもので開発に必要な情報量、動作実績も多い。

CGIはWWWサーバとプログラムとの間のゲートウェイ・インタフェース機能のことである。CGIプログラムの開発言語にはシェルスクリプト・Perl・C/C++等、標準入出力と環境変数を扱える言語であれば何を用いてもよい。今回利用したPerlは、文字列処理・字句解析処理などのテキスト処理に優れていることからCGIで最も多く利用されている。

WWWサーバにも様々なものが存在するが、WWWサーバとしてワールドワイドで60%以上のシェアを誇るOSS (Open Source Software) のApacheをその動作実績から採用した。また、OS (Operating System) もOSSのLinux、サーバ機器はIA-32 (Intel Architecture 32) PCを採用することで低コストなシステムを構築した。表1に開発システムの選定結果を示す。

表1 選定システム

サーバ機器	OS	WWW サーバ	開発言語
PC	Linux	Apache	Perl
IA-32	Vine Linux 2.1	1.3.23	5.005
自作機	OSS	OSS	OSS

### 3.2 システム要件

アンケート調査のシステム要件について以下に述べる。

#### 3.2.1. ユーザ認証

アンケート調査には、不特定多数の回答者を想定する場合と事前に回答者を特定する場合がある。今回のアンケート調査は、後者の特定の回答者に限定したものである。この場合、アンケートページへのアクセスに何らかの制限をかける必要がある。

アクセス制限のかけ方には、IPアドレスによる制限とユーザ認証による制限が考えられる。IPアドレスによるアクセス制限では、クライアントがグローバルIPアドレスであることが望ましいが、最近ではFW (Fire Wall) で組織内のIPアドレスを隠蔽する形で運用されている場合がある。また、全ての回答者のIPアドレスを事前に調査するのは現実的ではないため、IPアドレスによるアクセス制限は採用できない。従って、ユーザ認証によるアクセス制限を採用した。

ユーザ認証によるアクセス制限としては、アンケート回答時に回答者自身が固有の情報 (E-Mailアドレス等) を入力してユーザ登録させる方法も考えられるが、アンケート回答者以外が登録することも可能なため採用できない。本システムでは、アンケート実施者が事前にア

## WWW を利用したアンケート処理

アンケート回答者ごとのユーザ登録を行い、ユーザ名とパスワードをE-Mailで配布するものとした。

ユーザ認証の実現方法としては、プログラムで独自の認証機能を実装することも可能であるが、セキュリティホールになる可能性が高いためWWWサーバのユーザ認証機能を利用するものとした。

### 3.2.2. 任意項目の回答・訂正

今回のアンケート調査は、第87回薬剤師国家試験問題評価を目的として実施した。その内容は項目が122項目で、その内120項目（問題121～240）には小3項目の回答がある。図3に小3項目がある回答項目の例を示す。

表示件数		1 件
個別問題評価		
問題121	解答 (必須)	1 (先生方が正解だと思われる番号を選択してください)
	評価ランク (必須)	A: 適切な問題である。
	コメント	

図3 小3項目のある回答項目

このアンケートは1人の回答者が362(2+120×3)項目の回答を行う必要があるため、1度に全ての回答を入力すると回答者に大きな時間的負担を強いることになる。開発するシステムでは、回答者の負担を可能な限り軽減するように作成する必要がある。

紙面でこのようなアンケートを行った場合、回答者は何時でも自由に中断することができる。プログラムでも同様に何時でも中断再開できるように作成しなければならない。WWWによるアンケート調査では、この部分を見逃したものが多く見られ、その例としては1ページ20～30項目の回答を4、5ページに渡って入力を完了させなければならないものもある。このような作成方法を採用した場合、回答者がアンケート調査自体に大きな義務感を持っているか、または魅力的な懸賞でもなければ回答を期待できない。

回答者による中断を実現する方法として、まず考えられるのが全てのアンケート項目を1ページに表示し、任意時点で送信してもらう方法が考えられる。この場合、必須項目の回答が問題になる。必須項目は未回答が認められないものであるため図3の場合、回答者が実際

には未回答であっても‘解答：1・評価ランク：A’が送信されることになる。これを回避する方法として必須項目にも未回答の選択肢を用意する方法が考えられるが、‘解答・評価ランク’が未回答で‘コメント’のみ記入時など、不正な回答パターンが増えるためプログラムでの処理が煩雑になる。また、回答者も回答操作が判りにくくなるというデメリットが多くなる。

ここでは、1ページに表示する件数を1, 3, 5, 10件の中から回答者自身が回答しやすい件数を選択できるようにした。表示件数の全ての必須項目を回答しない限り次へ進めないようにすることで、不正な回答パターンの増加を防いだ。

