

WEB 环境中 WORD 文档生成技术讨论

Discussion on Word Creation Methods in WEB

方路平 章文彬 (浙江工业大学软件学院 310032)

摘要:随着互联网的迅速发展,基于 WEB 的交互方式已经变得越来越普及。因为操作的方便,人们已习惯使用 IE 或者其他浏览器通过访问网页的形式与信息管理服务进行信息交互。同时 Word 也是很多用户经常采用的文字处理软件,很多时候我们需要将网页中的信息转换成 Word 格式来处理。本文将综合讨论在 WEB 环境中生成 WORD 的一些手段,以及它们各自的特点。

关键词:网页 Word ASP VBA

1 引言

随着因特网的迅速发展,基于 WEB 的交互方式越来越成为主流。因为其方便,快捷等优点深受用户喜爱。而 Microsoft 公司的 Word 以其文字处理能力,灵活多变的排版方法,所见即所得以及友好的用户界面等优点使其成为目前大多数用户都采用的主流文字处理软件。例如我们从 WEB 页面中填写了诸如个人信息,所申请的项目等等内容并且提交了,而同时又需要一份与刚才提交的内容一样的书面材料,怎么办呢? 本文就来重点讨论一下基于目前的技术在 WEB 环境中有关生成 WORD 的一些方法以及比较各自的特点。

2 以网页方式直接生成 WORD 文档

Office2000 或者以上的版本支持 HTML 作为内部文件格式,因此可以利用 WEB 文档与 WORD 文档的密切关系将 WEB 中的一些内容用 WORD 的形式保存下来。采用这种方式,我们无需在服务器中预先生成 WORD 文档,但当用户通过 IE 浏览器访问网页的时候,将自动生成相应的 WORD 文档。用户可以选择直接打开或者保存该文档在用户本地机上。相关代码如下:

```
<%
```

此处为数据库连接操作代码(省略)

```
Response.ContentType = "application/msword" 通知  
浏览器用 word 格式显示文档内容
```

```
response.AddHeader "content - disposition", " inline;  
filename = example1.doc" 文档命名
```

```
% >
```

```
<html > <head >
```

```
<title > Table_creation </title >
```

```
</head >
```

```
<body bgcolor = " #FFFFFF" >
```

```
<center > <h1 > Table </h1 > </center >
```

```
<table border = 1 cellspacing = 0 cellpadding = 0 bordercolor = #000000 align = center >
```

```
<tr >
```

```
<td > 用户名 </td >
```

```
<td > 真实姓名 </td >
```

```
<td > 性别 </td >
```

```
</tr >
```

```
<%
```

```
rs.movefirst 移到第一个记录
```

```
do while not rs.eof
```

```
% >
```

```
<tr >
```

```
<td > <% = rs(" Te_Name" ) % > </td >
```

```
<td > <% = rs(" Real_Name" ) % > </td >
```

```
<td > <% = rs(" Sex" ) % > </td >
```

```
</tr >
```

```
<%
```

```
rs.movenext 移到下一个记录
```

```
loop
```

```
关闭记录集代码(省略)
```

% >

```
</table> </body> </html>
```

当用户浏览该网页的时候,由于有了 Response.ContentType = "application/msword" 这条语句,浏览器会提示直接打开或者以文件的形式保存到客户本地机上。如图 1 所示。

