

データベース入門 資料 7

XSLT による Web ページ用文書の生成

平成 23 年 5 月 30 日

目 次

1	XML 文書から Web ページ用文書を作る	1
1.1	データと表現形式	1
1.2	HTML	1
1.3	XML から HTML への変換	2
1.4	参考 : Web ブラウザを使った XSL 変換 — XML 文書への処理命令の追加	4
2	練習問題	4

1 XML 文書から Web ページ用文書を作る

1.1 データと表現形式

これまでは、主に XML 文書から必要なデータを抽出する目的で XSLT を用いてきた。その一方で、XML 文書の表示の形式を決める用途に XSLT を使うことができる。

例えば Web ページはデータを提示するのに適した表現形式の一つであり、XML 文書から Web ページ用の文書を作ることができれば、人がデータを閲覧するには便利である。それに対し、用紙に印刷する文書は Web ページとは異なる見栄えのものが好ましいものの、意味的に同じデータを Web と印刷用に各々用意するのは効率的ではない。

このような場合、データは一つの XML 文書として作成しておき、データの表現を定める役割を XSLT テンプレートに任せればよい。これにより、データを更新する必要が生じた際には元の XML 文書だけを更新し、XSLT テンプレートを適用するだけでよくなり、複数の文書に分散したデータを、それぞれ更新する必要はなく、データの保守が容易になる。

このように、XSLT による XML の変換を使えば、データを格納したひとつの XML 文書を、様々な形で表現することが可能になる。

以上のような、データそのものと表現形式を切り離して、データを一元化するという考え方は、データ処理一般において重要である。人が読む通常の文章の作成においても、文章の意味・論理構造と表示形式を分離を分離して、文章を作成することには意義がある。XML の利用は、これらを実現する方法の一つである。

1.2 HTML

HTML (HyperText Markup Language) は Web ページ記述用のマークアップ言語である¹。XML と HTML は共に SGML というマークアップ言語を基に開発されたものなので、両者の記法や用語は似ている。XML との違いは使えるタグの名前が予め決められている点や、開始タグのみの要素があること、タグ名に大文字・小文字の区別がない等である。また、HTML はあくまで Web ページ (人間が読むもの) の記述用言語であり、XML のような汎用性のあるデータ記述言語ではない。

以下に簡単な HTML の例を示す。ここでは、body 要素の中のタグ (h1, h2, a, p, table) の使い方が分かれば十分である²。(Web ブラウザでの表示はこちら³)。

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">

<html>
<head>
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
<title>sample HTML</title>
</head>
<body>
<h1>HTML 文書のサンプル</h1>
<h2>ハイパーリンク</h2>
```

¹HTML の上位言語として XHTML (eXtensive HyperText Markup Language) がある。これは HTML を XML の仕様に則った形にしたものと言えるが、ここではより一般的で記述が容易な HTML を使う。

²body 要素の内容だけを記述しても Web ページは表示される

³<http://echoes.hak.hokkyodai.ac.jp/db/550/sample.html>

外部文書へのリンクには href 属性を持つ a タグを使います

(例: データベース入門)。

同じコンピュータ内のファイルにリンクを張るときには、

href 属性値にファイルのパス名が使えます。

同じディレクトリ内のファイルであれば、href="ファイル名" でいいです。

<p>p タグは段落を作ります。</p>

<h2>表</h2>

表を作るには table タグを使います。border 属性は罫線の太さ指定です。

表の中に行を作るのは tr タグです。

行の中に何か書くには td タグを使います。行や列や表題は th タグです。

```
<table border="1">
  <tr>
    <th></th><th>列 A</th><th>列 B</th>
  </tr>
  <tr>
    <th>行 1</th><td>1A</td><td>1B</td>
  </tr>
  <tr>
    <th>行 2</th><td>1A</td><td>2B</td>
  </tr>
</table>
</body>
</html>
```

サーバ上の Web ページではなく、自分で作成したファイルを Web ブラウザで直接に表示するには、ブラウザのメニューからファイルを開く操作を行うか、URI 欄に file:// に続きファイルの絶対パス名 (例: file:///pub/db/xml/sample.html) を入力すればよい。

1.3 XML から HTML への変換

XSLT を使って XML 文書を HTML 文書に変換すれば、XML 文書内のデータを Web ページとして見ることができる。

以下に、参考文献の入った XML 文書⁴ (/pub/db/xml/ref.xml) から HTML 文書を出力するための XSLT スタイルシート例⁵ (/pub/db/xml/ref2html.xsl) を示す。これまでに示した XSLT テンプレートと異なるのは、次の 3 点である。

1. xsl:output 要素の method 属性値を html にすることで、テンプレートに記述したタグを出力させる⁶。
2. テンプレートに XSLT の要素以外のもの (HTML のタグ) を記述している。

⁴<http://echoes.hak.hokkyodai.ac.jp/db/550/ref.xml>

⁵<http://echoes.hak.hokkyodai.ac.jp/db/550/ref2html.xsl>

⁶<xsl:output method="text"> はテンプレートに記述したタグ等は除いて、文字データのみ出力する指定である。

3. HTML の href 属性値を XML 文書から得るために、属性値テンプレート $\{location-path\}$ を用いている。

XSLT スタイルシートのテンプレート記述において、タグの属性値部分に XML 文書のデータを埋め込みたければ、 $\{ \}$ 内にロケーションパスを記述すればよい。これを属性値テンプレート (attribute value template) という。

下記のスタイルシートでは、`<a href="{uri}">` において、XML 文書内の uri 要素の文字列値を取得し、それを Web ページのリンク先 URI として使う。

