

XML を用いたアンケートシステム

楯 武士

E-mail : t.tate@nttcom.co.jp

NTTコミュニケーションウェア株式会社

〒 261-0023 千葉県千葉市美浜区中瀬 1-6 NTT 幕張ビル 7F-WN

概要

XML(eXtensible Markup Language)はドキュメント管理やデータ交換などの分野での利用が注目されている技術である。日本における XML の利用を促進するために Java コンソーシアム XML 部会(旧 XML 研究会)では、XML の応用事例の研究や XML 関連技術の評価のためのプロトタイプ開発を行っている。

本報告は、XML の利点とされる点を実際のシステム開発を通して検証し、その考察と述べる。

Enquete System using XML

Takeshi Tate

E-mail: t.tate@nttcom.co.jp

NTT Communicationware Corporation

NTT Makuhari Bldg., 7F-WN, 1-6 Nakase Mihama-ku Chiba-shi, Chiba 261-0023 Japan

Abstract

XML (Extensible Markup Language) is focused in many field (e.g. Document Management, EDI) since its specification is published.

Java Consortium XML Subcommittee(former XML Study Group) develop enquete system using XML technologies. In this report, we introduce how develop this system and show how practical XML is and consider merits of XML.

1. はじめに

XML(eXtensible Markup Language) は

W3C(World Wide Web Consortium)において
1998年2月にその仕様が策定されて以来、ドキュ
メント管理やデータ交換などの分野での利用が

注目されている技術であり、数多くの応用規格が登場している。今日、日本における XML の普及は米国と比べて大きく遅れているのが現状である。そのため日本における XML の利用を促進するために Java コンソーシアム XML 部会(旧 XML 研究会)が 1999 年 9 月に発足し、国内外の XML 応用事例の収集・研究や XML 関連技術の評価のためのプロトタイプ開発を行っている。

XML 部会のプロトタイプ開発活動で XML 技術を用いたアンケートシステムの開発を行った。このシステムはアンケートデータの蓄積に XML を使い、データ処理も XML 技術を最大限活用することでシステムの利便性の向上を目指している。本報告は、アンケートシステム開発を通して行った XML の利点とされる点の検証の結果を述べる。

2. 背景

XML 部会プロトタイプ開発活動では、最初に XML をデータとして保有するのに適したシステムについての思索を行った。これにより XML は以下のような利点があると判断した。

1 構造的

DTD によって文書を構造的に表現することができるため、複雑で曖昧な文書構造でも正確に把握できる。また、DTD(Document Type Definition)を用いた妥当性の検証も容易に行える。

2 汎用性

特定のプラットフォームに依存しない独立したデータが表現でき、他システムでのデータの流用が容易に行える。

3 柔軟性

データフォーマット(DTD)の変更に際して DTD の変更だけで柔軟に対応でき、旧データフォーマットとの互換性の維持も可能である。

4 インターネットとの親和性

XML は SGML をベースに、また HTML を参考に策定されたため、インターネットとの親和性が高い。

5 オープンな仕様

W3C で策定されたオープンな仕様であるため、様々なベンダによる応用規格やその実装が急速に増えている。

この結果、プロトタイプ開発の対象をメールや Web などのインターネットを経由して行われるアンケートに決定した。アンケートは質問と回答の組からなり、回答のフォーマットも規定されているため文書が構造的であると判断したためである。そこで我々はアンケートの回答データを XML で保存することとした。

今回のシステムは Java コンソーシアムの事務局が行っているアンケートの配布・回収・集計作業を対象とした。その際に以下の点を留意して検討を進めた。

- **運用者にプログラム知識を要求しない**

アンケートの作成は事務局担当者が行うが、担当者に対してプログラムの知識、あるいは XML 関連のスキルを求めない。

- **メンテナンスが容易**

この開発は XML 部会メンバの有志が短期間開発によって行うため、メンテナンスにリソースを割くことが難しい。そのためメンテナンスの稼働を最小に押さえたシステムを構築する。

- **現在のアンケートの形態を変更しない**

次節で述べる現在 Java コンソーシアムで行われているアンケートに関する作業の変更を最小限にする。

3. 従来の作業内容

現在、Java コンソーシアム事務局が行っているアンケートは、1) セミナーやミーティングなどイベントへの出欠席の確認 2) イベント後の意見・感想の収集 3) 活動内容に対する意見収集などのアンケートの収集が行われている。1回のアンケートの流れは以下のようになっている(図 3-1)。