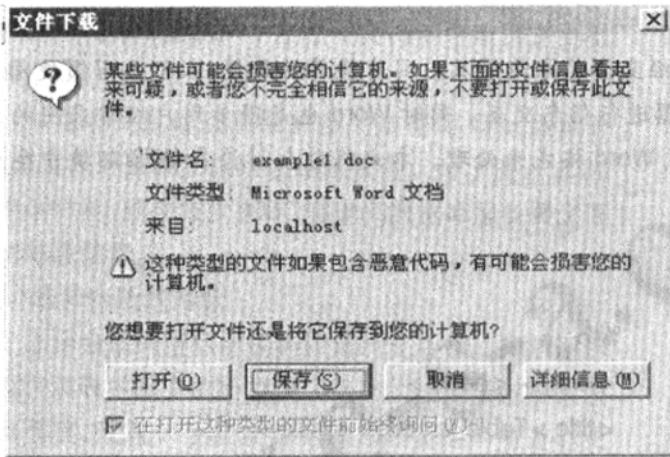


图 1 浏览器提示信息

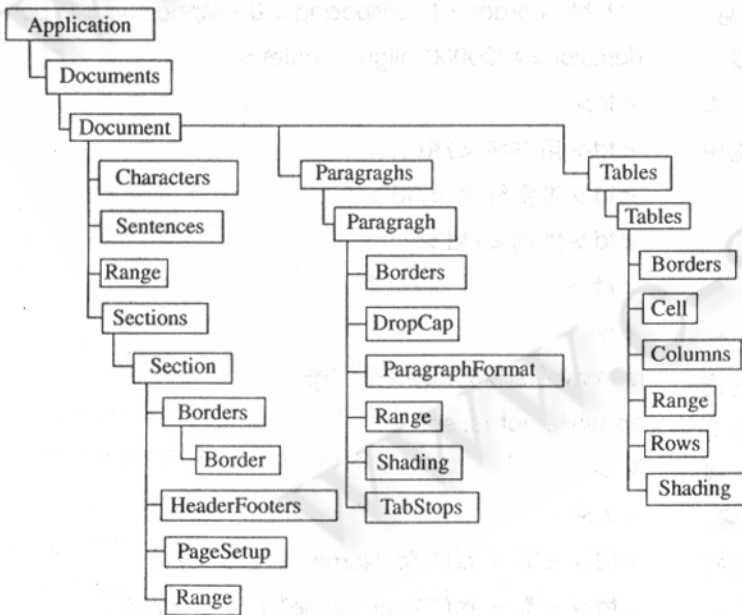


图 2 WORD 部分对象结构图

按照上面的方法虽可以生成 WORD,但是生成的 WORD 没有页眉页脚或者页码等等 WORD 应有的效果,也就是不能很好地控制所生成 WORD 文档的样式。

下面的方法可以很好地解决这个问题。

3 使用 VBScript 生成 Word 文档

在 ASP 中默认脚本是 VBScript,所以可以利用 VBA 编程进行网页到 Word 文档的转换。VBA (Visual Basic for Application) 是 Microsoft Office 系列的内置编程语言,是应用程序开发语言 VB (Visual Basic) 的子集。VBA 的独特之处在于它由应用程序控制,反过来,它又可以增强该应用程序的功能。

要通过网页生成对应的 Word 文档,实现这一要求的关键在于 ASP 中如何正确使用 MS Office COM 对象,首先必须在客户机上正确安装 Word 应用程序,其次必须理解 MS Word 对层次关系结构,并利用 VBA 编程,具体内容可在 MS Office 帮助中“与编程有关的信息”一节中查阅。

WORD 的部分对象结构如图 2 所示。

在 Word 中,几乎所有的操作都要调用 Document 对象本身或其内容。当使用 VBA 操作 Word 时,Document 对象表示一个打开的文档,而且所有的 Document 对象都是 Application 对象的 Documents 集合的成员。文档是一个由字符、单词、句子和段落组成的集合,字符组成单词,单词组成句子,句子组成段落等等。因此,每一个 Document 对象都具有 Characters、Words、Sentences 和 Paragraphs 四个集合。此外,每个文档具有一个包含一个或多个节的 Sections 集合,每一个节都有一个包含该节页眉和页脚的 HeadersFooters 集合。经常用到的 Word 对象及方法主要有:

- (1) Paragraphs 集合对象,即文本中段落对象的集合。用 Paragraphs (Index) 可对文本中的指定段落进行必要的编辑或访问。
- (2) Range 对象,该对象代表文档中的一个范围。每一个 Range 对象由一起始和一终止字符位置定义。VBA 中用于返回 Range 对象的 Range 方法的语法为: expression. Range (Start, End), 其中 expression 表达式是必须的,它返回一个 Document 对象。
- (3) Sections 集合对象,该对象包含了选定内容、范围或文档中的所有节。
- (4) Selection 对象,返回文档中当前选定的文本,

在需要编写代码对用户突出显示的文本进行操作时可以使用该对象。

(5) `InsertParagraphBefore` 方法,在指定的所选内容或区域前插入一个新的段落。

(6) `InsertParagraphAfter` 方法,在指定的所选内容区域之后插入一个新的段落。

下面我们讨论它们的使用。假设我们要把图 3 网页中的表格生成 word,图 3 中表格内的数据来自数据库。

学生信息	
姓名	学号
xjb	123456

生成

图 3 所要生成的 WORD 内容

代码如下:

```
<script language="vbscript">
```

```
Sub buildDoc()
```

定义数组变量,存放表格中的数据(开始数组没有具体大小,以后根据网页中的表格的实际大小通过 `redim` 确定数组大小)

```
with objWordDoc. Application. ActiveDocument. sections
```

```
(1)
.headers(1).range.text = "word example" '定义页眉
```

```
.footers(1).range.text = "student information" '定义页脚
```

```
end with
```

```
Dim theArray()
```

set table = document.all.data 把 html 文档中的表格的结构和数据赋值给 table

row = table.rows.length row 为 table 的行数

column = table.rows(1).cells.length colnum 为 table 的列数

```
Set objWordDoc = CreateObject("Word.Document") '创建一个 Word.Document 的对象
```

```
objWordDoc. Application. visible = true
```

```
redim theArray(column,row) 根据网页中的表格的实
```

际大小重新确定数组 theArray 的大小

以下两层循环将 html 文档的表格中的纯文本数据赋值给数组。

VB 数组从 1 开始而且为先行后列。而表格数组从 0 开始且先行后列

```
for i=0 to row - 1
```

```
for j=0 to column - 1
```

```
theArray(j+1,i+1) = table.rows(i).cells(j).innerTEXT
```

```
next
```

```
next
```

显示表格标题(以下代码略。)

