

# 数字化时代通用古文字字库的开发研究

李 未<sup>1</sup> 马晓丽<sup>2</sup> 盖海红<sup>2</sup> 张庆民<sup>2</sup>

(1 石家庄幼儿师范高等专科学校 河北石家庄 050228; 2 石家庄信息工程职业学院 河北石家庄 050035)

**【摘要】** 随着计算机信息技术的快速发展和越来越多古文字资料的出土,在数字化时代背景下,通用古文字字库开发的信息化发展成为必然趋势。通过计算机技术与古文字的结合,全世界正在推进构建通用古文字字库。本文主要以甲骨文、金文的已识字为研究对象,首先探析数字化时代背景下的古文字字形整理,其次介绍古文字字形整理的重要意义,最后就如何开发数字化时代通用古文字字库提出具体的对策,创建一种分类别、分层次的通用古文字字库,将不同字形的古文字在同一个平台中体现。

**【关键词】** 数字化时代通用;古文字字库;开发研究

DOI: 10.18686/jyxx.v2i1.32815

基于数字化时代,开发创建一个通用古文字字库现代信息化平台,对古文字实行统一编码处理,为每个古文字字形都设置一个相对应的码点,以方便用户输入和调用。

## 1 数字化时代背景下的古文字字形整理

当前,我们对古文字字形的整理建立在古文字学本体的研究基础之上,相较于计算机处理的古文字字形整理,两者之间的差异性较为显著。在古文字学的探索研究中,要求字形整理做到全面性,各种出土文献的古文字都要收录在其中。不过,计算机对古文字字形的整理,则注重差异,编码字符集中的每一个古文字都具有不同的意义,因而古文字的各种字形需要经过处理之后才能收录到计算机之中。古文字字形整理应充分结合考虑古文字字库的收字范围、字形选取标准、编码和存储方式等因素。

### 1.1 进入古文字字库的字符范围

古文字分为三大类,分别为已识字、歧释字及未识字,因此,我们主要从这三方面介绍:第一,古文字中的已识字,指的是经过考释研究,已经得到学界公认的古文字;第二,古文字中的歧释字,这种古文字包括多种考释结论,且还没有经过学界公认。这些古文字虽然能隶定,但是现代字形中没有与之相对应的;第三,古文字中的未识字,这些古文字未经过考释、没有办法隶定,也没有与之对应的现代汉字。

### 1.2 古文字的编码方式

当前,学界对古文字的编码方式主要有两种,第一种,运用所有古文字编码的方式,第二种,缺少对应现代汉字的古文字编码,即便有对应的也会将原有的码位覆盖和替换。上述两种古文字编码方案在古文字字形整理中的应用,既可以针对古文字的所有字形整理出结构特点不同但是字形完备的字形,又可针对同一古文字的不同异形字。不过,这两种编码方案相对计算机出来而言各有不足之处,究其根本还是因为古文字中的异形字繁多,不同时代的字形又各具特点,因此如果针对每个字形单独编写码点,那么必然出现一个古文字拥有多个码位的现象,这样也不方便对古文字的检索和统计。

通用古文字字库的构建是基于古文字中的已识字,

因此,为满足古文字的研究需求,只要利用现代汉字的编码,为与之相对应的不同字形的古文字设置不同的字库即可,不必再为古文字中不同字形进行单独编码。

### 1.3 古文字字形在字库中的储存方式

在字库中,古文字有两种存储方式,第一种存储方式是按照古文字类别,构建相应的大字库,也就是将同一类别的古文字字形统一收录在一个字库之中;第二种存储方式是将不同类别的古文字字形,按照层次划分归属构建成为较小的字库。这两种存储方式在实际运用中各有利弊。

当前,Unicode 国际字符编码体系,极大地简化了古文字字体的制作,不过,两种存储方式相比较,第一种有着较强的实践操作性,能为同平台处理古文字带来方便。不过这种方法对于同一个古文字的不同字形,就需要单独进行输入或者调用,在实际应用中这种分开操作又非常不方便。第二种存储方式对于古文字研究人员来说应用更具有科学性,能让我们对古文字字形进行分门别类的整理时,可以根据时代来划分古文字,为同一种古文字创建不同的子字库。

### 1.4 古文字字样的选取标准

关于古文字字样的整理,主要是选取古文字中的某一代表字样作为“主形代表”,基于数字化背景,运用计算机信息技术整理古文字字形,前提是需要整理出与该古文字有区别的字形,并将之收入古文字字库。因而,古文字字形整理具有多层次性和多角度的特点。古文字字样主要选取标准有时代、字形结构、笔形、笔画及字形等。

## 2 古文字字形整理的重要意义

### 2.1 弘扬中华民族传统文化的需要

数字化时代背景下,古文字字形整理对于古文字界、信息技术及中国传统文化的深入探索和传播具有重要意义。古文字属于汉字中的重要内容,我国五千年历史文化的传承就是以古文献为载体,其主要分为传世文献和出土文献两大类,而古文字就属于先秦出土的古文献载体。

