

## 分散環境下での Web コンテンツ管理システム

中山浩太郎\* 田中成典\*\* 古田均\*\* 中村健二\*\*

インターネットの爆発的な普及に伴い、広く一般に Web コンテンツを管理する必要が発生してきた。しかし、いまだに Web コンテンツ管理では FTP・サーバ・アップロードなどの専門的な知識が必要となり、Web コンテンツ管理の障害となっているケースが非常に多い。そこで、本研究では、HTTP の拡張プロトコルである WebDAV と HTML パーサ技術を利用し、ブラウザのみでコンテンツ管理のできるシステムを開発した。これにより、専門的な知識を必要としないコンテンツ管理手法の提供に成功した。また、本システムの最大の特徴はブラウザのプラグインとして開発している点であり、このことにより、上記利点だけでなく、オーサリングツールや FTP ソフトなどが不必要、Web ページの WYSIWYG 編集が可能、高速なコンテンツ編集が可能などの利点を実現することができた。

### Web contents management system for distributed environment

Kotaro Nakayama, Shigenori Tanaka, Hitoshi Furuta and Kenji Nakamura

Recently, with increment of people that use Internet technology, web management become more important. However, It needs much technical knowledge such as "FTP", "Server", "Upload" and so on. So, we developed a web management system that uses an extension of HTTP "WebDAV" and HTML parser. Most important point of the system is that the system is developed as a plug-in for Web browser. In the system, we can do web management by using only Web browser. The results of our experiment prove the advantages in easy to use, WYSIWYG, speedy renewal and so on.

#### 1. はじめに

近年、わが国では本格的なブロードバンドの普及および e-Japan 戦略により、インターネット利用者が爆発的に増加してきた。インターネットは、平成 14 年末の段階で利用者人口が約 7,000 万人と、実に総人口の 57% が利用する重要な情報インフラへと成長した [1]。インターネットサービスの中でも、Web ページの利用用途は特に幅広く、企業、自治体、個人、商用などの様々な場面で Web サイトを開設する機会が増加してきた。このような状

況の中、誰でも簡単に利用できることを目標とした Web ページ作成ツールが数多く登場し、ユーザの利便性向上に貢献した。しかし、Web サイト開設を必要とするユーザ層を分析すると、①Web に関する知識が必要である、②更新作業を行う場所が限定される、③更新に手間がかかるなどの要因が Web サイト管理の障害となっているケースが未だに多いという結果が得られた [5]。また、現状の Web サイト管理では、Web ブラウザ・オーサリングツール (ビルダー)・FTP ソフトなどの複数のアプリケーションを操作してコンテンツ管理を行う必要があり、少量の修正であっても

\*関西大学大学院 \*\*関西大学総合情報学部

\* Kansai University Graduate School

\*\* Faculty of Informatics, Kansai University

時間がかかるという問題がある。

そこで、本研究では、Web ブラウザのみでコンテンツ管理ができるシステムの構築を目指す。

## 2. 要素技術の概要

本章では、本研究の基盤となる技術について解説する。

### 2. 1 WebDAV

WebDAV は、Web-based Distributed Authoring and Versioning の略であり、閲覧専用のプロトコルであった HTTP を拡張し、編集 (Authoring) やリビジョン管理 (Versioning) を可能にした規格である[3]。WebDAV は、未だ仕様が確定していない部分が多く、発展途上の規格であるが、Apache プロジェクト[7]の「Apache HTTP Server」や Apache Jakarta プロジェクト[8]の Servlet コンテナ「Tomcat」では既に実装されるなど、次世代のコンテンツ管理規格として注目を集めている。しかし、サーバサイド技術が発展する一方で、WebDAV に対応したクライアントアプリケーションは未だ少ないのが現状である。

### 2. 2 HTML パーサ

通常、HTML や XML などのタグ文書を扱うアプリケーションを開発する場合には、タグ文書は直接扱わず、構文解析するパーサを利用するのが一般的である。特に、最近では XML 文書を構文解析するための「XML パーサ」や「XML DOM」が数多く開発されているが、HTML にも、構文解析をするための HTML パーサが存在する。HTML パーサを利用することにより、プログラムから HTML の構造や内容を抽出することや、

編集することができる。例えば、HTML パーサを利用すれば、HTML 内の A タグの href 属性だけを抽出するなどの処理が非常に容易に実現できるため、ページ内のデッドリンクを調査するデッドリンクチェッカーや Web サイトの自動巡回ロボット、アクセシビリティチェッカーなどのプログラムに HTML パーサの技術は利用されている。HTML パーサの概念を図 1 に示す。

