

数字技术破解千年甲骨谜团

甲骨文,这一源自3000年前的神秘文字,蕴藏着古人的文化精髓和先民的创造智慧,承载着华夏文明的变迁。然而,在这些镌刻于龟甲、兽骨的特殊符号被彻底破译前,甲骨文却面临着风化侵蚀、流失散佚等挑战。随着科技的进步与跨学科合作的深化,这一古老神秘的文字,正在实现跨越时空的永生。

破解谜团需迈过三道槛

自“甲骨文之父”、金石学大家王懿荣偶然揭开甲骨文的面纱,百余年间,为了破译“文明密码”,我国学者焚膏继晷,至今仍面临三大难关:风化侵蚀严重、大量流失逸散,释读难度极大。

3000年来,现存的甲骨,每一刻都在与风化、虫蛀等自然力量进行着无声的抗争。甲骨本身极为脆弱,在出土、保存、展示等各个环节,都极易因自然与人为因素受损。

中国社会科学院学部委员宋镇豪曾无奈地表示:“存在社科院的甲骨,仅仅过了几年,有些字就已经模糊不清,无法辨认。”

此外,因历史原因,大量甲骨散落于海外。日前,安阳师范学院甲骨文信息处理教育部重点实验室(以下简称“甲骨文实验室”)主任刘永革介绍,殷墟出土的甲骨总量约为16万片,现分别收藏于国内外181家博物馆、图书馆、科研院所及高校等机构中。其中,国内约有12万片。

当前,尚有数万片甲骨被私人藏家持有。而由于私自滥掘和盗卖,3万余片甲骨流失海外,散落于卡内基博物院、苏格兰皇家博物馆、大英博物馆和普林斯顿大学等地,还有不少未公诸学界,这都构成了中国文化遗产无法估量的损失。

2016年,中国文字博物馆发出的甲骨文破译“召集令”——“破译一个字,奖金十万元”,曾让这一冷门学问成为热点话题。高价悬赏,激发了更多人参与甲骨文释读的热情,一定程度上也揭示了甲骨文字研究的艰巨性。2018年,经专家委员会认可,复旦大学出土文献与古文字研究中心研究员蒋玉斌,凭借对甲骨文“蠡”字的释读拿到奖励。此后再无获奖。

据统计,甲骨文单字共有4000余个,其中仅有约三分之一的字形被成功破译。在这些古老的甲骨文里,一部分象形文字由于它们的形态与现代汉字差异较大,很难准确辨识其含义。对于同一个甲骨文字符,不同学者的解读往往存在显著差异。

宋镇豪表示,容易识读的字大都已



这是在殷墟博物馆新馆拍摄的考古发掘出土的甲骨文骨片。

新华社记者赵鹏 摄

经破译,剩下的都是难度大的。只有认真探索,根据甲骨文例、语境、文字属性、字体构型分析,结合金文等其他古文字以及往后的简牍文字,查其流变,集结成不同历史时期每个单字形体构型的信息包,才有可能破解谜团。

数字缀合技术显身手

9月15日,在甲骨文实验室,甲骨文字形识别与分析研究室主任张展向记者边演示边介绍。只见他登录“殷契文渊”网站后,在字形库中选择甲骨文“焚”,所有包含该字形的甲骨片信息悉数显现。

“这是该实验室打造的甲骨文大数据平台。该平台于2019年向全球学者开放。这一平台集甲骨文文字库、著录库、文献库三库于一体,目前已上线甲骨著录153种,甲骨图像多达23.9万余幅,甲骨论著共计3.3万余种。”张展说。

“数字化、数据化、智能化”是“殷契文渊”的三个关键词。刘永革介绍,该平台的建设初衷是对一百多年来所积累的相关数据进行全面系统的整理,建成世界最大的甲骨文数据中心。

“这个甲骨文形似电视机,还‘长’了两个耳朵。其实,它代表‘灯’,寓意一种祭祀形式。”张展向记者展示道,“这个字是凭借数字技术缀合破译的。”

甲骨缀合,是一项根据甲骨形态及所刻甲骨文,精确复原甲骨碎片位置的拼接操作。在数千年的埋藏过程

中,甲骨因地层压力、水湿浸润及发掘翻动等因素,往往由完整一片断裂成多片。数字缀合,正是突破这一研究瓶颈的关键技术。

张展进一步介绍,拼接后的甲骨文如获新生,成为历史考古的新材料,价值倍增,因此这项技术被称为甲骨文的“再发掘”。该实验室正致力于利用人工智能等技术,对甲骨碎片进行自动缀合,目前已缀合70组甲骨碎片。