- **アンケート配布**

アンケートの実施者(イベント主催者など)があらかじめ作成したアンケートをイベントの前後にメールで対象者に配布、またはイベント開催時に紙で配布する。

- **アンケート回収**

メールで配布したものについては、アンケート回答者が回答を記入し、事務局担当者のメールアカウントに返信する。イベント開始前に紙で配布されたものについてはイベント終了時に事務局担当者が回収する。

- **アンケート集計**

メール、紙いずれの場合も事務局担当者が手作業にて集計し、集計結果レポートを作成し、アンケートの実施者に報告を行う。

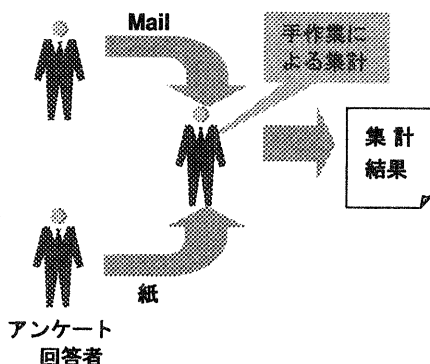


図 3-1 従来のアンケート処理

次に従来行われているアンケートのフォーマットについて整理する。現在の Java コンソーシアムのアンケートは一般的なアンケートと同じように次の2つの回答の形式がある。

- **記述式**

回答者が回答内容を記述する形式

- **選択式**

回答者があらかじめ用意された選択肢の中から回答を選択する形式

例として実際に Java コンソーシアムで用いられたアンケートを示す(図 3-2)。これは回答者の氏名、所属企業名を記述し、参加するワーキンググループを選択するアンケートの一部である。

Q1: 氏名
[植 武士] 記述式

Q2: 企業名
[NITコムウェア]

Q3: 参加 WG
 : 応用技術 WG
 : 基盤技術 WG
 : モデル WG 選択式

図 3-2 従来のアンケート例

4. アンケートシステム

前節で示した従来の作業内容を検討した結果、以下のシステム化を行うこととした。

- メールによるアンケート回収は従来通りに行う。アンケートの配布についてはシステムの対象外とする。
- 集計についてはシステムが自動的に行う。

このときに事務局担当者がデータを多角的に分析を行えるような機能を盛り込み、さらに集計結果のグラフ表示を行うことで分析を支援する。

- 紙のアンケートはシステム化が困難であると判断し、今回のシステム化の対象外とする。
- 紙によるアンケートの代替手段として Web によるアンケート投入機能を追加した。イベント時にアンケート用 URL を周知し、イベント参加者が Web ブラウザで回答を投入することで、従来は紙によるアンケートを実施していた不特定回答者から回答を収集することができる。

図4-1にシステムの概略図を示す。

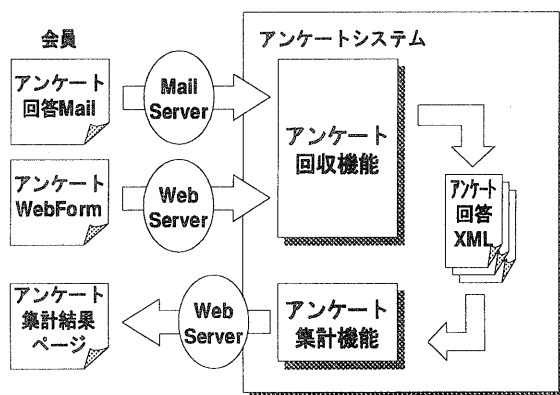


図 4-1 システム概略図

4.1. アンケート回収

アンケートの回収は、メール、Web どちらの方法で回収した場合も、結果的には、一箇所に集約され、それぞれのアンケート毎のファイル(アンケート回答 XML)として管理される。システムのバックエンドでは、特にデータベースで管理する方法はとっていない。これは、アンケート特性が、個々のアンケート結果を後で更新するような場面がなく、静的に蓄えられる性質があることから本方式を選択した。

4.1.1. メールによるアンケート回収

メールによる回収では、所定の返信メールアドレスに送られたメールをシステムが自動的にサーバから取得し、XML 文書形式に変換する機構とした。また、アンケート回答メールを XML 形式で運用できないか検討を進めたが、回答者が安全かつ容易に XML 文書を編集できなければならない問題のため、回答データの解析を独自に行う必要が生じた。アンケート回答メールはプレーンテキスト形式を採用することとし、返信内容からデータが取得できるよう独自の識別子を規定することでデータ解析を行う方法とした(表 4-1)。図 4-2 にメールによるアンケートの例を示す。