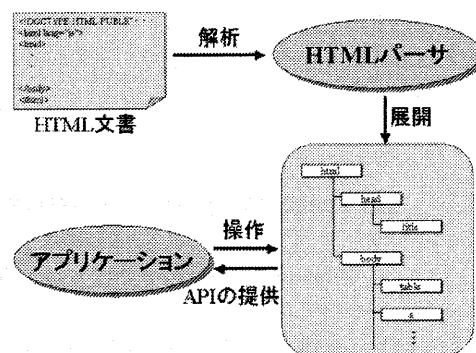


図 1 HTML パーサ概要

## 3. 研究の目的と対象

従来の Web サイト構築では、Web ブラウザの他に FTP ソフトやオーサリングツールなどのソフトを使いこなす必要であった。このことにより、Web サイト管理者は、ブラウザで閲覧しているページがファイルシステム上では実際にはどこにあるのかを常に把握しておく必要があった。例えば、Web サイトの更新時には更新対象のファイルがどこに存在しているかを検索する作業必要となり、アップロードの際にはファイルシステムのどこに配置するかを指定しなければならなかった。これは、サーバサイドの知識や Web の仕組みを知らない者にとっては Web サイト更新の大きな障害となっている。また、Web サイト管

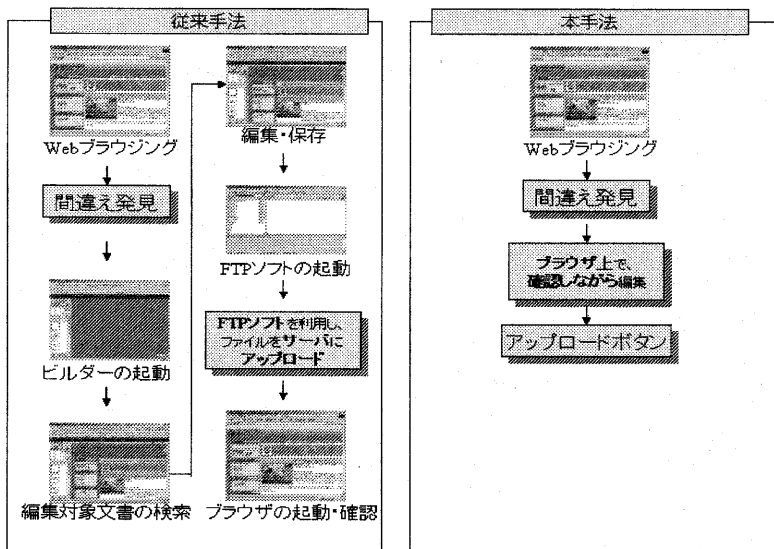


図2 従来手法との作業

理の専門家であってもこれらの作業には大きな負担がかかる。

そこで、本研究では、Web更新作業をWebブラウザのみで可能にする更新手法を提案する。本手法を利用することにより、閲覧・間違い発見・編集・アップロードまでの流れをWebブラウザのみで円滑に行うことができるため、従来のWebサイト更新で必要であった多くの作業工程を省略できる。従来の手法と本手法の作業工程比較を図2に示す。

#### 4. システムの概要

本研究では、WebDAVとHTMLパーサを利用し、コンテンツ管理システムを開発した。本システム構成を図4に示す。

本システムは、WYSIWYGインタフェース「WebON」とWebDAVを利用したアップロード機能の2つのモジュールから構成される。各モジュールにおける処理の流れを以下に示す。

#### 4. 1 WYSIWYGインタフェース「WebON」

「WebON」は、Webブラウザ上からの編集作業を可能にするインタフェースである。WebONは、筆者らが開発したWebページのオーサリングシステムである。最大の特徴は、Webブラウザのプラグインとして開発している点である。このことにより、Webブラウザ上でWebページのWYSIWYG編集を可能としている。WebONの編集画面を図3に示す。



