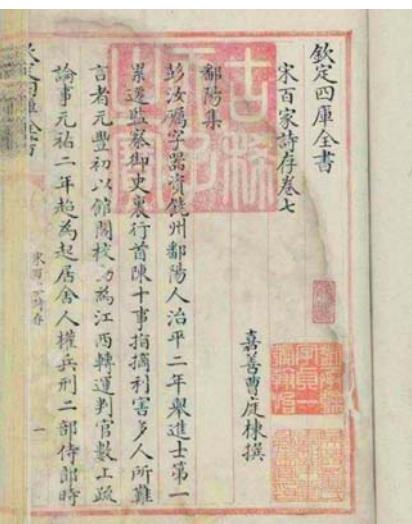




【文化观】

AI让古籍里的文字“活”过来



□云韶

这两天,一个名为“汉典重光”的数据库引发舆论关注。据报道,这个数据库由四川大学、美国加州大学伯克利分校、阿里达摩院等机构联合研发,字数超过了三万字。含40余种珍贵宋元刻本、写本;明清至民国时期著名学者钱谦益、翁方纲、王韬的抄本、稿本;著名藏书楼嘉业堂、密韵楼的抄本,还有清文澜阁《四库全书》零本等。

事实上,古籍数字化并不是什么新鲜事。比如国家图书馆的“中华古籍资源库”已在线发布超过3.3万部的古籍影像;中华书局的“中华经典古籍库”已发布3000多种、15亿字的点校本古籍;爱如生公司的“中国基本古籍库”收书1万种,既有可供检索的全文,又提供古籍原版图像……这些古籍数据库,已经成为很多文史研究者须臾不可离的助手,极大缓解了读者阅览古籍的难题。

不过这些数据库的建立之初,基本上都是人工录入,制作起来耗时耗力。而“汉典重光”项目最大的亮点便是用AI替代人工录入,大幅压缩了专家标注的工作量。就像报道中所说,“百万字书籍的数字化工作量从1000天降低到了35天,效率比人工专家录入方案提升近30倍”。笔者坚信,倘若这一技术能够推广,古籍数字化的进程将会驶入快车道。

若果真如此的话,实为古籍之幸。元代著名散曲家张养浩有一首《山坡羊·潼关怀古》,其中言道:“伤心秦汉经行处,宫阙万间都做了土。”土石所垒的建筑,尚且敌不过时间的冲击,更何况写在竹帛上的文字,记在纸张上的书籍?我们常说我国历史源远流长,古籍浩如烟海,殊不知,我们现在所看到的古籍只不过是“惊鸿一瞥”,翻阅史书的“经籍志”,看看古代的目录学著作,成千上万的古籍其实早已消失在历史的长河之中。

为了保存古籍,古人想了很多办法。最常见的是保存“副本”。像《永乐大典》《四库全书》这样的大部头都有多种版本,分藏在不同的藏书楼中。但无论藏多少部,这些古籍也是纸做的,水、火、地震、虫害,都会让它们烟消云散。此外,因邦交、贸易、战乱等,历史上中国古籍时有出海,近代以来的战争和动荡更加剧了古籍的损毁和流散。据不完全估计,散居海外的中国古籍超过40万部、400万册,包括甲骨简牍、敦煌遗书、宋元善本、明清精椠、拓本舆图、少数民族文献等。

即便这些古籍足够幸运,能够完好地

保存至今,也不能够随意借阅。有测试表明,一部宋元古籍,离开专用书库,置于普通阅览室中供人翻阅一小时,其寿命就会缩短数月。如今,幸好有了数字化的手段,我们可以利用现代信息技术对古籍文献进行加工处理,使其转化为电子数据形式,通过光盘、网络等介质让古籍得到保存和传播。数字化无疑给了古籍新的生命。

特别是, AI技术假如能够得到广泛运用,不仅会降低古籍数字化的门槛,还会提高古籍数字化的准确性,这对于延续古籍的生命大有裨益。网络上传播的一些数字化古籍因为输入错误或者文字识别技术上的不足,篇章搞乱者有之,错字错句者有之,如果从网络上直接下载复制这些文献资料而不核对原文,一定会出现许多硬伤。而AI技术的运用,则解决了这一问题。如“鍊”“鍊”二字, AI能够很轻松地区分它们。“鍊”的CJK的字符代码是“934A”,“鍊”的CJK的字符代码是“932C”,因此AI会按两个不同的字来对它们作处理。但对审校专家和技术人员来说,人眼区分“鍊”与“鍊”却难得多。

