

膨大な情報量を持つWebでは、ユーザーにとって欲しい情報だけを見やすく提供するパーソナライズ技術が重要である。そこで、XML(eXtensible Markup Language)ページに含まれる任意の部分ページをパーツ化し、好みの部分ページをドラッグ&ドロップを用いて一つのページへ合成編集できる、クライアント中心のWeb情報のパーソナライズシステムeXcart_{TM}を開発した。eXcart_{TM}を用いることにより、ブラウザ上で閲覧しながら、サイトを越えて好みの記事の切抜きを集めたり、異なるサイトの商品カタログを同一ページで統一表示して見ることができる。

Personalization technology, which provides the information required by individual users in a well-ordered manner, has become increasingly important due to the enormous volume of contents of the World Wide Web.

Toshiba has developed eXcart_{TM} as a client-side Web personalization system that enables users to collect sub-resources of different Web pages defined as Web parts from different Web sites with a drag-and-drop operation, and then merge them into an original page. Users can therefore collect their favorite articles across Web sites, or compare goods offered on different sites, in a unified page view with eXcart_{TM}.

1 まえがき

最近、Webサービス⁽¹⁾、ESI(Edge Side Includes⁽²⁾)、シェアポイントポータルサーバ⁽³⁾など、Webページをパーツ化し、再利用しようという動きが多く見られるようになってきている。これらの共通の技術基盤はXML⁽⁴⁾である。

XMLは、Webページとしての利用を想定して設計されているが、キラーアプリケーションソフトウェアの欠如により、今までWebページとしての利用を見かけることはほとんどなかった。しかし、これらの再利用の動きを見るに、今後XMLのWebページとしての重要性が高まりそうである。

ここで紹介するeXcart_{TM}はXMLを用いたWebのパーソ

ナライズ技術の一つであり、ユーザーがブラウザ上で、パーツ化されたWebページをサイトを越えて自由に組み合わせ所望のオリジナルページを作成するという、閲覧時のWeb合成機能を提供する。これにより、オープンで自由度の高いパーソナライズ機能をユーザーに提供できる。eXcart_{TM}を用いたWeb情報の合成イメージを図1に示す。

以下2章でパーソナライズ技術に関する背景を述べ、3章でeXcart_{TM}のモデル・構成を説明し、4章でeXcart_{TM}の適用事例として電子ちらし作成システムについて述べる。

2 背景

2.1 Webのパーソナライズ技術

WWW(World Wide Web)は情報基盤として急速な普及を遂げ、あらゆる情報がWWWを通じて提供されるようになってきている。従来は、企業紹介や観光情報など、手作りのコンテンツで構成される静的で非定型的な文章中心のWebページが基本であったが、データベースなどバックエンドシステムから機械的に動的生成される定型的なWebページがとて増えてきている。

このようなページは、概してページ全体としてのまとまりは弱く、複数の独立したサブコンテンツの寄り集まりとして構成されていることが多い。その結果、ユーザーは、ページ全体に対してではなく、各サブコンテンツに対して興味を示すことが多くなり、次に示すようなサブコンテンツ単位で情

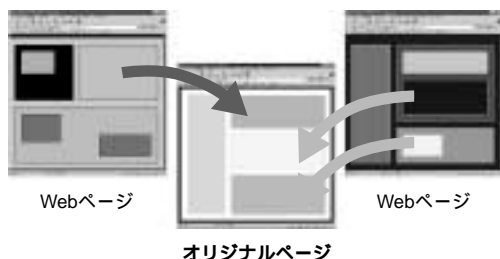


図1. 閲覧時のWeb合成機能の実現イメージ ユーザー所望のWeb情報を組み合わせてオリジナルページを作成する。

Concept of collecting and rearranging Web information across different sites with browser

報を取り扱いたいというニーズが高くなっている。

- (1) サブコンテンツを分類・整理したい。
- (2) サブコンテンツ単位で保存したい。
- (3) 興味のあるサブコンテンツどうしを比較したい。

例えば、商品カタログが列挙されている複数のショッピングサイトから、興味のある商品どうしを比較するとか、複数のニュースサイトから、興味のある記事だけを組み合わせた新聞切抜きを作成する、といった利用ケースが考えられる。

しかし、Webブラウザは、従来の文章中心のWebページを読み進める用途に対しては最適であるが、このようにサブコンテンツ単位でWebページを取り扱うことはできない。