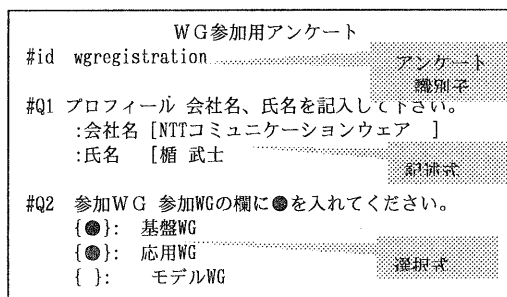


図 4-2 メールによるアンケート例

表 4-1 メールによるアンケートの識別子

識別子	意味
#id	アンケートの識別子
#Qn(n=1,2,3...)	質問文の開始を示す識別子
[]	記述式アンケートの解答欄
{ }	選択式アンケートの解答欄

このようなことから、メールによるアンケート回収機能は、内部で XML パーサを利用せず、XML 文書を生成している。

4.1.2. Web によるアンケート回収

メールによるアンケート回収機能と同様、Web フォームによる回収機能でも入力データを XML

形式として扱えないか検討を行ったが「担当者
にプログラム知識を要求しない」制約から、フォー
ムにスクリプト言語を用いたクライアント処理
を組み込むような方法は避けなければならなかつ
た。これにより Web フォームに投入された回
答データをクライアント側で XML に変換し回答
機能へ渡す方法がとれず、XML 文書を回収機
能の入力とすることはできなかった。

Web フォームによるアンケート回収機能では、
フォームからの返信(HTTP-POST)データを解
析する前段とアンケート回答 XML ファイルを生
成する後段の2フェーズからなり、後段にXSLT
によるデータ変換を利用することで、出力形式
への柔軟な対応等の汎用性を高めている。

図 4-3 に Web によるアンケート回収処理の概
要を図 4-4 に Web でのアンケート投入例を示
す。

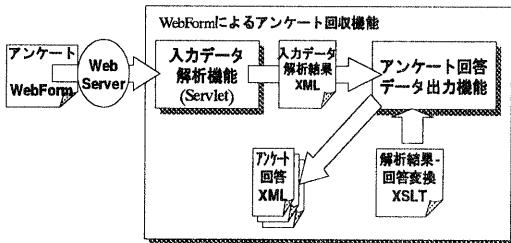


図 4-3 Web によるアンケート回収

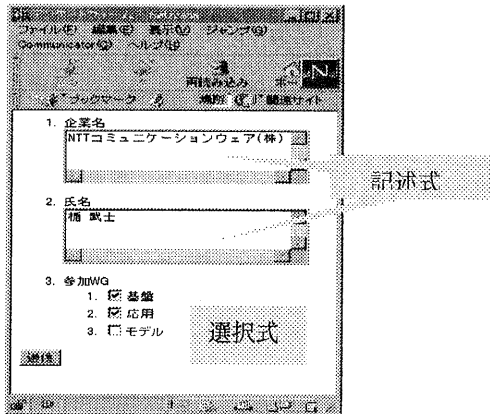


図 4-4 Web によるアンケート例

4.1.3. アンケート回答 XML

前述のようにメール、Web どちらの方法で投入
されたアンケートも同じ DTD(図 4-5)に従う XML
が生成される。表 4-2 に各要素名と属性名の説
明と図 4-6 に生成される XML の例を示す。

```
<!ELEMENT questionnaire_idv (header?, body)>
<!ATTLIST questionnaire_idv formid CDATA
#REQUIRED>
<ELEMENT header (formtitle?, formversion?,
formdate?, receivedate?)>
<ELEMENT body (question)+>
<ELEMENT question (answer*)>
<!ATTLIST question
number ID #REQUIRED
name CDATA #IMPLIED
type (select|descript) #REQUIRED>
<ELEMENT answer (selection* | descript*)>
<ELEMENT selection (#PCDATA)>
<!ATTLIST selection no CDATA check
(yes | no) #REQUIRED>
<ELEMENT description (#PCDATA)>
<!ATTLIST description name CDATA #IMPLIED>
```

図 4-5 アンケート回答 XML の DTD

表 4-2 XML ファイルの要素と属性

要素/属性名	要素/属性	意味
questionnaire_idv	要素	ルートエレメント
formid	属性	どのアンケートの回 答かを示す識別子
header	要素	付加情報を記載す るヘッダ要素 (optional)
question	要素	質問の回答を表す 要素
name	属性	質問名(optional)
number	属性	質問番号
type	属性	記述式か選択式か を示す識別子
answer	要素	質問の回答を表す 要素
description	要素	記述式の回答を表 す要素