古籍数字化,不仅有利于保护古籍,还方便了古籍走进普通人的生活。在过去很长一段时间,古籍对普通人来说是陌生的,一方面是因为古代文盲率较高,人们看不懂,另一方面则是古籍备份少,人们难以得到。古籍善本长期以来秘不示人,几乎成了传统,历史上有的藏书家在书上盖有这样的印章:“鬻及借人为不孝。”古籍既不能卖,也不能借。钱谦益是明末排名第一的藏书家,绛云楼不慎失火,大部分藏书化为灰烬。他的好友曹溶去安慰他,同时也批评他吝于把藏书示人,致使绛云楼所藏孤本绝迹人间。为什么不给人看呢?主要还是为了保护。

如今,人们的受教育率大幅度提高,读书识字已经不是什么难事。而古籍数字化之后,古籍“母本”不再需要冒着各种风险“抛头露脸”,甚至可以“颐养天年”,与此同时,数字化的古籍可以分身千百次、走出“深闺”,实现同一时间满足不同地域、不同读者的阅读需求,做到了一对多、点对面、虚对实的变化。

这样的变化有利于提高我们的文化自信。钱穆先生在《国史大纲》开篇曾提出,“当信任何一国之国民,尤其是自称知识在水平线上之国民,对其本国已往历史,应该略有所知”。钱穆先生为什么要有这个建议,就是因为有很多人“站不起来”了。近代以前,我们国家一直是世界强国之一,我们的人民对自己的文化有着强烈的认同感和自豪感。鸦片战争以后,我们国家陷入了内忧外患的境地,中国人民经历了战乱频仍、山河破碎、民不聊生的深重苦难。由于一直受人欺负,很多地方不如人家,许多民族自信心不足。甚至出现了“全盘西化论”的极端论调。毛泽东同志曾经讲过,有些人做奴隶久了,感觉事事不如人,在外国人面前伸不直腰,批评的就是这种现象。

只有知道我们的生命缘起何处,才能知道我们的脚步迈向何方。要解决这一问题,人们就要多读优秀的古籍,特别是要反复品读古籍中的经典。因为,大家只有熟悉了我们自己的历史,才会知道我们能够在世界民族之林傲立五千年,秘密是什么。大家只有熟悉我们自己的历史,才会知道我们被列强蹂躏了百年之后,还能够再次站起来、富起来、强起来的力量是什么。

□郑学富

童年时光天真无邪、无忧无虑。现在的儿童游戏花样齐全,品类繁多。在古代却没有这些现代化的游戏,那时候儿童玩游戏是置身于大自然,亲近山川河流,童年趣事最多的是垂钓、牧牛、放风筝、捕知了等。让我们从浩瀚的古诗词里,去寻找古代儿童玩的游戏活动。

“小娃撑小艇,偷采白莲回。不解藏踪迹,浮萍一道开。”唐代诗人白居易的《池上》一诗,勾画了一幅儿童采莲图。大和九年(835年),时任太子少傅分司东都洛阳的白居易,一日游历于池边,见山僧下棋、小娃撑船而作此诗。在莲花盛开的夏日里,一个活泼可爱的小男孩,划着一条小船,偷偷地去池塘中采摘白莲花。他兴高采烈地划着满载“战利品”的小船而归,却不知道掩盖自己“偷窃”的踪迹,水面的浮萍上留下了一条船儿划过的痕迹。诗情画意,有景有色有行动描写,细致逼真,富有情趣;而小主人公天真幼稚、活泼淘气的可爱形象,也就栩栩如生,跃然纸上了。

“牧童持蓑笠,逢人气傲然。卧牛吹短笛,耕却傍溪田。”唐代诗人崔道融的《牧童》诗将一个放牛儿童演绎得活灵活现。诗人在外出游览途中,见到一个牧童身穿蓑衣,头戴斗笠,横坐在牛背上,悠然自得地吹着短笛,碰见行人更是显得非常神气。牛耕田时,他却在溪边的田头玩耍。诗人将牧牛儿童描写得调皮可爱、可亲,情趣盎然。

“牛儿小,牛女少,抛牛沙上斗百草。”唐代诗人贯休《春野作》中的诗句生动描绘出小孩子欢天喜地玩斗草游戏的场景。司空图《灯花》中的“明朝斗草多应喜,剪得灯花自扫眉”的诗句,写的也是想象赢得斗草游戏的喜悦。斗草,正是古代儿童喜欢玩的游戏之一。

