

データ・書式の共有による書類作成支援システムの構築

安田 愛[†] 藤澤 公也[†]

東京工科大学大学院 バイオ・情報メディア研究科[†]

はじめに

会社や学校では、何度も同じ類の書類を利用することがある。また、書類も紙で配布されるのではなく、電子ファイルで配布することも増えてきている。しかし、提出は印刷した紙を提出する場合が多い。書類の処理も含めたすべての仕組みを電子化するシステムも提案されている[1]が、この場合はすべての利用者は新たなツールの使い方を覚える必要がある。

これらの問題点を解決するために、書類作成支援システムの開発を行っている[2]。このシステムでは、利用する書類及び処理手順が変わらないことを目的として、設計・実装を行った。

入力データの共有を可能とするために、データベースを利用して入力データと書類テンプレートを管理する。紙に手書きで記述をし、書類を作成する方法と、このシステムを利用して入力データを差し込むことで作成する方法が共存できるようになる。このシステムを利用して書類を作成した場合、データが入力された PDF ファイルが出力される。これを印刷することで、手書きで作成した書類と同様に、紙媒体で提出が行える。また、書類の見た目の変化が起こら

ないように、書類を画像化し、背景とした PDF ファイルの入力部分にフォームを埋め込む。これを書類テンプレートとして利用する。このシステムは、「書類テンプレート登録」「入力データ登録」「登録済みデータの書類テンプレートへの差し込み」の機能からできている(図 1)。

書類作成支援システムの問題点と解決策

入力データの共有を行い、書類の作成を行うシステムとして、書類作成支援システムの 3 つの機能を実装していたが、このシステムには 2 つの問題点が残されていた。1 つ目の問題は、書類テンプレートを作成する際に、既存の市販ソフトウェアである「やさしく PDF へ文字入力 PRO v. 6.0」[3]を利用している点である。すなわちこのソフトウェアをインストールしているコンピュータを利用しないと書類テンプレートを作成できない。2 つ目の問題は、手動での入力を行っている点であり、入力間違えが発生しやすい。例えば、消耗品を購入する書類を作成する場合には、商品の品番や品名を記述しなくてはならない。型番等は英数字の羅列であることが多く、1 文字間違えると別の商品になってしまう場合がある。

本研究では、これらの問題点を解決するために以下の 2 つの機能(図 2)の実装を行った。(1) 書類テンプレートの作成を Web ブラウザ上から行えるようにする。(2) Web 上の情報を出力しデータとして入力できるようにする。

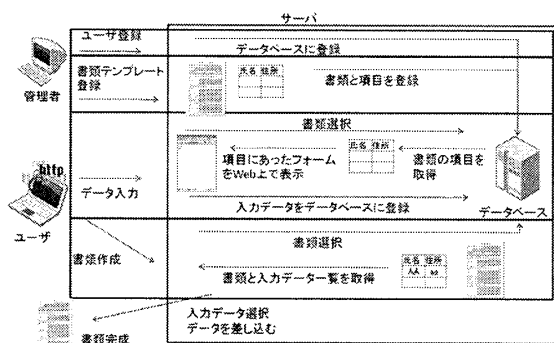


図 1 ベースとなる書類作成支援システム

The documents generating support system with sharing templates and data
[†]Ai YASUDA, Kimiya FUJISAWA
 Graduate School of Bionics, Computer and Media Sciences, Tokyo University of Technology

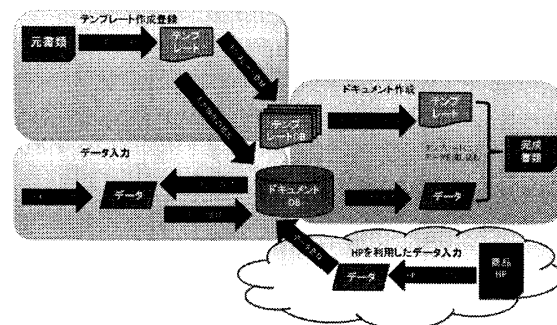


図 2 追加する 2 つの機能の位置づけ

Web ブラウザ上からの書類テンプレート作成

本システムでは、入力すべき項目にフォームを付加した PDF ファイルの書類を書類テンプレートとして登録する。ここでは、既存の市販ソフトウェアを使うのではなく、書類テンプレートを Web ブラウザ上から作成を行えるようにした。