运用科学有效的方式针对古文字字形整理,将古文字纳入国际标准传播平台,对弘扬和传承中华民族传统

文化起到推动作用,同时也为相关领域的研究提供先进的现代化研究手段,以促进实现中文信息处理的完整性。

## 2.2 创新改革落后存储传播检索方式

古文字研究在资料存储、传播、检索方面,相对绝大多数的学科来说较为陈旧落后。一方面,在活字印刷技术到电脑排版印刷技术的漫长发展历程中,古文字从未被主导的汉字印刷排版所接纳。古文字资料面临的印刷、传播问题显而易见;另一方面,古文字的检索手段也并不像现代通用汉字一样成熟,这也给古文字资料的查找带来限制。总之,正是因为古文字资料查找困难,才导致古文字研究中对资料的掌握问题频发。

## 2.3 为古文字研究创新引入现代化手段

古文字资料系统的局限性,限制了古文字的研究,特别是在研究方法的引入和实际运用中偏旁分析法的使用较为随意,而现代汉语研究中较为常用的数理统计、计算机自动分析等并没有得到应用。对于古文字的研究,在搜集和查找资料时,研究者需要投入精力,减少在材料的综合思辨方面的投入,培养能够优秀的古文字科研人员。

# 3 数字化时代通用古文字字库开发的对策

## 3.1 古文字字库的设计理念

通用古文字字库的设计应当遵循以下设计理念。

一方面,凭借“字位”概念,古文字中的甲骨文与金文属于一种字体,它们作为已识字,与现代汉字具有传承对应关系,因而按照“一字一码”的标准原则,可以将甲骨文和金文当成字体处理。因为在同一个字库中,一个码位与一个字符相对应,所以需要构建不同的古文字字库,才能实现不同字形的古文字共用一个码位。

另一方面,按照时代分期制作古文字字库,展现古文字字形的发展和演变过程。通用古文字字库的创建,能将一个古文字的所有字样制作成不同的字库。因此,应以字库的分类作为依据,重新统计字库的数量。整合古文字现有研究发现,时代分期是构建通用古文字字库的最有效解决方案。

## 3.2 古文字字库的制作

古文字字样整理出来以后,都是图片的格式,因而需要借助相应的软件或者工具,将其转化成文字的格式,如此才能收录进古文字字库,这也就是古文字字库的制作过程。通用古文字字库一般采用 True Type 字体格式制

作,就是以“宋体——方正超大字符集”作为基础,运用 Hyper Snap-DX 截图软件和 Font Creator 字体制作软件来完成字库的制作。

首先,运用截图软件切分和提取古文字字样的图片,进行编辑后,运用字体制作软件,将字样图片转变成 TTF 字体格式,这样是为了保证字体不会出现生字形的讹变。制作而成的字体安装在电脑后就可以输入所需的甲骨文或者金文字体。之后,在 WPS 或者 word 这些软件上安装,不管哪一种输入方法,都可以任意调用和处理古文字。

## 3.3 积极发挥数字化优势

基于数字化平台,古文字的研究也要积极调整传统理念和方法,以全新的研究视角和研究层次进行探索。

首先,顺应数字化时代发展,加强古文字研究新思维的孵化,一方面加强数字化功能的推介工作;另一方面,对新研究成果的储存中,重视研究资料处理过程的存储。

其次,加大力度积极构建数字化平台,实现资源的共建和共享,并设置相应的机制对古文字数字化平台进行管理,最大限度地实现平台资源的共享。

最后,积极创建古文字字库数字化平台的监督检验功能,在不断完善过程中根据学术规范的监督检验需求,重新关联数据和创建路径,运用数字化手段支持古文字研究的学术规范系统。

# 4 结语

数字化时代,加大力度构建通用古文字字库,不仅方便古文字的使用,而且能够创建一个具有实用性、高效性的古文字数字信息化平台。本文基于对甲骨文和金文已识字的研究,首先探讨古文字字形整理的原则,其次介绍古文字字形整理的重要意义,并就如何构建通用古文字字库提出具体的对策,推动实现古文字的信息化处理与调用,在普通输入法的状态下能高效准确地输入和显示古文字字形。

**作者简介:** 李未(1979.9—),男,河北石家庄人,本科,副教授,研究方向:电子商务、行政管理。

**基金项目:** 本论文系 2019 年度石家庄市科学技术研究与发展计划项目《数字化时代通用古文字字库的开发研究》(项目编号:191130441A)的研究成果。

## 【参考文献】

- [1] 刘根辉,张晓霞.古文字字形整理与通用古文字字库开发研究[J].古汉语研究,2016(3):51-56.
- [2] 柳申奥.我国计算机字库字体的产生与设计应用研究[J].美术教育研究,2017(7):78-79.
- [3] 石义祎.中文字体的设计与应用研究[J].文艺生活·文海艺苑,2019(6):167-168.
- [4] 张晓霞.面向计算机处理的古文字字形研究暨字库开发[D].湖北:华中科技大学,2014:104.
- [5] 张峰,谭生力.论古文字中戈字变体及相关诸字形音义[J].江汉考古,2016(4):113-119.
- [6] 王翠叶.关于数字化时代的汉字应用问题[J].海外华文教育动态,2016(2):75-80.
- [7] 尚伟.基于认知心理视角的古文字信息处理研究[J].情报科学,2013(7):135-139.
- [8] 王颖.汉字书法字体的数字化生存[J].包装工程,2018(22):46-50.