そこでわれわれは、Webページ内に含まれる独立性の高いサブコンテンツを、Webページ本体から取り出したりはめ込むことのできるWebパーツとしてモデル化し、Webブラウザ上で、ユーザーの意図でページ内のWebパーツをドラッグ&ドロップして自由に組み合わせる機能を提供する、Web情報のパーソナライズシステムeXcart_{TM}を開発した。

2.2 XML

Web情報の組合せ機能を実現するには、閲覧しているWebページを複数のWebパーツに分解し、それらWebパーツを組み合わせるWebページを合成する仕組みを、ブラウザの動作するクライアントマシン上に導入する必要がある。

Webパーツ処理を実現するには、Webページからデータを取り出したり追加したりするなど、データ加工が不可欠である。しかし、HTML(HyperText Markup Language)は次の理由からデータ加工は苦手であるため、HTMLベースでWebパーツ処理を実現するのは現実的でない。

- (1) 使われているタグはもっぱら表示方法のために用意されている。
- (2) データと表示情報が混在しているため、データを構造的に取り出すのが難しい。
- (3) 文法があいまい(方言が多い)であり、情報処理のための共通API(Application Programming Interface)が用意されていない。

一方、XMLは、上記のようなHTMLの欠点を補う次世代のWebページ記述言語として、W3C(World Wide Web Consortium)で制定された経緯があり、データ指向であるため、データ加工処理に適している。主に次のような特徴がある。

- (1) XMLページでは、データ(XML文書)と表示情報(スタイルシート)が分離されている。
- (2) XML文書は、データ構造を表すユーザー定義タグを使ってデータを記述する。
- (3) スタイルシートは、関連づけられたXML文書のデータをどのように表示するかを記述する。
- (4) XML文書のデータを操作する標準APIとしてDOM(Document Object Model)が用意されている。

このようにデータと表示が分離されているため、レイアウトなど表示方法にとらわれず、DOMを通じて、データ構造を取り出したり操作したりすることができる。

以上を踏まえ、eXcart_{TM}の合成機能の実現にXML技術を用いる方針とした。

3 Web パーツ合成システム eXcart_{TM}

3.1 Web パーツモデル

ここでは、eXcart_{TM}の設計のベースとなっているWebパーツモデルについて述べる。

Webパーツとは、XMLで書かれたWebページ(XMLページ)において、パーツ組合せが有用な再利用性の高いXMLのサブリソースである。例えば書籍XMLページでは、図2のように各bookタグ以下のXML文書が該当し、そのコンテンツをWebパーツコンテンツと呼ぶ。

```
<?xml version= 1.0 " encoding= "Shift_JIS "?>
<bookstore>
  <name>東芝書店</name>
  <location>Yokohama</location>
  <books>
    <book pz : component= " yes " stocks= 10 ">
      <author>東芝太郎</ author>
      <name>東芝の歴史</ name>
      <price>2000</price>
    </book>
    <book pz : component= " yes " stocks= 4 ">
      <author>東芝花子</ author>
      <name>東芝物語</ name>
      <price>1000</price>
    </book>
  </books>
</bookstore>
```

図2 . Web パーツコンテンツの例 各書籍情報がeXcart_{TM}専用タグ属性によってWebパーツコンテンツに指定されている。

Example of Web part content

どのサブリソースがWebパーツコンテンツであるかは、Webページを提供するWebサイト側がデザインする。この指定にはeXcart_{TM}が定めた専用のタグ属性pz:componentを使い、任意のタグ上に書き込むことができる(図2)。

また、Webパーツは単独もしくは任意の組合せで表示可能でなければならない。これに対して、Webパーツコンテンツが所属するスタイルシートは、ページ全体に適用することを前提として作成されているため、そのまま利用することはできない。Webパーツコンテンツを表示可能にするためには、Webパーツコンテンツを収容していたコンテナ文書もあわせて必要となる。このことから、各Webパーツは、表示に関する付加情報として、スタイルシートとコンテナ文書を所有するものとする。また、スタイルシートとコンテナ文書を合わ

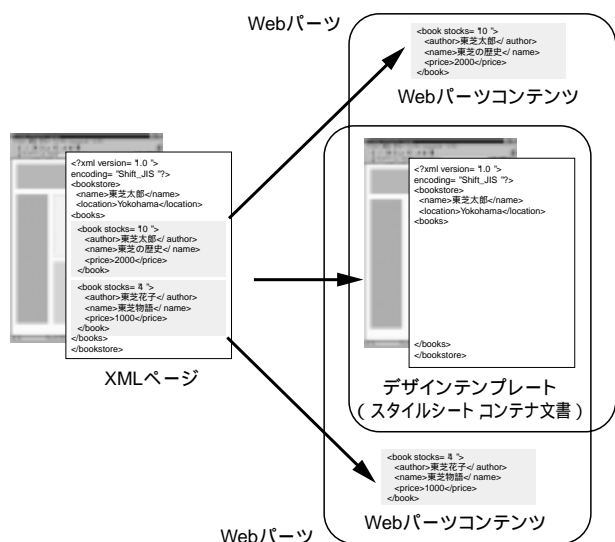


図3 . Web ページのパーツ化 各 Web パーツにはコンテンツ以外に表示のためのデザインテンプレートが与えられる。
Mechanism of decomposition into several parts of Web page