次に、任意回答項目からの再開機能が必要になる。任意回答項目からの再開で前回何処まで回答したか、回答者の記憶を期待することは現実的ではない。そこで、前回何処まで回答したかを回答項目の一覧表から回答者自身が確認できるようにした。また、この一覧表から回答を再開できるように作成した。

データの訂正には、送信前の訂正と送信後の訂正がある。送信前の訂正は、送信確認画面で回答者に送信内容の確認をさせ、誤りがあればWWWブラウザの戻る機能を利用して1つ前のページへ戻ることで訂正を行う。この訂正機能で、ある程度の訂正を期待できるが、人間が回答するものであるためデータ送信後の訂正機能も必要となる。この送信後の訂正は回答項目の一覧表から行えるようにした。この訂正画面では訂正前の回答内容を含んだ形で表示し、訂正作業が効率よく行えるように配慮した。

### 3.2.3. 集計処理

データ処理において、データ送信時に不正データのチェックを行うことで異常なデータを排除できるが、集計処理時にも同様の排除処理を行うように作成した。

アンケートの集計処理を行う上で問題になるのが、図3を例にすれば‘コメント’の回答者による自由記入の処理が挙げられる。‘解答・評価ランク’は回答件数を集計するだけであり最もコンピュータ処理に向いている。

‘コメント’の自由記入は単純に集計処理することが不可能で、最もコンピュータ処理が難しいものの一つである。コンピュータによる文書内容の意味の解釈は現時点では難しいため、ここでは単純に自由記入の回答内容を一覧で表示するものとした。

‘解答・評価ランク’は、回答人数とその割合、また同時に横棒グラフをリアルタイムに表示するように作成した。

集計結果は、回答項目全てを一括表示するものと、任意回答項目のみを表示するものを作成した。また、集計結果をHTMLで出力することにより、結果をファイルに保存してWWWサーバが無くても容易に参照できるようにした。

### 3.2.4. 管理機能

本システムでは、ユーザ認証機能を利用しているためユーザの登録、削除、パスワードの強制変更機能が必要である。また、回答者の回答内容の編集機能も必要である。

回答内容の編集機能は、紙によるアンケートの場合はアンケート実施者が集計処理時に不適切な内容について対処できる。しかし、プログラムで集計処理を行う場合はコンピュータが不適切な内容かどうかを判断できないため、管理者が不適切な回答内容のチェックと編集を行う必要がある。

一般に、システムの運用・管理は1人で行うものではなく、複数の管理者により運用・管理すべきである。そのためには、管理者登録機能が必要である。本システムは、これらのシステム要件とセキュリティを踏まえて開発を行った。

## 4. プログラム開発

アンケート調査のデータ処理プログラムはPerlで作成した。プログラムの機能には、回答処理・送信確認・回答確認・回答編集・集計処理と管理機能などを実装した。以下に、WWWを利用したアンケート処理に必要なCGIプログラム固有の問題について述べる。

### 4.1 変数宣言

Perlは汎用のスクリプト言語であり、CGI専用の言語ではない。スクリプト言語はその特徴の一つとして、手軽にプログラムを記述できるように変数宣言をしなくても良いという言語仕様になっている場合が多い。Perlもこの言語仕様で作成されている。

変数宣言は、プログラムの生産性とメンテナンス性を維持するためには必要不可欠なものである。今回作成したPerlプログラムでは、`use strict`を利用して変数宣言無しを禁止し、全ての変数宣言は`my`を利用して変数のスコープを可能な限り限定した。また、グローバル変数の使用は必要最低限に止め、`Global`などのように一部に大文字を利用してローカル変数と区別した。

### 4.2 データの送受信

HTTPプロトコルでクライアントからサーバへデータを送信する方法としてはGETとPOSTの2つがある。GETは、送信データを環境変数で渡すため多量のデータ送信には不向きである。そのため、アンケート調査ではPOSTを用いるべきである。POSTは標準入力としてプログラムに送信データを渡すため、送信データ量の制限がない。

FORMの送信データは`x-www-form-urlencoded` (RFC1738)でエンコードされて送信されるため、プログラムでデコードする必要がある。本システムではPerlの置換機能を利用してデコードした。

### 4.3 文字コード

漢字コードにはEUC, JIS, SJISなどがある。JIS, SJISはPerlプログラム内での文字コードの取り扱いが煩雑になる。特にSJISは漢字コードの2バイト目に'¥'などのPerlで特別な意味を持つ文字があるため問題がある。そこで、Perlプログラム内での漢字コードは、文字の取り扱いが容易なEUCで統一した。

送信データの漢字コードは、WWWブラウザが自動的に決定するためプログラム内で扱う漢字コードに変換する必要がある。漢字コードの変換には、jcode.plを利用した。

また、送信データの改行コードも、Windows(CR/LF), Macintosh(CR), UNIX/Linux(LF)と統一されていないため、本システムではPerlの置換機能を利用してサーバOSの改行コードで統一した。