甲骨脆弱,令人揪心。对此,研究人员通过建立数字化全信息数据模型,为甲骨文物赋予永恒的“生命”。

2022年,甲骨文实验室联合腾讯公司推出“甲骨文全球数字焕活”计划。目前,他们已经取得“微痕分析”和“甲骨三维建模”等技术突破,构建起包含三维模型、微痕增强图、微痕灰度图、数字摹本、数字拓片及科学影像层等在内的“甲骨文全信息数据模型”,实现了甲骨实物在数字空间的高保真还原和保护。

“这是一项比绣花还精细的工作,也是一项堪比科学实验的复杂工作。”刘永革说。

在用高精度相机拍摄甲骨片时,研发人员需综合运用高清拍摄、微距拍摄、三维建模、红外线拍摄、多光谱拍摄等一系列技术手段,对每一片甲骨拍摄150余张照片,确保进行全信息采集。研发人员再将不同甲骨图片进行精确对齐与叠加,每个图层都显示不同的信息。

“比如这个字,它在甲骨拓片上是看不到的,甚至实物都看不清。”腾讯SSV数字文化实验室首席架构师王朝阳指着一张甲骨图片介绍,“但是借助微痕增强、高保真等技术手段,我们得以窥见这些‘隐藏内容’。”

通过微痕分析,记者不仅能看清甲骨片上的文字刻写风格和笔画叠压关系,还能深入了解钻凿形态、卜兆裂纹、材质纹理等信息。

AI协同打造“朋友圈”

人工智能,能否成为解锁古代“密码”的钥匙,为甲骨文释读带来新可能?

“我们的‘人机协同助力甲骨文考释’项目,目前已成功构建了包含173万

个甲骨文的庞大字库。该项目的推进,就是为了解决甲骨实物‘看不清’、文字‘查不到’等基础性问题。”刘永革说。

“AI是一股不可忽视的力量。”王朝阳告诉记者,“在‘殷契文渊’AI协同平台上,我们事先采集高清的甲骨信息,然后把实物文字和标准文字进行关联,供专家学者查找。”

该平台已有4500个标准字形,其中1500个已被破译、3000多个未被确认。使用者只要输入当今汉字或者文字图片,就可以查询到甲骨片或者文字。

“这个平台,不仅仅是一个‘人机协同’的搜索工具,更是一个‘人人协同’的社交媒介。”王朝阳说,“在这里,每位专家都拥有标注的权限,只要他们愿意,便能将个人的研究信息与同仁共享,实现知识的‘相互看’,彼此间灵感碰撞,共同激发新的研究思路。在这样的协同合作下,我们将会实现‘1+1>2’的效果!”

王朝阳进一步举例说明:“以‘车’字为例,甲骨文中‘车’字的出现频率极高,但要将其聚合起来却并非易事。而现在,借助计算机匹配与AI协同技术,我们可以迅速在甲骨片上找到并标注出‘车’字。”

此外,甲骨文单字的写法千变万化,为了准确识别一个文字的标准字形,开发人员需要了解与其相似的文字有哪些,这个文字在哪些文献中出现过,以及它的偏旁部首是什么。为了解答这些问题,我们与中国文字博物馆和正大文字中心紧密合作,将更多珍贵资料导入了平台。

他补充道:“再如‘丁’字,在这片甲骨上是这样写的,而在其他甲骨上又是如何呈现的呢?只需轻轻一点,相关信息便会立刻呈现在眼前。通过字形匹配,众多信息被紧密关联起来,为甲骨文的整理与研究带来了极大的便利。”

此外,甲骨文实验室还免费提供了甲骨文手写字、甲骨文字检测、甲骨文字识别等三个开放数据集,这些数据集已被国内外各大研究机构和科研院所广泛使用,下载次数已超过3.5万次。它们为甲骨文的“数字化”与“智能化”研究提供了有力的数据支持。

据科技日报



游人在河南省安阳市殷墟遗址甲骨碑廊游览。(资料照片) 新华社记者鲁鹏 摄