図3 WebON 編集画面

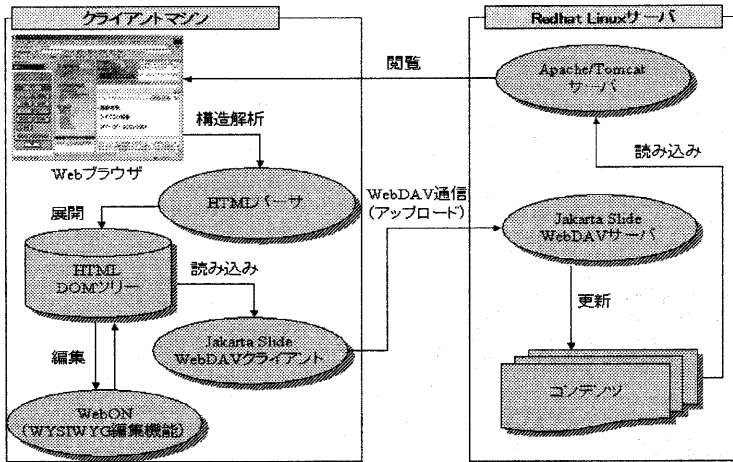


図3 システム構成

WebONには、文章の編集機能だけでなく、画像の追加・削除機能、フォントの設定機能、リンクの指定機能、表（テーブル）の生成機能、位置ぞろえ機能、タイトル設定機能など、Web ページ編集に必要な一通りの機能を実装している。これらの機能はすべてHTMLパーサ技術を利用して実現した。HTMLパーサは無数にあるが、本研究では、Microsoft社が開発したMSHTMLをHTMLパーサとして採用した。MSHTMLは、Internet Explorerに組み込まれているHTMLパーサであり、ページのレンダリングの際にも利用されているため、安定性が高くパージング速度が速いのが特徴である。WebONにおける各機能の処理手順を以下に示す。

(1) HTML DOM の作成

HTMLは、<HTML>タグをルートノードとするツリー形式のデータ構造である。本システムでは、HTMLをパーサによりパージングし、DOMツリーを獲得し、プログラムから処理しやすい形式に変換する。

(2) 選択ノードの特定

DOMツリー内のノードから、ブラウザ上のカーソル位置に対応するノードを検索し、特定する。

(3) 属性の獲得

ダイアログを表示し、ユーザから属性情報を獲得する。例えば、リンクを付与する場合はhref属性の値をURLとして獲得し、画像追加の場合はsrc属性の値をURLとして獲得する。

(4) ノード追加

HTML DOMツリーに新しいノードを追加する。例えば、Webブラウザで選択中に箇所に対してリンクを付与する場合は、手順(1)で獲得したノードを<A>タグと</A>タグで囲む必要がある。このためには、①式に<A>タグを示すノードを作成、②選択中のノードのコピーを子ノードとして登録、③選択中のノードを削除するという3段階のプロセスを実行し、新しいノードを作成する。

4. 2 WebDAV を利用したアップロード機能

本研究では、オープンソースプロジェクトの Apache Software Foundation が開発している WebDAV クライアントライブラリ「Jakarta Slide」[8]を利用してクライアントプログラムを作成し、ブラウザからのアップロード機能を実現した。本クライアントプログラムにおける WebDAV 通信でのアップロード手順を以下に示す。

(1) HTML ソース獲得

編集中の HTML ソースをブラウザから獲得し、一時ファイルとして保存する。

(2) HTTP Session の確立

HTTP の Session を開始することにより、サーバアプリケーションとの通信が可能な状態となる。HTTP の Session を開始するには、HttpClient クラスの startSession メソッドを利用する。また、HTTP Session を確立する際には、ユーザ認証を行い、セキュリティの向上を実現した。本システムでは、HTTP リクエスト内にユーザ ID とパスワードを付加することにより、ユーザ認証を行った。HTTP リクエスト内にユーザ ID とパスワードを挿入した場合の URL の構文を以下に示す。

```
http://<USER_ID>:<USER_PASS>@
<SERVER_NAME>/

<USER_ID>      : ユーザ ID
<USER_PASS>    : ユーザパスワード
<SERVER_NAME> : サーバ名もしくは IP アドレス
```

(3) ファイル Lock

ファイルのアップロードの際には、アップロード中に他者が同時更新を行わないように Lock 処理を行う。Lock 処理には、WebDAV の lock コマンドを利用する。

(4) ファイルアップロード

ブラウザから取得した HTML ソースをサーバに送信する。アップロード処理には、WebDAV の Put コマンドを利用する。

(5) ファイル Unlock

ファイルのアップロード処理が終了したタイミングでファイルにかけていた Lock を解除するために、Unlock 処理を行う。Unlock 処理には、WebDAV の unlock コマンドを利用する。

(6) HTTP Session の切断

通信が終了した場合は、HTTP の Session を切断する。Session の切断には、HttpClient クラスの endSession メソッドを利用した。

5. 実験計画

本章では、本論文で提案した方法と、従来の Web サイト構築の手法を比較し、新手法の有用性を実証する。本実験に用いたサーバ環境を表 1 に示す。

表 1 環境一覧表

項目	環境
OS	Red Hat Linux 9.0
Web サーバ	Apache
J2EE コンテナ	Tomcat4.1
WebDAV サーバ	Slide1.0
WebDAV クライアント	Slide1.0
JDK	J2SDK SE 1.4
Web ブラウザ	Internet Explorer 6.0
クライアント開発環境	Visual C++ 6.0

