

当古籍保护遇上互联网:

一场技术与公益的双向奔赴

■ 本报记者 皮磊

6月10日是“文化和自然遗产日”，不少人选择走进国家典籍博物馆观展，在古籍里感受文化遗产的魅力。在这里，你不仅能看到难得一见的珍贵文献，还能看到来自全国各地的修复精品，展品包括《永乐大典》“湖”字册、元刻本《西夏文大藏经扉画》、元至顺二年（1331）刻本《永类铃方》及样式雷图档、黄河舆图和金石拓本等。其中，国家图书馆藏《永乐大典》“湖”字册更是国宝级藏品。

”

这场展览，是国家图书馆（国家古籍保护中心）与中国文物保护基金会、字节跳动公益联合举办的“同心护珍宝 聚力续华章——字节跳动古籍保护专项基金成果展”，展期两个月，面向社会公众免费开放。展览分为“古韵新生：各地精品荟萃”讲好河川故事；舆图”镂于金石：探源古文字”世界记忆：样式雷图档”四个部分，精选35种39件已经修复完善的珍贵古籍予以展出。除了以实物展出古籍修复成果，展览还展示了古籍修复人才培养与古籍活化利用成果。

2021年6月，字节跳动与中国文物保护基金会发起设立古籍保护专项基金，支持国家图书馆等机构古籍修复、人才培养、古籍活化与数字化等。两年时间里，字节跳动不断加强与各方合作，发挥“技术+公益”优势，在推动古籍修复、古籍活化与数字化、社会参与及倡导等方面取得了显著成效，让古籍保护这一“专业而又枯燥”的话题进入大众视野，并且吸引了越来越多不同群体的关注。

一场与时间的赛跑

古籍是人类文明和文化传承的重要载体。但对一些人来说，由于缺乏接触和了解的渠道，古籍保护仍是一个遥远而又陌生的话题。那么，什么是古籍？古籍又有什么特征？

国家古籍保护中心办公室主任王红蕾介绍说，古籍主要指书写或印刷于1912年以前具有中国古典装帧形式的书籍，包括汉文古籍和少数民族文字古籍，以及甲骨、简帛、敦煌遗书、碑帖拓本、古地图等文献，具有重要的文物价值、学术价值和艺术价值。

在此次展览中《永乐大典》“湖”字册等多个国宝级珍品的展出让观众大饱眼福，也证明了运用科技手段修复珍贵古籍取得的丰硕成果。据介绍，截至2021年年底，全国已累计完成古籍普查登记数据270余万部，累



古籍修复师修复古籍（图片来源：字节跳动公益）

计修复古籍超过385万叶，制定颁布古籍保护国家标准及行业标准17项。

“有一些珍贵的典籍是在古籍普查过程中陆续发现的。相关成果我们已经通过全国古籍普查登记基本数据库和全国古籍普查登记目录向全社会发布，并且实现了全时全域共享。有612家收藏单位、140余万条古籍普查目录已正式出版。”王红蕾告诉记者。

她谈道，自2007年中华古籍保护计划启动实施以来，经过各方不懈努力，我国古籍保护工作取得了瞩目成就，评选并公布了六批《国家珍贵古籍名录》，累计收录古籍13026部；评选六批共203家全国古籍重点保护单位，覆盖全国31个省、自治区、直辖市，包括博物馆、档案馆、高校图书馆等机构。此外，目前已有20个省份建立了省级古籍名录。

而这其中，社会力量的参与非常必要。从目前相关机构出台的各项政策措施及具体实践来看，民间机构、公益机构的参与也发挥了重要作用。

可以说，古籍保护是一场与时间的赛跑。2021年6月，字节跳动与中国文物保护基金会联合设立“字节跳动古籍保护专项基金”，初期启动资金1000万元人民币，定向资助中华珍贵古籍修复和专业人才培养、古籍活化与数字化等项目，其中重点包括《永乐大典》“湖”字册等国家珍贵典籍和特藏文献。

这一专项基金的成立，也是企业、基金会等社会力量参与古籍保护与修复的典型范例。记者了解到，当时中国文物保护基金会、国家图书馆、字节跳动三方分别签订了《中国文物保护基金会与国家图书馆（国家古籍保护中心）战略合作框架协议》《中国文物保护基金会字节跳动古籍保护专项基金设立协议》和《古籍保护与利用公益项目合作备

忘录》。这意味着，三方将在古籍保护、古籍科研、古籍数字化、人才培养、古籍活化与利用等方面展开系列深度合作。

那么，当古籍保护这一专业议题遇上互联网，又能擦出怎样的火花？拥有前沿技术的互联网企业又将从哪些方面介入古籍保护？

科技与公益的双向奔赴

据介绍，字节跳动古籍保护专项基金首批启动资金计划在一到两年内，开展100册珍贵古籍修复，其中优先修复《永乐大典》、黄河和长江文献等重点保护古籍，以及培养100位古籍修复的专职人才。该专项基金启动一个月后，字节跳动公益即联合中国文物保护基金会、国家图书馆发起“寻找古籍守护者”活动，招募创作者通过抖音、西瓜视频、今日头条等平台，创作相关内容、推动古籍活化、助力传统文化传承。