文字コードで容易に解決できないのが、半角カタカナである。解決方法としてjcode.plを利用して半角カタカナを全角カタカナに変換する方法が考えられるが、入力内容によっては漢字コードの自動判定に失敗し正常に変換できない場合がある。対処方法は、プログラムでは半角カタカナを全角カタカナへ変換するように作成し、回答者には入力画面で半角カタカナを使用しないように注意を促す2段構えにした。

### 4.4 XSS(Cross Site Scripting)対策

WWWでユーザによる入力処理を行う上でセキュリティ面から気を付けなければならないものにXSS対策が挙げられる。掲示板を例にXSSの動作の仕組みを示す。

#### XSSの動作

- a) 悪意を持った人間が入力フォームにHTML/Java Scriptなどを記述しWWWサーバに送信する。
- b) 掲示板プログラムは、受信した内容をそのままデータファイルに保存する。
- c) 別のユーザが掲示板にアクセスする。
- d) 掲示板プログラムはデータファイルからXSS目的で入力された内容を含んだHTMLを生成しクライアントに送信する。
- e) クライアントのWWWブラウザは受信したHTMLを解釈し実行する。

この仕組みを悪用することで、WWWブラウザが保持する個人情報の盗み出しやユーザにコ

表2 サニタイジング処理する文字

変 換 前	変 換 後
<	&lt;
>	&gt;
&	&amp;
"	&quot;
'	& # 39;



## WWW を利用したアンケート処理

ンピュータウイルスプログラムを実行させるなど幅広い応用が可能である。

XSSは、ユーザの入力した内容からHTMLを生成し出力するd)の段階でサニタイジング処理することで防ぐことができる。

サニタイジング処理するものを表2に示す。サニタイジング処理は、表2のHTMLで特別な意味がある文字を変換し無効化することである。

HTMLのリストボックスを利用している項目などは、リストに有るものしか入力できないと勘違いしがちであるが、これらの項目も悪意を持った者が独自の送信プログラムを作成し任意のデータを送信することは容易である。

サニタイジング処理は、ユーザの入力項目全てに対して行わなければならない。

### 4.5 ファイルのロック

WWWを利用したデータ処理では、ファイルのロック処理は非常に重要である。WWWによるアンケート調査の場合、複数の回答者がデータの送信を同時に行うことが考えられる。ファイルのロック処理を行わずに同時にデータを受信した場合、同一ファイルへの同時書き込みや書き込み途中のファイルの読み出しが発生する。このような状態になった場合、ファイルの書き込み・読み出しともにその内容は保証されない。

ファイルのロック処理は、ファイルの書き込み・読み出し双方で行う必要がある。ファイルのロック中は、同時起動された他方の書き込み・読み出しが中断されるためロック処理は可能な限り短時間で完了させる必要がある。そのためには、データファイルを1つにまとめずに複数のファイルに分割するなどして書き込み・読み出し時間を短時間で完了できるようにする。

### 4.6 管理機能

本システムの管理者は、システム管理者と一般管理者に分かれる。システム管理者はプログラム内で定義しプログラムの変更をしない限り削除不可能とし、一般管理者は一般ユーザとして登録されたユーザを管理者に昇格するという手順で登録するものとした。これはシステムを複数名で運用する場合、操作ミスなどによって誤って管理者を削除させないためである。また、一般管理者も一般ユーザに降格されない限り、削除されないように配慮した。

## 5. システムの構築と運用

アンケート調査にWWWサーバを利用することは、InternetにWWWサービスを提供することになる。サービス提供者はそのサービスのセキュリティに責務があるが、この責務を果さずに構築・運用されているものがしばしば見受けられる。従って、基本的なセキュリティに配慮したシステムの構築と運用について以下に述べる。

サーバ構築時のセキュリティ対策の基本ポリシーは‘不要な機能（サービス）は動作させな

い’である。この基本ポリシーを終始一貫することが重要である。構築したWWWサーバも、この基本ポリシーに従いWWWサービスを提供するのに必要な機能以外はインストールしない形で構築した。

アンケート調査では、回答者の個人情報を扱う場合があるため、回答内容を記録したデータファイルの取り扱いには十分な配慮が必要である。これを考慮したWWWサーバの一般的なディレクトリ構成を図4に示す。図4のhtdocsは、HTMLなどのコンテンツを保存するディレクトリである。

データファイルの保存場所がよく見られる失敗として、htdocsにデータファイルを保存するというものがある。これは‘HTMLページからリンクしていないから大丈夫’という勘違いによる失敗である。この勘違いはWWWの動作を理解していないために発生するものである。WWWの基本動作を示す。