本実験では、Web サイト管理者と Web の知識を持たない者を被験者として採用した。従来の手法としては、オーサリングツール、FTP ソフトと Web ブラウザを利用してコンテンツ管理を行った。実験内容としては、Web コンテンツを複数用意し、その更新時間を比較

した。

## 6. 実験結果の評価

本章では、実際に被験者に本システムを利用させた。

まず、Web サイト管理者を被験者として従来手法と本手法の比較を行った。その結果、従来のオーサリングツールを利用する手法では、①Web ブラウジング、②修正箇所の発見、③オーサリングツールの起動、④編集対象文書の検索、⑤編集・保存、⑥FTP ソフトの起動、⑦サーバの指定、⑧アップロード先へ移動、⑨ローカルコンテンツの検索、⑩アップロード、⑪Web ブラウザでの確認、という11もの工程を経て Web サイト管理を行っていることがわかった。一方、本研究での手法を利用することにより、工程③～④および⑥～⑪までの工程を省くことができる。このことにより、作業時間が著しく減少したことを確認することができた。これは、Web サイトの規模が大きくなればなるほどその傾向が強かった。

次に、Web の知識を持たない者を被験者として従来手法と本手法の比較を行った。その結果、どちらの手法も編集作業までの時間はそれほど変化がなかったが、アップロード作業の段階で著しい差が生じた。これは、従来方法の場合は、Web の知識がなく、概念の習得に大きな時間を費やしたのに対し、新手法が Web の知識を必要としなかったことに起因していた。

## 7. おわりに

本論文では、HTML パーサ技術と WebDAV 技術を利用した新しい Web サイト管理手法を提案・実装した。特に、ブラウザ上での

WYSIWYG 編集を可能とする「WebON」と WebDAV によるアップロード機能を開発することにより、Web ブラウザのみでの Web ページ更新が可能な手法を実現した。また、その評価を行った結果、従来の Web サイト構築に比べ、専門的知識を必要としないことと多くの作業工程を省くことを確認した。

次に、本手法の今後の課題を以下に列挙する。

- バージョン管理機能

WebDAV には、バージョン管理のための仕様が策定されている。しかし、未だバージョン管理機能を実装したサーバは非常に少なく、機能の開発には時期尚早である。しかし、Web サイトを複数人で管理する際には必須の機能であるため、今後のサーバでの実装の普及に伴い、開発を進めていく必要がある。

- サーバサイドプログラミング機能

最近では、データベースを利用した動的な Web サイトが急増してきた。動的な Web サイトを構築するためには、Java サーブレット・JSP・ASP・PHP・CGI などのサーバサイドプログラミングが必要である。一方、Web アプリケーションでは、データベースへのデータ参照・追加・編集・削除など定型的な処理が多く、汎用的なインタフェースを作成することができる。そのため、本システムを拡張し、Java などのソースコード自動生成機能を付与することにより、動的な Web サイトの構築にも対応することができる。

最後に、本研究は、日刊工業新聞主催キャンパスベンチャーグランプリにおいて、情報通信部門最優秀賞を受賞している。受賞に際

して有益なコメントと機会を与えてくださった  
審査員の方々に心から感謝の意を表する。

#### 参考文献

- [1] 総務省：平成 15 年度版情報通信白書
- [2] World Wide Web Consortium：  
DOM (Document Object Model) specification,  
<http://www.w3.org/DOM/>
- [3] Y. Goland, Microsoft, E. Whitehead, UC  
Irvine, A. Faizi, Netscape, S. Carter, Novell,  
D. Jensen, : HTTP Extensions for Distributed  
Authoring – WEBDAV, RFC (Request for  
Comments ) , February 1999 ,  
<http://www.ietf.org/rfc/rfc2518.txt>
- [4] Greg Stein : WebDAV Resources (Web page),  
<http://www.webdav.org/>
- [5] 田中成典, 古田均, 中山浩太郎ほか: WYSIWYG  
機能付き汎用 Web サイト構築支援システムの研  
究開発, 情報処理学会第 64 回全国大会, 講演論  
文集 p4-337
- [6] 田中成典, 中山浩太郎, 江川学, 杉町敏之, 野中  
一希, 福島賢治, 渡部菜月, 田中誠人, 山下智子:  
ステップアップ Visual C# .NET 入門, 工学社, 2002  
年 7 月
- [7] Apache Software Foundation : The Apache  
Software Foundation , <http://www.apache.org/>
- [8] Apache Software Foundation : The Apache Jakarta  
Project, <http://jakarta.apache.org/>