

# 甘肃省文化资源普查名录 数据采集与提取研究

□ 张 诤 马廷旭

为贯彻落实党的十八大关于扎实推进社会主义文化强国建设的精神,全面盘清甘肃省文化资源家底,准确掌握文化资源状况,科学评价文化资源价值,建立分类分级管理制度,对全省文化资源进行有效保护、理性挖掘、传承创新,不断推进甘肃文化大省和华夏文明传承创新区建设,提升甘肃省的文化软实力,根据《甘肃省华夏文明传承创新区建设总体方案》,2013年5月,甘肃省正式启动了文化资源普查和分类分级评估工作。该项工作主要分为三个阶段:1.基于各厅(局)报送的文化资源统计数据,编制《甘肃省文化资源名录》初编,用以指导后续阶段的工作;2.分类分级实施全省文化资源普查工作,建立《甘肃省文化资源数据库》;3.对全省普查的文化资源进行分类分级评估,建立《甘肃省文化资源分类分级评估项目库》。计划从2013年5月启动到2015年5月结束,前后历时两年时间。

## 一、《甘肃省文化资源名录》编制面临的主要问题

为了能够尽快系统全面的摸清甘肃省各厅(局)掌握的文化资源案头数据,按照甘肃省委办公厅和省政府办公厅联合印发的《甘肃省文化资源普查暨分类分级评估工作实施方案》的具体要求,甘肃省文化资源普查办公室依据甘肃省的文化资源特点,借鉴国内其他省、市文化资源普查中的文化资源分类办法,并结合甘肃华夏文明传承创新区建设需要,将甘肃省文化资源从总体上划分为历史文化、少数民族语言文字、非物质文化遗产、自然景观文化、宗教文化、文学艺术、饮食文化、建筑文化、节庆(会)及赛事文化、文化之乡、地名文化、新闻出版、广播影视、社科研究、文化类高等教育、文艺机构和团体、文化产业、文化人才等18个大类并分别由14个牵头厅(局)负责报送。要求在2013年10月底前基于各厅(局)报送的文化资源统计数据,初步形成省、市两级《甘肃省文化资源名录》初编。

## (一)甘肃省文化资源名录数据的采集标准问题

文化资源名录数据是文化资源基础数据的索引数据。基于各牵头厅(局)已拥有的案头文化资源数据编制《甘肃省文化资源名录》初编,可以为后续甘肃省文化资源全面普查工作提供直接的参考和指导。按照《甘肃省文化资源普查暨分类分级评估工作实施方案》的具体要求,甘肃省文化资源普查办公室需要采集十四个牵头厅局目前已拥有的18大类文化资源案头数据,从中抽取用以编制《甘肃省文化资源名录》初编的标准数据。但是14个牵头厅(局)所拥有的各大类文化资源均有各自不同的元数据标准,因此,制定统一规范

的甘肃省文化资源名录数据采集标准,是甘肃省文化资源普查办公室编制《甘肃省文化资源名录》初编的基础。

## (二)甘肃省文化资源名录数据的采集处理方式问题

由于甘肃省文化资源普查办公室人力有限,同时14个牵头厅(局)按照甘肃省文化资源普查办公室制定的名录数据采集标准报送的名录数据数量非常大,如果按照传统的各厅(局)人工纸质报送,省普查办集中人工整理的名录数据采集处理方式,将很难在《甘肃省文化资源普查暨分类分级评估工作实施方案》要求的时间内完成《甘肃省文化资源名录》初编的编制工作。因此,需要借助现代信息技术,设计一种准确性和效率更高的名录数据采集处理方式。

## 二、甘肃省文化资源普查名录数据采集模板设计

针对上述《甘肃省文化资源名录》初编编制过程中面临的主要问题,甘肃省文化资源普查办公室技术组根据《甘肃省文化资源普查暨分类分级评估工作实施方案》的具体要求,结合各厅(局)拥有的18大类文化资源数据的特点和元数据标准,首先制定出了统一的文化资源名录数据采集标准,同时考虑到经费投入限制及各厅(局)文化资源数据报送人员的软件使用习惯和水平,技术组最终基于微软OFFICE办公软件中的EXCEL电子表格软件,按照制定的文化资源名录数据采集标准,设计开发了省、市两级文化资源名录数据采集电子模板。

为了保证文化资源名录数据采集的准确性和数据一致性,电子模板设计时充分利用了EXCEL数据表中的格式设置功能,对电子模板里的相应属性字段进行了数据输入内容及格式的限定。例如,设定了资源“分类名称”“所在市州”等属性字段为下拉列表式选择字段,限定了资源“目前现状”“重要价值”等属性字段的输入字数,限定了“资源图片编号”字段的输入格式,保证了名录数据输入的准确性和一致性。同时,为了保证各厅(局)报送的18类文化资源电子模板中采集的名录数据与相应的文化资源图片文件的关联性,技术组设计了电子模板中资源图片的编码规则。“资源图片编号”=市(州)编码(01—14)+县(区)编码(01—87)+资源图片序号(T00001,T00002……,T99999)。例如,与兰州市城关区报送的文化资源名录数据相关联的图片文件的命名编码为:0101T00001-0101T99999。