古时,在端午节前后,有一种特别流行的游戏:斗草。斗草最早见于文献是在魏晋南北朝时期,唐朝后渐渐成为妇女和孩童所喜爱的游戏。

梁朝人宗懔在《荆楚岁时记》上有“五月五日,四民并踏百草,又有斗百草之戏”的记载。唐朝时斗草之戏最盛,据记载,唐中宗时安乐公主在端午节斗百草,为了使自己所采的花草种类繁多,她派人快马加鞭去远处摘取。为了保证独此一份,她还命令在摘取后,把其余的花草都剪掉。从宋朝开始,平日里也经常有斗草游戏。到了清代,斗草游戏仍然盛行,不过在诗文的记载中,斗草游戏仅以女性为限了。《红楼梦》第六十二回中记载,“宝玉生日那天,众姐妹们忙忙碌碌安席饮酒作诗。各屋的丫头也随主子取乐,薛蟠的妾香菱和几个丫头各采了些花草,斗草取乐。”

唐代诗人胡令能隐居莆田时,闲来无事,前往农村访友,路过小河边时,看到一个头发蓬乱、面孔青嫩的小孩在河边学钓鱼,幼稚顽皮,天真可爱,他侧着

身子随意坐在草丛中,野草掩映了他的身影,正在聚精会神、一心一意地钓鱼。胡令能忘记了朋友的住处,遂上前问路,听到有人问路,小儿害怕应答声惊跑了正在上钩的鱼儿,从老远就招手而不回答。待胡令能走近跟前,才附在其耳朵边告诉朋友的住处。胡令能被这个不拘形迹、专心致志于钓鱼的蓬头稚子所感染,写下了《小儿垂钓》:“蓬头稚子学垂纶,侧坐莓苔草映身。路人借问遥招手,怕得鱼惊不应人。”将一个钓鱼的儿童描写得活灵活现,形神兼备,意趣盎然。这首诗别有情趣,在平淡浅易的叙述中透露出几分纯真和无限的童趣。

“篱落疏疏一径深,树头花落未成阴。儿童急走追黄蝶,飞入菜花无处寻。”这是宋代诗人杨万里的《宿新市徐公店》的诗句。有一次,诗人住宿在浙江德清县新市镇的徐家客店里,门外的篱笆稀稀落落,一条小路通向远方,路旁树枝上的桃花、李子花已经飘落了,但树叶还没有长大得很茂密,不远处的一片田野里盛开着黄色的油菜花,恬淡自然,宁静清新的田野风光令人向往。儿童们张开双手扑打打,两脚跌跌撞撞追逐着蝴蝶。小小的蝴蝶飞入这黄色的海洋里,儿童们东张西望,四处搜寻,可是已分不清哪是蝴蝶,哪是黄花,再也找不到蝴蝶了。焦急、失望,表现了儿童们的天真和稚气。

“昼出耘田夜绩麻,村庄儿女各当家。童孙未解供耕织,也傍桑阴学种瓜。”宋代诗人范成大的《夏日田园杂兴·其七》一诗描写了儿童学种瓜的天真情趣。

“牛儿小,牛女少,抛牛沙上斗百草。”唐代诗人贯休《春野作》中的诗句生动描绘出小孩子欢天喜地玩斗草游戏的场景。司空图《灯花》中的“明朝斗草多应喜,剪得灯花自扫眉”的诗句,写的也是想象赢得斗草游戏的喜悦。斗草,正是古代儿童喜欢玩的游戏之一。

古时,在端午节前后,有一种特别流行的游戏:斗草。斗草最早见于文献是在魏晋南北朝时期,唐朝后渐渐成为妇女和孩童所喜爱的游戏。

梁朝人宗懔在《荆楚岁时记》上有“五月五日,四民并踏百草,又有斗百草之戏”的记载。唐朝时斗草之戏最盛,据记载,唐中宗时安乐公主在端午节斗百草,为了使自己所采的花草种类繁多,她派人快马加鞭去远处摘取。为了保证独此一份,她还命令在摘取后,把其余的花草都剪掉。从宋朝开始,平日里也经常有斗草游戏。到了清代,斗草游戏仍然盛行,不过在诗文的记载中,斗草游戏仅以女性为限了。《红楼梦》第六十二回中记载,“宝玉生日那天,众姐妹们忙忙碌碌安席饮酒作诗。各屋的丫头也随主子取乐,薛蟠的妾香菱和几个丫头各采了些花草,斗草取乐。”

唐代诗人胡令能隐居莆田时,闲来无事,前往农村访友,路过小河边时,看到一个头发蓬乱、面孔青嫩的小孩在河边学钓鱼,幼稚顽皮,天真可爱,他侧着

短史记

古诗词中觅童真

