

WWW を用いた文書管理システム (OHMS) の実装

3D-4

田中 智之 今郷 詔 岩崎 雅二郎 西村 美苗
 (株) リコー 情報通信研究所

1 はじめに

近年、コンピュータのネットワークへの接続が一般的になり、ネットワーク上の情報にどこからでもアクセスできるようになってきた。ネットワークを利用して、オフィスに溢れている様々な情報をうまく共有するオフィス文書管理システムを構築することによって、情報流通の手間・コストを大きく削減するだけでなく、情報の流通路を組織の木構造から独立させ、情報の保有者と潜在的需要者とを直接結ぶことが可能になる。

そのために我々は WWW をベースにして OHMS (Office Hyperdocument Management System) を開発した [1]。

本稿では、ブラウザとデータベース部とを仲介する部分を中心に、OHMS の実装について述べる。

管理機能や、文書に対する全文検索機能を備えている。

イメージサーバ 紙文書のスキャニングおよび OCR 処理を行なう。

OHMS ゲートウェイ OHMS ブラウザからは WWW サーバとしてアクセスされ、ODR 内の情報を HTTP に合わせて OHMS ブラウザに送る。また OHMS ブラウザのコマンドボタンが押された時に、対応する ODR 操作も行なう。

OHMS ゲートウェイインタフェース (OGI) は、HTTP サーバと ODR サーバとの仲介役を果たしており、個々の機能を実行する CGI スクリプトの集合として実装されている。

2 OHMS の構成

OHMS は図 1 に示すように、ユーザが使用するブラウザの部分と、文書・リンクを管理する OHMS サーバの 2 つの部分からなる。

ブラウザは Mosaic などの WWW クライアントそのものである。ただし、通常の WWW クライアントにはファイルをサーバに登録する機能が欠けているので、ファイル登録コマンドをインストールしておく必要がある。

一方、OHMS サーバは次の 3 つの部分から構成される。

ODR (Office Document Repository) サーバ 文書データ・リンクデータ・検印データを記憶するデータベースである。OODBMS 上に構築しており、それぞれの文書・リンクに対するアクセス権

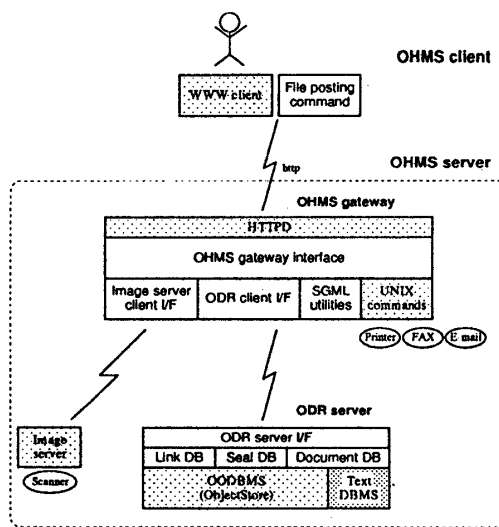


図 1: OHMS の構成

2.1 OHMS ゲートウェイ

OHMS のユーザーインターフェイスは WWW クライアントを利用しているため、そのままでは文書进行操作することができない。文書管理システムとしては文

Implementation of the Office Hyperdocument Management System (OHMS) using WWW
 Tomoyuki TANAKA, Satoshi IMAGO,
 Masajirou IWASAKI, Mina NISHIMURA
 Information and Communication R&D Center
 RICOH Co.,Ltd.

書の登録、更新、削除、検索などの様々な操作をクライアントから行う必要がある。OHMS ではサーバにこれらの機能を CGI スクリプトとして付加し、文書を表示する時点でこれらの機能を提供する操作ボタンを文書へ付加することによりクライアント側からの操作を可能としている。

2.1.1 文書表示

文書データベースには一つの文書に対して、タイトルや文書タイプ等の書誌情報、リンク情報が関連して納められている。OGI は文書表示のリクエストを受けると文書タイプに応じて以下のような変換を行いブラウザへ返す。

HTML/プレインテキスト HTML テキストに変換。

SGML テキスト HTML に変換する。ただし、変換には時間がかかるため一度変換したものは HTTP サーバ上にキャッシュとして保存し、以降のアクセスにはキャッシュを返す。

イメージ・その他 データ本体へのリンクを記述した HTML テキストを返す。これはテキストデータ以外をそのまま表示しては後述する操作ボタンを付加することができないからである。

OHMS 文書間のリンクが設定されている場合には、本文以降に HTML のリンクとして付加する。さらに、OHMS の操作ボタン (OGI スクリプトへのリンクを記述してある) を文書の先頭と末尾に付加する。操作ボタンについては、どの文書に対する操作を行うかを OGI のスクリプトに知らせるため、文書の ID を OGI スクリプトへのリンクの URL の一部として付加している。

2.1.2 文書ファイル登録

クライアントマシン上の文書ファイルを OHMS に登録する機能である。OHMS のユーザインターフェイスとして使用する WWW クライアントにはクライアントマシン上のファイルをサーバに送る機能は無いため、専用のプログラムを用意し、それを用いて文書の登録を行なう。したがって、OHMS に文書を登録する手順は次のようになる。

1. 表示文書上の「新規登録」ボタンを押すと登録画面に切り換わるので、タイトルと文書タイプを設定する。

2. 「ファイル選択」ボタンを押すと、これらの情報がコンテンツタイプ「application/x-ohmspost」なるデータとしてブラウザに送られる。
3. ブラウザは送られてきたデータタイプに対応するヘルパーアプリケーションとして OHMS の文書登録プログラムを起動する。

文書登録プログラムはユーザが指定したファイルを WWW サーバへ HTTP プロトコルで接続し POST メソッド [2] を用いて送信する。

2.1.3 文書画像登録

文書画像登録のリクエストを受けると、OGI はイメージサーバにスキャンおよび OCR 処理を依頼する。OCR 結果は、先頭部分数行をユーザの文書タイトル入力を補助するために画面に表示した後、ODR に渡される。ODR では OCR 結果自体を保存する訳ではなく、全文検索用索引を作るための材料としてのみ使用する。ユーザが実際に検索キーとして使用する文字列は、文書中に何度も出現していたり、見出しやタイトルのように大きな文字でプリントされている傾向が高いため、品質の悪い画像でも検索が可能な場合が多い。

3 おわりに

WWW の枠組みを利用して文書管理システム OHMS を作成した。WWW を利用することで、様々なプラットフォームからシステムを使用でき、ユーザに解かりやすいインタフェースを提供することができた。

今後は一層の高速化と操作性の向上を目指していく。

参考文献

- [1] 岩崎雅二郎 他, WWW を用いた文書管理システム (OHMS) の概要, 情報処理学会第 51 回全国大会論文集, 1995.
- [2] T.Berners-Lee, R.T.Fielding and H.F.Nielsen, Hypertext Transfer Protocol - HTTP/1.0, Network Working Group Internet Draft, December 19, 1994.