甘肃省文化资源普查办公室通过电子邮件和构建QQ群的方式将18大类文化资源名录数据电子模板分发给了14个厅(局)的相关名录数据报送人员,各厅(局)数据报送人员根据电子模板的填报说明进行了18大类文化资源名录

数据的填报,并通过电子邮件方式将填报完成的名录数据报送汇总到了甘肃省文化资源普查办公室指定的电子邮箱。

### 三、甘肃省文化资源普查名录数据自动提取转换接口设计

按照《甘肃省文化资源普查暨分类分级评估工作实施方案》的具体要求,甘肃省文化资源普查办公室要在2013年10月底前基于各厅(局)报送的文化资源名录数据,编制完成《甘肃省文化资源名录》初编。由于14个牵头厅(局)报送的名录数据数量非常大,并且都是以EXCEL电子模板形式报送的数据,要将电子模版中的数据通过人工转换的方式形成WORD格式的甘肃省文化资源名录文档,需要的人员数量和工作量将非常大,在人力有限的情况下,甘肃省文化资源普查办公室技术组结合微软OFFICE办公套件的技术特性,利用OFFICE VBA编程工具,针对各厅(局)报送的每一类文化资源名录数据的特点,自主开发了提取EXCEL电子模板数据自动生成WORD格式文档的名录数据提取转换接口。以省文物局报送的甘肃省二级文物名录数据为例,电子模板中一共填报了6121条文化资源名录数据,通过在计算机上运行名录数据提取转换接口程序,可以自动完成对这6121条数据的提取,并按程序设定的文档格式自动生成一个甘肃省二级文物名录WORD文档。具体接口程序如下所示:

```
Sub wwsz2()
'定义WORD应用程序变量
Dim objWordApp As Word.Application
'定义word文档变量
Dim objWord As Word.Document
Dim objSel As Word.Selection
Dim strTitle As String
Dim str As String
On Error GoTo errHandle '错误处理
Sheet1.Select
Set objWordApp = New Word.Application
Set objWord = objWordApp.Documents.Add
objWord.Application.Visible = True
Set objSel = objWord.Application.Selection
With objSel
.InsertAfter Text:=strTitle & Chr(13)
.ParagraphFormat.Alignment = wdAlignParagraphLeft
.Font.Size = 16
ThisWorkbook.Activate
End With
'选择内容
For j = 1 To 6121
For i = 1 To 7
ReDim arr(8) As String
arr(i) = Sheet1.Cells(j, i)
Set objSel = objWord.Application.Selection
With objSel
.InsertAfter Text:=arr(i) & Chr(32)
.InsertAfter Text:=strTitle & Chr(13)
```

```
.ParagraphFormat.Alignment = wdAlignParagraphLeft
.Font.Size = 16
.EndKey Unit:=wdStory
Application.CutCopyMode = False
ThisWorkbook.Activate
End With
Next i
Set objSel = objWord.Application.Selection
With objSel
'设置word文档格式
.InsertAfter Text:=strTitle & Chr(13)
.InsertAfter Text:=strTitle & Chr(13)
.ParagraphFormat.Alignment = wdAlignParagraphLeft
.Font.Size = 16
.EndKey Unit:=wdStory
Application.CutCopyMode = False
ThisWorkbook.Activate
End With
Next j
'设置word文档保存路径
objWord.SaveAs Filename:="C:\test\wordtest.doc", _
FileFormat:=wdFormatDocument
objWord.Close
objWordApp.Quit
errExit:
Set objSel = Nothing
Set objWord = Nothing
Set objWordApp = Nothing
Exit Sub
errHandle:
MsgBox Err.Description
Resume errExit
End Sub
```

甘肃省文化资源普查办公室利用技术组自行设计开发的名录数据采集电子模板和名录数据自动提取转换接口,对14个厅(局)报送的甘肃省18大类文化资源近10万条名录数据进行了准确高效的处理,最终编制完成了近50万字《甘肃省文化资源名录》初编。

[2014年度甘肃省社科规划项目《甘肃省文化资源普查技术支撑研究》项目编号:14YB119阶段性成果]

### 参考文献

[1]张诤,张亮亮.甘肃“云”上的文化传承[J].信息化建设,2014(1).

[2]张诤.甘肃省文化资源普查技术支撑问题及对策[J].发展,2015(4).

(第一作者单位:甘肃省委党校信息中心,甘肃省文化资源普查办公室;第二作者单位:甘肃省文化资源普查办公室,甘肃省社会科学院)