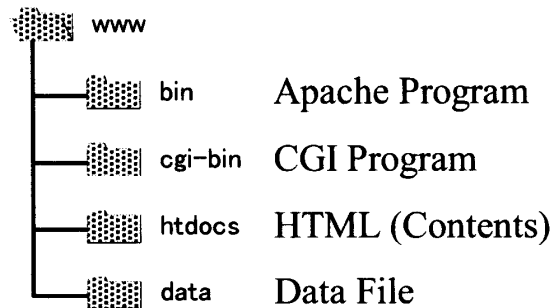


図 4

#### WWWの動作

- a) クライアントはWWWサーバにHTTPプロトコルでファイルの送信をリクエストする。
- b) リクエストを受け取ったWWWサーバはhtdocs配下にあるリクエストされたファイルをHTTPプロトコルでクライアントに送信する。
- c) a, bの手順を繰り返し行いWWWブラウザは受信したファイルを解釈しその結果を画面に表示する。

このWWWの動作を見ればページにリンクしていないファイルも、ファイル名を知っていれば閲覧でき、またファイル名を知らなくてもデータファイルは、その内容から連想しやすいファイル名が付けられるため閲覧される危険性があることが理解できる。

このように、htdocsは公開情報を保存するディレクトリであり、非公開情報を保存してはならない。データファイルなどの閲覧を禁止すべきファイルは、図4のdataディレクトリのようにHTTPプロトコルによるアクセスが発生しない場所に保存する。

また、ファイルのパーミッションにも配慮が必要である。UNIX/Linuxのパーミッションは、User/Group/Other毎にRead/Write/Executeのパーミッションが設定できる。データファイルは、UserのみにRead/Write、Group/OtherはNoneのパーミッションに設定する。運用によっては、GroupにRead/Writeパーミッションが必要な場合もあるが、OtherにRead/Writeパーミッショ

ンが必要になる場合は、システムや運用の見直しを行い、そのようなパーミッション設定を絶対行わないようにしなければならない。

## 6. おわり

アンケート調査にWWWを利用することは多くのメリットがあるため良く利用されるようになった。しかし、そのシステムや運用状況には大きな問題点がある。すなわち、アンケート調査にWWWを利用する場合、以下の点について十分な配慮を行ったシステムの開発・運用を行う必要がある。

- a) サービス提供に利用するInternet技術の正しい理解
- b) セキュリティを考慮したシステム運用
- c) 適切なデータ管理

これらは、何れもサービス提供者が対処すべきものである。

開発したアンケート調査システムは、これらの点に十分配慮して開発・運用を行った。アンケート調査にWWWを利用することは、アンケート実施者・回答者双方にとってペーパーレス・リアルタイムの集計など多くのメリットがあり、有効な手段であると考える。課題としては、WWWサーバのセキュリティホールや開発したプログラムのセキュリティ対策など継続的な対応が必要になる点が挙げられる。

最後に、第87回薬剤師国家試験問題評価のアンケート項目を提供して頂いた薬学部塩見浩人教授と田村豊助教授に深謝の意を表する。

## 参考文献

- 1) The Apache HTTP Server Project, <http://www.apache.org/>, The Apache Software Foundation
- 2) Larry Wall, Tom Christiansen, Randal L. Schwartz 共著 近藤嘉雪訳, プログラミング Perl 改訂版, オライリー・ジャパン, 1998
- 3) Randal L. Schwartz 著 近藤嘉雪訳, 初めての Perl, ソフトバンク, 1995
- 4) IPA ISEC セキュア・プログラミング講座, <http://www.ipa.go.jp/security/awareness/vendor/programming/>, 情報処理振興事業協会セキュリティセンター
- 5) 大島智樹, 平尾純一著, Perl/CGI プログラミング&サンプル, 秀和システム, 1998
- 6) jcode.pl official page, <http://srekcah.org/jcode/>
- 7) 筒本和広, 瀬島紀夫, 黒瀬能聿: WWW を利用した理解度調査システムの開発, 日本教育情報学会年会論文集14, 1998, P.224~ P.225
- 8) 筒本和広, 瀬島紀夫, 黒瀬能聿: WWW を利用したコンピュータ・リテラシ教育支援システム, 日本教育情報学会年会論文集13, 1997, P.116~ P.117

瀬 島 紀 夫・筒 本 和 広・三 谷 康 夫

## Data Processing of WWW Based Questionnaire

Norio SESHIMA, Kazuhiro TSUTSUMOTO, Yasuo MITANI

In this paper, a WWW based questionnaire processing system is proposed in order to develop the paperless procedure in questionnaire surveys, the efficient data input and its prompt publication of the results. More concretely, we propose the questionnaire processing system for the purpose of evaluating the 87th pharmacist national examination question in Japan. That is, in this study, after describing the basic mechanism of WWW for developing the system, we propose its specific data processing method using WWW and the basic security measures.

[Key Words: WWW, CGI, Perl, Internet, Security]