当传统的事业遇上互联网，一切都被按下了快进键。

2022年3月，字节跳动与北京大学合作成立“北京大学—字节跳动数字人文开放实验室”，将人工智能技术应用于古籍资源的智能化整理。当年10月，由该实验室研发的古籍数字化平台“识典古籍”测试版正式上线。据了解，未来三年“识典古籍”将陆续完成一万种古籍的智能化整理工作，基本覆盖儒家、道家和佛学的核心典籍目录，届时将全部免费开放。

字节跳动公益古籍项目负责人家瀛告诉记者，“识典古籍”平台主要应用了字节的三大基础能力，包括OCR（文字识别）、自动标点和命名实体识别。其中，文字识别技术是对古籍的影印版文字进行单个切分，再进行文字识别和顺序识别；自动标点技术是通过序列标注的方式，对

古籍自动进行标点划分；命名实体识别技术是通过序列标注识别文本中的人名、地名、书籍、时间、官职等信息。

“这三大技术对应的实际上就是古籍整理最基础的三个步骤。目前行业内文字识别准确率平均为93%至94%，而识典古籍平台的准确率已经达到了96%至97%。我们对自己提出了更高的要求，就是希望智能识别的能力达到出版水平。”家瀛表示，“在这个项目中，我们会邀请公司内部专门的技术团队来做技术优化和迭代，提供技术支持。同时，我们也会邀请一些古籍保护领域的专家和前辈为项目提供指导。”

据她介绍，除了在古籍整理方面的技术投入和创新，字节跳动也将相关技术应用在了增强内容传播、提升用户互动体验等方面。

如，字节跳动联合中国第一历史档案馆、敦煌研究院、国家图书馆、甘肃简牍博物馆，与PICO、抖音共同打造“古籍寻游记”。该项目以VR互动纪录片为核心，依托PICO的6DoF交互技术以及视频云的光场扫描技术还原“古文献四大发现”——殷墟甲骨、居延汉简、敦煌遗书、明清档案，让用户以更沉浸的方式体验古籍记载的历史。

“古籍保护内容其实是包罗万象的，远比我们最初想象的宽广得多，因此这一事业也特别需要社会力量的参与。我们看到，在互联网技术的加持和引导下，这个话题实现了跨界联合，不论是五六岁的孩子还是青年插画师，都能够以相应的方式参与到古籍保护议题中。这也是为什么古籍保护会成为字节跳动主要的公益方向之一。”家瀛说。

科技赋能开拓新发展方向

据了解，古籍保护分为原

生性保护、再生性保护和传承性保护。其中，原生性保护主要是指古籍修复，先让古籍的生命延续下来，这也是整个保护工作的第一步；再生性保护即古籍数字化或出版，以现代的方式让古籍实现再生，进而被保护下来；而传承性保护主要是指古籍活化，就是把古籍里的知识提炼出来，让大家能够理解、参与进来，只有这样才能产生更多创意或创新性传承发展空间。

自2021年6月设立至今，字节跳动古籍保护专项基金各项工作都取得了一系列重要成果。如在古籍修复方面，资助国家图书馆、天津图书馆等10家国家级古籍修复中心修复完成104册（件）珍贵古籍；在人才培养方面，先后依托云南省图书馆、宁波市天一阁博物院、国家图书馆培养107位古籍修复师；在古籍活化方面，推出“全民晒书”“寻找古籍守护者”等话题，拍摄《穿越时空的古籍》纪录片，引发社会共鸣。

公益力量的参与、互联网技术的加持无疑加快了古籍数字化进程。但不同于其他领域，古籍保护有其自身的特殊性和复杂性。

“古籍保护工作目前主要面临资金不足、专职人才匮乏、社会化传播不足等三大挑战。”通过两年时间的探索，家瀛对古籍保护工作有了更为深刻的认识，在她看来，古籍保护并不是一个冷门的事业：“有很多人在为此坚守和负重前行”。

也正是有了这两年时间的尝试，字节跳动更多扮演一个“撮合平台”的角色：撮合内部和外界力量，整合更多资源，让更多业务板块、企业及社会公众关注古籍保护事业并主动参与进来，积极贡献智慧和力量，最终实现人人公益的愿景。

她谈道：“如果古籍没有得到及时保护和传承，可能就慢慢消失了，所以这是一场与时间的赛跑，我们必须着急。但古籍保护又是需要一定专业支撑才能做的事情，处理不当可能会造成不可挽回的损失，因此又急不得。这看似非常矛盾，却是古籍保护的现状。”

记者了解到，由于古籍保护公益项目的延展性，目前字节跳动内部有很多来自不同业务板块的员工参与，数量超过了2300人，字节内部还专门成立了一个古籍保护公益社团。家瀛谈道，“我们将古籍整理技术开放给员工、志愿者及专业人士使用，让他们在参与中帮助改善相应的技术体验。我们还联合了一些古籍馆藏单位、社会合作团体等，如创作者工作室、出版社等，希望把这些社会力量连接起来，转化成我们的智囊团、资源库，大家一起推动古籍保护公益事业。”