

# 利用 ADO stream 实现 office 文档向 SQL server 数据表字段的直接导入导出

李明 赵恒永 邓亚

(北京化工大学信息科学与技术学院,北京 100029)

**摘要:** 分析了笔者在做实际项目中所遇到的问题,提出了可行的解决办法,介绍了利用 ADO Stream 对象实现 Word、Excel 以及其他文档格式的文件向 SQL Server 数据表 image 字段直接导入导出的方法。并且探讨了 ADO Stream 对象模型及其相关主要技术。

**关键词:** SQL Server;ADO Stream;导入导出

**中图分类号:** TP311.52; TP315

## 引言

随着计算机高新技术向各个领域的渗透,实现办公无纸化管理已经不再是梦想。笔者在为某大型国有企业专利事务所开发专利管理系统的过程中所要实现的公文无纸化管理中一个重要的功能就是把各项专利申请在申请过程中的所有公文都以电子文档的形式进行系统备份,即进行电子归档。

在整个专利申请的过程中所要用到的申请文档有很多种,如说明书、权利要求书、说明书摘要、费用减缓请求书等。首先遇到的问题就是如何设计存放所有专利申请文档数据表的结构。

(1) 所有的专利文件都要在 SQL Server 7.0 中存储,操作系统为 Win 2000。

(2) 不管某项申请是否申请成功,向专利局递交申请书的时候,专利局会分配一个申请号,而这个申请号是唯一标识该申请的关键字。管理员通过申请号实现对各项申请在整个申请流程的监控。

(3) 客户要求寄往专利局的所有文件都存放在数据库中,对于最初的申请文件,说明书、权利要求书、说明书摘要、说明书附图、摘要附图、请求书、费用减缓请求书、代理委托书可以从软盘中直接拷入;对后续的实审请求书、陈述书等表格可以直接拷入也可以由事务处理过程生成 office 文档后再存入。

(4) 所有的申请文档都是 Word、Excel 或者其他文本文件格式。

## 1 方案设计

如何设计存放所有专利申请文档数据表的结构,传统方案是为每个专利法所规定的申请文档分别建立一张表格,表结构是根据实际的申请文档的格式所定的。实际的申请文档的格式有几项内容:表结构应该有相应的字段对应;每个表格中的记录数应该相同,所有记录的关键字(申请号)一一对应。

传统方案的明显缺点为:首先,不易扩充性,如果专利局下发了新的申请文件格式,则需要新增加一张相应的数据表,考虑到数据的完整性,新增加的这张表格里要自动添加在其他表格中已经存在的申请号的全部新记录;其次,需要建立专利局法定申请文档数量的相应表格,使库结构变得繁琐;最后,安全性比较差,需要对字段进行加密处理。为了克服上述弊端,笔者自行设计了一个新方案,建立申请档案总表(表 1)。

除编号、申请号以及专利名称以外,专利法所规定的所有文档类型作为该表的其他字段名,并且字段类型是 image,字段里存入的数据值是文档的实际内容的二进制形式。本方案显著的优点为:易用性,所有的申请文档只存放在一个表格中,而且表结构和人们日常思维一致,申请号为主关键字,后面的字段是本申请的所有申请文档。安全性,后面申请文档的字段类型选用 image,而 image 类型是 SQL Server 支持的二进制数据类型。最大好处是它不随

收稿日期: 2002-07-06

第一作者: 女,1978 年生,硕士生

E-mail: mingli @grad. buct. edu. cn

服务器所使用的代码页不同而有不同的解释。所以,它常用来存储图象等数据类型。本文用它来存储 Word、Excel 以及其他形式的文档文件。用二进制数据类型来存储文件内容无疑增强了系统的保密性,相当于对字段值进行了加密。即使有非法用户盗用密码进入数据库,他所看到的字段内容也是一堆毫无意义的乱码。易扩充性,如果专利局下发了新的申请文件格式,只需要增加表的字段就可以了,而且不需要维护表格之间数据的完整性<sup>[1]</sup>。

表 1 申请档案总表

Table 1 Total table of application documents

序号	字段名称	类型	长度	备注
1	编号	nvarchar	4	与职务专利数据库中的编号是对应的
2	申请号	nvarchar	12	关键字
3	专利名称	nvarchar	50	与职务表中的名称是一致的
4	受理通知书	image	16	文档实际内容
5	权利要求书	image	16	文档实际内容
6	说明书摘要	image	16	文档实际内容
...	...	...	...	...