要素/属性名	要素/属性	意味
selection	要素	選択式の回答を表す要素
no	属性	選択肢番号
check	属性	選択されたかを示す属性(yes or no)

```

<? xml version="1.0" ?>
<questionnaire_idv
  formid="wregistration">
  <header/>
  <question number="Q1"
    name="プロフィール"
    type="descript">
    <answer>
      <description name="会社名">
        NTTコミュニケーションウェア
      </description>
    </answer>
  </question>
  <question number="Q2">
    <answer>
      <description name="氏名">
        榎 武士
      </description>
    </answer>
  </question>
  <question number="Q2"
    name="参加WG" type="select">
    <answer>
      <selecton no="1" check="yes">
      <selecton no="2" check="yes">
      <selecton no="3" check="no">
    </answer>
  </question>
</questionnaire_idv>

```

図 4-6 アンケート回答 XML の例

4.2. アンケート集計

回収されたアンケートは、アンケート集計機構で集計される。集計結果は事務局担当者以外の方がネットワークを介して参照できるように Web ブラウザで閲覧できるようにした。また、様々なビューで内容を確認できるような機能を付加した。この処理は、ある質問に何人の人が選択肢1をチェックしたかということを集計するといったロジックが記述された集計用 XSL スタイルシートと、集計結果を HTML や SVG(Scalable Vector Graphics)グラフに変換する結果表示用 XSL スタイルシートの2種のスタイルシートを組み合わせることによって

行われる。この際に集計毎に集計用 XSLT スタイルシートと結果表示用 XSLT スタイルシートをユーザに選択させるのではなく、予め想定される組み合わせを変換規則用 XML に記述しておき、ユーザに変換規則用 XML を選択させることでシステムの利便性の向上を行っている(図 4-6)。

この結果、多角的な分析を実現するとともに、分析ロジックと表示方法の分割が図れ、スケーラビリティの向上に繋がった。

また、これらの対策により、集計機能は集計規則をすべてプログラムから除外することができ、集計用途に限らず広範囲にデータを加工する機能として汎用化することができた。

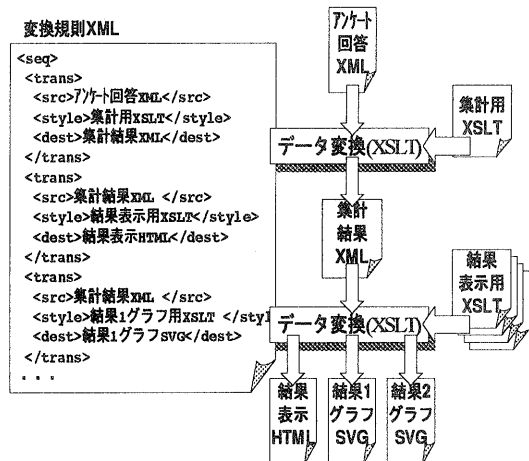


図 4-6 集計処理概要

5. 実装

本システムは、XML 部会メンバーが分割した機能を1つずつ担当し開発を行った。この開発に当たり、以下の点を重視した。これら項目に対しても XML 技術を適用することで達成できた。

- **機能間インターフェースは正確かつ簡潔**

機能間入出力は XML ファイルとし、インターフェース規定は DTD のみで行うことで、機能間の結合度を疎にし、インターフェースを理解する上

での意思疎通に時間をかけずに正確に行うことができた。

• 開発規模の抑止

XML Parserを用いることで、入出力部を独自開発しない。実際にこれによって集計機能は0.3KLで実現している。

6. 考察

本システムの開発から得られたことを元に、2章で挙げたXMLの各メリットについて考察する。

1. 構造的

汎用的かつシンプルなDTDを設計することで、記述式、選択式からなる一般的なアンケートに対して構造的なXMLを容易できた。このDTDをインターフェース定義として利用することで、アンケート回答XMLの構造に即した集計ロジックを構築することができた。

2. 汎用性

今回のアンケート回答XMLデータを他のシステムへポーティングすることは行っておらず、汎用性は未検証である。このXMLの構造は単純であるため、データのポーティングも容易に行えると考えられる。また、DTDを用いて妥当性を検証することで、ポーティング時の構造のチェックを容易に行うことができる。しかし、複雑な構造を持つXMLのポーティングは変換作業が大きくなるという問題がある。つまり、XMLの汎用性(データのポーティングの容易さ)とデータ構造の表現能力はトレードオフの関係にあることが言える。

また、今回のシステムはJava言語で開発したためにシステム全体の汎用性も向上した。

3. 柔軟性

今回、アンケート回答XMLのDTDの変更は行っておらず、柔軟性は未検証である。しかし、

今後アンケートフォーマットの変更に伴うアンケート回答XMLに対する要素や属性の追加に際して、DTDの変更を行うときに新たな要素(または属性)をオプションに定義すれば、XMLデータを変更せずに利用することができる。