せてデザインテンプレートと呼ぶこととする。

これらを図示すると図3のようになる。同じ Web ページ内の各 Web パーツは、デザインテンプレートとして同じスタイルシートとコンテナ文書を持つことになる。

3.2 システム構成と動作

上記 Web パーツモデルに基づいて設計した eXcart_{TM} システムの構成を図4に示す。eXcart_{TM} は、Web パーツを取り扱えるよう Web ブラウザを拡張し、Web ブラウザが Web パーツを含むページを開いた場合に、Web パーツのドラッグ&ドロップ操作による情報抽出と蓄積から、収集された Web パーツを組み合わせた合成ページの出力までを処理する。技術的特長の一つは、各 Web パーツがコンテンツとビューを持っており、合成ページを作成する際に、コンテンツの組合せ及び適用するビューを対話的に選択できる点である。

図5は eXcart_{TM} の実行画面例である。Web ページの左上側に表示されているエクスプローラパー領域は、Web パーツを蓄積するカート機能を果たしており、ユーザーは、Web ページからこのカートへ Web パーツをドラッグ&ドロップすることで、任意のサイトから Web パーツを収集できる。

収集する Web パーツは、データ構造に互換性があれば、デザインが一致している必要はない。各 Web パーツのデザインは、左下のプレビューミニペインで確認することができる。

また、eXcart_{TM} は、カート上に集められた Web パーツを組み合わせて Web ページを合成することができる。このとき、合成対象となる Web パーツのデザインが複数種類ある場合、eXcart_{TM} はどの表示形式をベースに合成するかをユーザーにたずね、選択された Web パーツのデザインテンプレートを元に、各 Web パーツを一つの Web ページへ合成し、

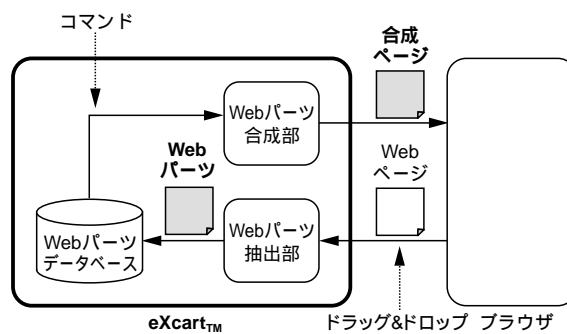


図4 . eXcart_{TM} のシステム構成 eXcart_{TM} はブラウザの拡張機能として動作し、Web パーツを分解・合成する。
System configuration of eXcart_{TM}



図5 . eXcart_{TM} の実行画面例 ページ上の任意の Web パーツを eXcart_{TM} のカート領域へドロップすることで、Web パーツを蓄積できる。
Example of eXcart_{TM} screen display

ブラウザ上に表示する。

4 適用事例 電子ちらし作成システム

eXcart_{TM} の適用事例として、東芝ドキュメント(株)と試作した電子ちらし作成システムについて述べる。

東芝ストアなど電化製品の小売業者は、通常、商品を掲載したちらしを印刷会社に作成依頼している。このちらしの発注は、まとまった枚数の単位で行わねばならないため、品ぞろえに応じて掲載商品を頻繁に変えたり、顧客に応じて結婚祝いや学業祝いなどきめ細かく対応したバージョンを作成するのは難しい。

電子ちらし作成システムは、この問題を解決することを目的として、eXcart_{TM} を使って開発されたシステムであり、小売業者自身が、Web パーツ化されたデザインテンプレートや