实现本方案的难点在于如何实现 Word、Excel 以及其他文档格式的文件与 SQL Server 中二进制数据的相互转换,即实现 Word、Excel 以及其他文档格式的文件向 SQL Server 数据表 image 字段的直接导入导出。

## 2 程序实现

### 2.1 ADO Stream 对象模型简介

利用 ADO Stream 对象实现 Word、Excel 以及其他文档格式的文件向 SQL Server 数据表 image 字段的直接导入导出。

ADO Stream 对象是 Active Data Objects(ADO) 2.5 里引入的,它可以极大的简化对 SQL Server 数据库的二进制大数据对象导入导出以及修改的程序代码的编写。一个 stream 对象可以用来操纵包含流数据的字段和记录。它可以从下述 3 种方法中得到<sup>[2]</sup>。

· 从连接到一个文档、文件夹或者记录对象的 URL (Uniform Resource Locator, 在 Internet 的 WWW 服务程序上用于指定信息位置的表示方法) 里取得。

· 为应用程序存储数据初始化一个 Stream 对象而得到。

· 通过打开一个与一个记录对象相关联的缺省的 Stream 对象而得到。

利用 Stream 对象的属性和方法,可以做如下工作:

· 向一个 Stream 对象里输入字节或者文本数据;

· 从一个 Stream 对象里读出字节数据;

· 把在 ADO 缓冲里存在的任何数据写到一个附属的对象里;

· 把一个 Stream 对象里的数据拷贝到另一个 Stream 对象里;

· 控制从一个源文件里读取数据的行方式;

· 把流数据存储在文件中去或者反向操作;

· 确定一个流对象里的数据类型<sup>[3]</sup>。

### 2.2 程序源代码

WriteFileToField() 函数完成把数据从一个源文件里写到数据表的某一个字段里,有关的程序源代码如下:

```
Public Sub WriteFileToField ( sFilename As String, fField As ADODB.field)
```

```
// 定义了两个参数, sFilename 表示的是源文件的路径名, 注意这个路径名是一个完整的路径名包括文件的扩展名。比如 Word 文件扩展名为 doc, Excel 文件的为 xls, 文本文件为 txt。 fField 表示的要把源文件数据导入的字段, 数据类型为 image。
```

```
Dim mStream As ADODB.Stream
```

```
// 定义 mStream 为一个流对象
```

```
Set mStream = New ADODB.Stream
```

```
// 初始化 mStream 流对象
```

```
mStream.Type = adTypeBinary
```

// 设置流对象 mStream 数据类型的为二进制类型

```
mStream.Open
```

```
// 用 Open 方法打开 mStream 流对象
```

```
mStream.LoadFromFile sFilename
```

// 把数据从 sFilename 所指定的源文件中导入到 mStream 流对象中

```
fField.Value = mStream.Read
```

// 利用 Read 方法从 mStream 流对象中读出数据到所指定的字段中

```
mStream.Close
```

```
// 关闭 mStream 流对象
End Sub
WriteFieldToFile() 函数把一个数据表字段的二进制内容导入到一个源文件中去,源代码如下:
Public Sub WriteFieldToFile(sFilename As String, fField As ADODB.field)
    // fField 表示的要导入到源文件数据数据表的字段,数据类型为 image。注意:如果在此操作之前 fField 里存入的数据值是 Word 文档的内容,则在把数据值导出的时候 sFilename 里应指定扩展名为 doc。
    Dim mStream As ADODB.Stream
    Set mStream = New ADODB.Stream
    mStream.Type = adTypeBinary
    mStream.Open
    mStream.Write fField.Value
    // 把 fField 字段的数据内容写入到 mStream 流对象
    mStream.SaveToFile sFilename, adSaveCreateOverWrite
    // 利用 SaveToFile 方法把 mStream 流对象里的内容写到 sFilename 所指定的文件里去, adSaveCreateOverWrite 参数表示如果文件存在则以覆盖的方式写入。如果文件不存在,则生成新的文
```

件。

```
mStream.Close
// 关闭 mStream 流对象
End Sub
```

通过 sFilename, fField 两个参数就可以灵活地控制数据表 image 字段与 Word、Excel 以及其他文档格式的文件相互转换。以上是笔者在实际编程过程中探索出来的方法,增加了编程的灵活性,并且简化了代码量。

### 3 结束语

利用 ADO Stream 对象来实现 Word、Excel 以及其他文档格式的文件向 SQL Server 数据表 image 字段的直接导入导出是笔者在开发项目过程中探索出来的,且已得到很好的应用,程序的源代码可直接移植到其他应用中而无须重写代码。

### 参 考 文 献

- [1] 袁鹏飞. SQL SERVER 7.0 数据库系统管理与应用开发[M]. 北京:人民邮电出版社,1999
- [2] Gunderloy M, Chipman M. SQL Server 7 In Record Time[M]. 北京:电子工业出版社,1999
- [3] Gunderloy M. Visual Basic Developer's Guide to ADO [M]. 北京:电子工业出版社,1999

## Implementation of importing and exporting data of an office document to SQL server field based on ADO stream object

LI Ming ZHAO Heng-yong DENG Ya

(College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China)

**Abstract:** A problem that the writer faced during the course of a actual subject was analyzed, and a viable solution was proposed to settle this problem. Implementation of importing and exporting data of an office document to SQL Server field based on ADO Stream object as well as the ADO Stream object and its main technique were introduced.

**Key words:** SQL server; ADO stream; import and export

(责任编辑 云志学)