また、DTDの変更が大きい場合、既存のデータとの整合性が取れなくなり、データの変換が必要になる。これも前項と同じように柔軟性(データフォーマットの変更に対するデータの耐性)と構造的なトレードオフとなると言える。

4. インターネットとの親和性

本システムでXMLとインターネットとの親和性について最も高く感じたのは、アンケート集計結果の表示に関する部分である。集計結果XMLを、XSLを利用してHTML、SVGに変換することで、Webブラウザへの表示を容易に行うことができた。また、今後WebブラウザのXML対応が進むことで、XMLとインターネットの親和性がより向上することは明らかである。

5. オープンな仕様

多くの分野でのXML応用規格が策定されているが、本システムの開発においてはXMLデータにアクセスするためのAPIにはDOM(Document Object Model)を、変換にはXSLを用いた。これらのXMLの応用規格を実装したツールを利用することで、前節に示したようにシステムの開発量が削減できた。また、集計結果のグラフィック化にSVGを用いて、結果分析の表現能力を高めた。これらのことから、XMLの応用分野の広さとそれら応用規格を実装したツールを利用することによる開発工数の削減を感じた。

次に本システムを通して得られたXMLの問題点を示す。

本システムでアンケートの回答数が増えるとそれに伴ってXMLのファイル数も増加する。このと

きに集計処理に多くの時間がかかるという問題点が顕在化した。この原因は次の 2 点であることが分かった。

1. XML は他のデータフォーマットと比べてタグなどの付加情報が多く、処理するデータサイズが大きくなること
2. パターンマッチが多く、一般に性能が劣るといわれる XSL の処理を多用したこと

1の問題点については、XML 対応のデータベースを利用して効率的にデータを抽出するなどの対応が可能であると考えられる。2については、XSL スタイルシートをプリコンパイルするツールが現在開発されている。これにより XSL による変換速度の向上が見込める。

以上から、XML はアンケートのようなインターネット上で交換される構造的なドキュメントデータのフォーマットとしての有効性が確認された。この際に DTD の汎用性を高めることがデータ交換の容易さや仕様変更に対する耐性の向上に寄与するが、文書構造の表現能力の低下を招くことが確認された。つまり、データの交換が多く、異なるタグセット間の変換が多い EC や EDI の分野のシステムでは DTD の汎用性を高めることが望ましく、逆にデータをローカルに蓄積し、複雑なデータ構造を持つドキュメントを管理するようなシステムでは特化した DTD を用意することでデータの処理が容易になると言える。

また、XML の幅広い応用規格の実装系を用いることでシステムの利便性を向上することができ、データのポーティングについてもその可能性を検証することができた。

7. おわりに

本報告は、XML 技術を用いたアンケートシステムの開発を通して、XML のメリットの検証を行った。このシステムでアンケートデータの蓄積に XML を用い、データ処理も DOM、XSL、SVG などの

XML の応用規格を実装したツールを最大限活用することで、システムの利便性の向上や開発工数の軽減を行うことができた。

また、XSL などの最新の仕様については性能を要求されるシステムでの使用は課題が残ることを示した。XML はアンケートのようなインターネット上で交換される構造的なドキュメントデータのフォーマットとして有効であることが確認された。

謝 辞

本システムの開発にご協力を頂いた Java コンソーシアム XML 部会の方々と Java コンソーシアム事務局の方を紹介し、感謝の意を表す。(企業名五十音順、敬称略)

NTT コミュニケーションウェア	山田 洋一
大建情報システム	大竹 浩
大建情報システム	花島 一雅
鉄道情報システム	青沼 慶伍
日立製作所	勝目 猛
日立製作所	森 徹
日本アイ・ビー・エム	吉田 敏幸
日本ユニシス	井上 祐司
日本ユニシス	藤井 淳一
日本ユニシス	牧野 友紀
日本ユニシス	宮崎 良地
ライトウェル	渡邊 光
Java コンソーシアム事務局	野田 博正

参考資料

- [1] XML 仕様
<http://www.w3.org/TR/1998/REC-xml-19980210>
- [2] SVG 仕様
<http://www.w3.org/TR/2000/CR-SVG-20000802>
- [3] XSLT 仕様
<http://www.w3.org/TR/xsl>
- [4] DOM 仕様
<http://www.w3.org/TR/REC-DOM-Level-1>
- [5] Java コンソーシアム XML 部会
<http://www.javacons.gr.jp/xml/>