```
<?xml version="1.0" encoding="UTF-8" ?>
<xsl:stylesheet version="1.0"
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
  <xsl:output method="html" encoding="UTF-8" />
  <xsl:template match="/">
    <html>
    <head>
    <title>Reference List</title>
    </head>
    <body>
    <h1>Reference List</h1>
    <h2>book</h2>
    <table>
    <xsl:for-each select="ref/book">
      <tr>
        <td><xsl:value-of select="@isbn" /></td>
        <td><xsl:value-of select="title" /></td>
        <td><xsl:value-of select="author|editor" /></td>
      </tr>
    </xsl:for-each>
    </table>
    <h2>web</h2>
    <table>
    <xsl:for-each select="ref/web">
      <tr>
        <td><a href="{uri}"><xsl:value-of select="title" /></a></td>
      </tr>
    </xsl:for-each>
    </table>
    </body>
    </html>
  </xsl:template>
</xsl:stylesheet>
```

上記スタイルシートによる変換結果をブラウザで表示するには、こちら⁷を参照。

1.4 参考: Web ブラウザを使った XSL 変換 — XML 文書への処理命令の追加

XML パーサを持った Web ブラウザは XML 文書の構造を解析して表示できるが、XSLT プロセッサの機能も併せ持つ Web ブラウザを使えば、XSLT による XML 文書の変換と結果の表示を Web ブラウザのみで行える。

そのためには、変換対象となる XML 文書内に、適用したいスタイルシートの情報を書き込む必要がある。これを XML では処理命令 (PI; processing instruction) という。処理命令を記述する位置は、ルート要素よりも前としなければならない⁸。

次のものは、XSLT による変換を行う処理命令の例である。適用したい XSLT スタイルシートは href 属性で指定する。Web サイト上にスタイルシートがある場合には、

```
<?xml-stylesheet type="text/xsl" href="http://echoes.hak.hokkyodai.ac.jp/db/550/ref2html.xsl"
```

のような記法が可能であり、XML 文書と同じコンピュータにあるファイルをスタイルシートとして指定するのであれば

```
<?xml-stylesheet type="text/xsl" href="/pub/db/xml/ref2html.xsl" ?>
```

のようにパス名を記述すればよい⁹。

処理命令が書かれた XML 文書を Web ブラウザで開けば、XSLT 変換の結果を得ることができる¹⁰。

2 練習問題

1. 参考文献の入った XML 文書 (/pub/db/xml/ref.xml) に、第 1.3 節の XSLT スタイルシート例 (/pub/db/xml/ref2html.xsl) を適用し、結果を出力しなさい。xsltproc コマンドを使うこと。

その出力をパイプと less を使って確認してから、リダイレクトを使ってファイルに保存しなさい。ファイル名の終り (拡張子) は .html とする。保存したファイルを Web ブラウザで表示しなさい¹¹。

2. 予備問題: 前の資料「6 XSLT の反復構造¹²」の第 3 章の練習問題 3.3 で作成した XSLT スタイルシートを、HTML を出力するように変更してみなさい。

さらに、そのスタイルシートと xsltproc コマンドを使って、郵便番号表の入った HTML ファイルを作成し、ブラウザで表示しなさい。

また、XML 文書に処理命令を追加して、xsltproc コマンドを使わずに、ブラウザのみで表を閲覧できるようにしなさい。

⁷http://echoes.hak.hokkyodai.ac.jp/db/550/ref_PI.xml

⁸もちろん XML 宣言より後ろ。

⁹XML 文書と同じディレクトリに存在する XSLT テンプレートを適用する場合、href 属性には、パス名に代え、ファイル名のみ記述してよい

¹⁰実際に処理命令を実行できるかどうかは、XML 文書を扱うプログラム (Web ブラウザ) に依存する。

¹¹firefox -no-remote を実行すれば、恐らく firefox をリモートのコンピュータで起動して、ウィンドウをローカルに表示可能

¹²http://echoes.hak.hokkyodai.ac.jp/db/550/xslt_foreach.html