在上面的网页中放置了一个按钮,可以先浏览网页,网页中的数据可以来自于模板文件或者是数据库。如果有必要生成 Word 文档的话,只需要点击一下那个按钮就可以在本地生成相应的 Word 文档。点击生成之后,系统会提示“在此页上的 ActiveX 控件和本页上的其他部分的交互可能不安全。你想允许这种交互吗?”。虽然可以选择 YES,也可以将 IE 的安全级别设为低,或在中级安全级别下将那个 ActiveX 控件设为启用。但这总是会让用户感觉不妥当,下面提供另外一种方法。

4 在服务器端生成 WORD 文档

以上介绍的几种方法都是在客户端生成 WORD 文档,需要使用 ASP、VBScript 等,而当某些 WEB 服务器没有配置这样的编译器的时候,上述方法就失效了。

下面讨论一种方法,在服务器端生成 Word 文档,客户端所要做的就是点击该链接,然后下载就可以了。当然采用这种方式时,服务器端需要安装 OFFICE 程序。

首先在服务器端启动一个监视进程。该进程从数据库某一特定表中的特定字段读取状态,该状态标识要求的 Word 文档是否已自动生成。所要生成的 WORD 文档的模板可以从某一文件中读取。该进程从数据库中读取相关数据,根据模板将内容填到相应位置。可以用 Delphi,VC,VB,C 等编程语言实现。如图 4 所示。

其中□表示数据库:用于存放从客户机发送过来的数据;○表示其中的监视进程:以一定的时间间隔从数据库中读取状态信息,采集数据表中的有关数据生

成 WORD 文档。□表示 WORD 文档的集合。

下面以 DELPHI 为例介绍如何生成 WORD 文档。

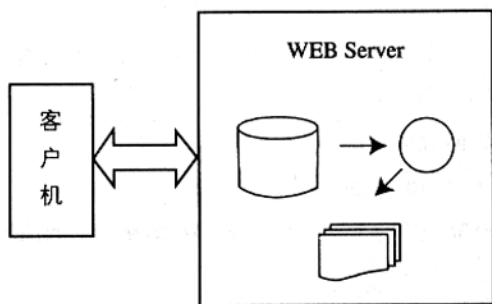


图 4 生成 WORD 文档过程模型

监视程序以某一时间间隔查询数据库,该任务可由一定时器完成。一旦发现某些数据需要转换成 Word 文档,则启动一个新的线程来处理。与数据库连接的方式较多,在此就不多赘述了。数据提取之后,将其填入一个 WORD 文档中。使用 Delphi 提供的 Servers 控件调用 Office,它使我们很容易把 OFFICE 中的应用程序(比如说 WORD)当作一个 COM 应用服务器进行控制。

先介绍一下其中一些相应的控件:

- (1) TwordDocument: 建立 Word 文件对象
- (2) TwordApplication: 启动 Word 并建立同 Word 的连接
- (3) TwordFont: 设置 Word 文件的字体
- (4) TwordParagraphFormat: 设置 Word 文件中的段落格式

第一步:通过 TwordApplication 实现 WORD 的启动。

Try

```
Wordapplication. Connect;
```

Except

```
MessageDlg('Word may not be installed', mtError, [mbOk], 0);
```

```
Abort;
```

```
End;
```

第二步:通过 TwordApplication 实现一个新文件的创建。

```
ItemIndex : OleVariant;
```

```
Template: = EmptyParam; //表示不使用模板
```

```
NewTemplate: = False; //表示新建文档的类型
```

为文档,

```
//添加新文档
```

```
WordApplication. Documents. Add ( Template, NewTemplate );
```

```
WordDocument1. ConnectTo ( WordApplication. Documents. Item( ItemIndex) );
```

第三步:通过 TwordDocument 实现数据的插入。

```
WordDocument. Range. InsertAfter( '这是个例子'+ chr( #13) );
```

```
WordDocument. Tables. Add ( WordDocument. Range, rownum, colnum, template, newtemplate ); //
```

其中 rownum 是表格行数, colnum 是表格列数

```
//向表格中插入数据
```

```
WordDocument. Tables. Item ( 1 ). Cell ( 1, 1 ).
```

```
Range. Text: = '姓名';
```

```
WordDocument. Tables. Item ( 1 ). Cell ( 1, 2 ).
```

```
Range. Text: = '学号';
```

5 总结

以上所讨论了在 WEB 环境中生成 WORD 文档的三种基本方法。其中第一种方法只是简单的把网页中的内容存成 WORD 格式。而第二和第三种方法通过调用 WORD 对象进行编程来很好地控制 WORD 文档的版式。因此,如果只需要生成简单的 WORD 文档,可以使用第一种方法,因为其实现比较简单。但如果需要进行复杂控制的话可以选用第二和第三种方法。

参考文献

- 1 北京博彦科技, Office VBA 编程高手, 北京大学出版社, 2000。
- 2 晶辰工作室, Word 2000 中文版 VBA 开发实例指南, 北京电子工业出版社, 2000。
- 3 陈宣华, 如何利用 VBA 在 ASP 中生成 WORD 文档试卷, 宁波高等专科学校学报, 2003(6)。