この機能の UI の実装には、JavaScript を用いた。ブラウザ上での描画には wz_jsgraphics.js 用い、ブラウザ上での範囲指定には JavaScript Image Cropper UI を用いている。また、実際の書類テンプレートの PDF ファイルの生成は、サーバ側で行っており、iText を用いている。クライアントとサーバのやりとりは、prototype.js による Ajax で行っている。

サーバでは元となる PDF を画像化し、HTML に埋め込んでいる。この画像に対して、どこにどのようなフォームを埋め込んでいくかを指定していく。データ入力エリアを範囲選択により行い、同時にフォーム名・データを差し込む際のデータの配置・フォントサイズ等を属性として付加し、これらの情報をサーバへ送信している。サーバ側では、元の PDF の書類に対して、入力フォームを付加した PDF を生成する。フォームを付加する際には、送られていた属性により、フォントサイズ等の設定を行う。ここで作成された PDF は、既存の書類作成支援システムでの書類テンプレートとして利用する。

Web 上の情報の利用

本システムでは、書類に記述すべきデータをインターネット上にある情報から抽出できるようにする。これによって、手動で入力することでの間違い等が軽減できる。

商品情報等を閲覧可能な Web サイトでは、WebAPI が公開されているサイトと公開されていないサイトが存在する。WebAPI を利用できる Web サイトに関しては、WebAPI を利用して必要な情報を XML 形式で抽出する。この Web サイト用の XSL を用意しておき、抽出した XML に利用することで、本システム用の XML に変換する。また、WebAPI が公開されていない Web サイトでは、HTML を XHTML 化し、対象の Web サイト用に用意された抽出用の XSL を利用することで、必要な情報が含まれた本システム用の XML を生成する。本システム用の XML は、タグ名と登録を行うテーブルの項目名が対応している。このタグ名と抽出されたデータを利用し、データベースへの登録が行える。

本研究では、2 つの Web サイト上からデータの

抽出を試みた。1 つ目は WebAPI が用意されている Web サイトである Amazon であり、2 つ目は WebAPI がない Web サイトである TSUTAYA online shop である。Amazon と TSUTAYA online shop でそれぞれの商品詳細ページの URL を利用することでその商品の情報を抽出する。Amazon の Web ページからのデータの抽出は、公開されている WebAPI である AmazonWeb サービスを利用して行った。TSUTAYA online shop の Web ページには WebAPI が公開されていないので、商品のページの HTML を XHTML 化して行った。

2 つの Web サイト上からそれぞれ本システム用の XML 形式で情報を抽出し、本システムのデータベースに登録できた。また、登録したデータを利用して書類も作成できた。他の Web サイトでも適切な XSL を用意することで、必要なデータの登録が容易になる。

おわりに

提案するシステムによって、既存の市販ソフトウェアやツールをインストールすることなく書類作成支援が行える。入力データの共有や Web 上の情報を利用してデータの登録が行え、登録されているデータを利用することで、書類作成の手間が軽減された。

今後の課題としては、書類テンプレート作成時に、一度作成したフォーム部分の変更や、既存のシステムに組み込むためのフォーム名等のルールに適應できるようにすることがあげられる。本システムで利用するルールに沿った名前付けは、書類テンプレート作成時に、入力する側に任されている。既存のシステムに組み込むためには、ある程度の自動化を行う必要がある。またデータ入力時に、他の Web サイト上からも抽出可能とする必要がある。Web ブラウザ上から簡易的に他の Web サイト用の XSL の作成が行えるようにする必要もある。

今後、今回浮かびあがった問題点を改善し、ユーザが利用しやすいものを実装していく。

参考文献

- [1] 森廣, 宮崎, 鶴, "Web 技術を利用したデータベースシステムによる業務効率化", 大分工業高等専門学校紀要 第44号, PP.37-42, 2006
- [2] 安田愛・藤澤公也, "書類入力支援システムの設計と実装", 第71回情報処理学会全国大会 1Q-1 (2009.3)
- [3] メディアドライブ株式会社「やさしく PDF へ文字入力 PRO v.6.0」: <http://mediadrive.jp/products/ypdf/index.html>