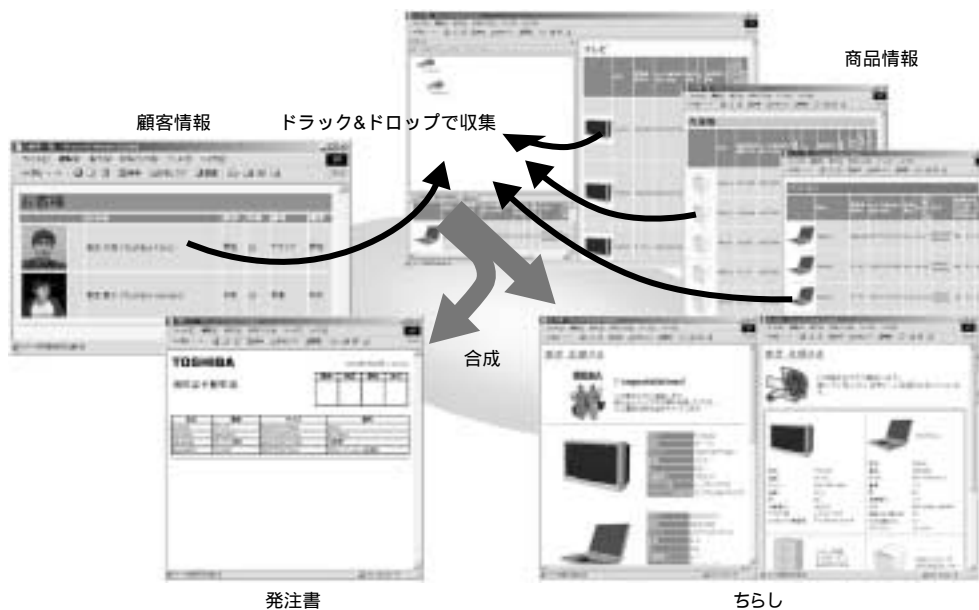


図6. 電子ちらし作成システムの利用方法 Webパーツ化された商品情報や顧客情報をインターネット経由で収集・合成することで、様々なバージョンのちらしや発注書をブラウザ上で簡単に作成できる。

Use model of electronic catalog authoring system

商品データを自由に組み合わせ、Webブラウザ上で様々なバージョンのちらしを手軽に作成することができる。

以下に、具体的な利用方法について説明する(図6)。

まず、ちらしに掲載する製品情報を収集する。この情報は、印刷会社やメーカーのWebサイト上でWebパーツとして公開される。希望する製品を見つけた場合、eXcart_{TM}にドラッグ&ドロップしてWebパーツを蓄積する。

そして、印刷会社が用意したデザインテンプレートを用いて、収集した製品情報を記載したちらしページを合成する。また、ブラウザの印刷機能を使うことで、紙ベースのちらしを作成することができる。

また、小売業者が自分の顧客をWebパーツ化していた場合、顧客情報を合わせて合成することで、特定顧客向けの専用ちらしを作成することもできる。図6の例では、顧客の名前入りのちらしを作成している。このように、Webパーツはタグ構成によって種類が区別されて(この場合では商品情報と顧客情報)、適切な表示が行われるようになっている。

また、メーカーから提供されたデザインテンプレートを使うことで、選択した電化製品の発注書を作成することもできる。このようにこのシステムでは、ちらしの作成から商品発注まで一貫して取り扱うことができる。

5 あとがき

独自のWebページのパーソナライズシステムとして、ユーザーが、ブラウザ上でパーツ化されたWeb情報を組み合わせることのできるeXcart_{TM}について述べた。

eXcart_{TM}の今後の課題としては、異なるスキーマを持つWebパーツの合成を実現すること、また、現在のコピーベースの合成だけでなく、リンクベースの合成を提供することなどが挙げられる。

また、このeXcart_{TM}技術をサーバサイドへ適用することにより、あらかじめ決められた合成ロジックに従って、Webパーツを自動的に加工・合成し配信することを可能とする、コンテンツの付加価値ポータルサービス技術としても展開していきたい。

文献

- (1) Web Services Description Language (WSDL) 1.1
<http://www.w3.org/TR/wsdl>
- (2) EDGE SIDE INCLUDES <esi>
<http://www.edge-delivery.org/index.html>
- (3) SharePoint Portal Server
<http://www.microsoft.com/sharepoint/default.asp>
- (4) eXtensible Markup Language (XML) 1.0
<http://www.w3.org/TR/REC-xml>



浜田 伸一郎 HAMADA Shinichiro

研究開発センター コンピュータ・ネットワークラボラトリー。
XML 応用システムの研究・開発に従事。
情報処理学会会員。
Computer Network Lab.



関 俊文 SEKI Toshitomi, D.Eng.

研究開発センター コンピュータ・ネットワークラボラトリー主任研究員、工博。分散システムに関する研究・開発に従事。
電子情報通信学会、情報処理学会、電気学会会員。
Computer Network Lab.