

# 守护文化遗产之根

——中国古籍数字化现状扫描

本报记者 张稚丹

屠呦呦从古医书中获得提取青蒿素的思路获诺贝尔奖，让人意识到古籍中蕴含着惊人的物质和精神财富。

对古籍的整理、保护和利用，就是最大限度地保护与传承中华文脉和传统文化。

然而，经历了数千年的古籍是如此脆弱，经不起翻阅等直接接触，绝大多数都存放在库里，连专业研究者都很难看到。保护，就难以利用；利用，就很难充分保护。现代数字化技术的出现，才让这对矛盾迎刃而解。

至此，我们方能了解2019年11月，国家图书馆（国家古籍保护中心）组织第四次古籍数字资源共享联合发布的价值——全国20家单位在线发布数字资源7200余部（件），全国古籍数字资源发布总数已超过7.2万部，普通读者也可以和研究者一样在网上免费共享古籍数字资源。

古籍数字化有多重要？目前中国的古籍数字化处于什么状况？本文试图做一个简单的勾勒。



## 数字化带来古籍重新发现

对搞文史研究的人来说，占有独家资料几乎意味着初步的研究成果。

北京大学教授荣新江说，古籍数字化带来了古籍发现的新时代，数字目录和古籍全文发布后，我们可能无意中就发现重要的研究素材。过去我们只知道大的馆藏，其实很多小馆藏着珍贵的古籍。

山东大学文学院院长杜泽逊说，因为有了数字化古籍，北京大学教授

修复和保护，新闻出版署下辖的相关出版社负责影印出版，全国高等院校古籍整理研究工作委员会负责古籍的整理研究和人才培养工作。

数字化浪潮改变了固有的秩序。古籍普查在线上登记，阅览在网上免费共享，数字化替身保护了图书原件。如果不拘泥于物质形态只考虑文化内涵，数字化甚至可以看作是流落海外珍稀古籍另一种形式的“回家”。

荣教授是研究敦煌学的，当国内出版社为海外收藏的敦煌文献做黑白图录时，1994年，英国已经开始做IDP

张丽娟和一位年轻人在国图收藏中发现了过去不知道的十行本《尚书注疏》的元刊原印本和杨复《仪礼图》元十行原印本。元十行本，清代以来用的都是明代历次修版，原印本的出现甚至会改变原有的结论。

荣教授希望古籍数字化工作更多地跟学术界沟通，根据需要决定哪些古籍先发布。比如做海上丝绸之路研究，广东、福建、浙江一带的地方图书馆馆藏家谱变得重要。如果国家优先支持开展这类古籍的数字化，将帮助我们占据学术制高点。

## 古籍数字化关系文化命脉

数千年遗存下来的古籍如浩瀚之海。据统计，仅汉文古籍品种就超过20万种，版本超过50万种。

2007年，国务院办公厅发文提出，在“十一五”期间大力实施“中华古籍保护计划”，国家图书馆（国家古籍保护中心）负责摸家底、古籍

（数字化的国际敦煌项目），将英、中、俄、日、德、法、韩所藏敦煌古文献的数字化资源汇聚到英国，全世界的研究者共享。荣新江惋惜我们动手有点晚：“在‘十三五’古籍出版规划讨论的时候，各出版社都在策划要出多少整理本古籍。我提出，为什么不搞古籍数字化，将国家用于古籍整理的经费支持数字化整理？”

他认为，古籍数字化关系到国家的文化命脉，国家应该设立专项基金，支持更大规模的数字化。

## 多头掘进 对接贯通

国家图书馆在古籍数字化方面起到了带头作用。2016年，作为“中华古籍保护计划”阶段成果的“全国古籍普查登记基本数据库”正式上线，包括书名、著者、版本年代、册数存卷的数据一目了然。

同年上线的“中华古籍资源库”备受读者欢迎。国家图书馆副馆长张志清说，因为拥有原版扫描的清晰数据图像，2016年点击率增长了1000%，2017年在此基础上又增加100%，2018年，过去排在地方志、中文图书、民国期刊之后的老四“中华古籍资源库”，页面浏览跃居第一位。

中华书局则利用已出版资源做古籍数据库。“中华经典古籍库”收录了中华书局及其他出版社正式出版的整理本古籍图书，资源涵盖经史子集各部，包含二十四史、通鉴、新编诸子集成、十三经清人注疏、史料笔记丛刊、古典文学基本丛书、佛教典籍选刊等经典系列，并提供便捷的阅读、查询、文献征引等服务。“西南联大数据库”“中华文史工具书数据库”“中华文史学术论著库”“海外中医古籍库”“历代进士登科数据库”之外，他们向外拓展，又设立了“木版年画数据库”“中华石刻数据库”。其中宋代墓志铭数据库为全国第一个在线出版的数据库。

上海图书馆是古籍收藏的大户，目前公益性发布馆藏8000余部家谱、470余部珍贵古籍。考虑到本馆影像资源丰富，但缺乏可以研究的文本资源，于是采用了“众包”模式，发动社会大众、志愿者参与馆藏文献的整理和研究，利用大家的知识盈余去共创价值。

中国哲人先贤对语言的思考，至少有2000多年历史。

《论语》的“正名”、《荀子》的“约定俗成”，都包含了对语言本质的深刻认识。在汉代，《尔雅》、扬雄《方言》、许慎《说文解字》、刘熙《释名》等早期语言文字学经典群星闪耀。

但研究发展最快、学术最繁荣的阶段，要数新中国成立以来的70年。《新中国语言文字研究70年》（中国社会科学出版社）一书，从三个阶段归纳70年发展主脉及推动力量。

新中国成立之初，党和政府对文化教育高度重视，扫盲和提高干部群众语文水平的需求直接推动语言文字研究。

1950年，《人民日报》发表《请大家注意文法》，报刊上相继推出吕叔湘等人的《语法修辞讲话》《语法讲话》，学术期刊进行多场语法大讨论，浓郁的语法学氛围，深化了对汉语语法特点的认识。

当时文字改革三大任务——简化汉字、推广普通话、制定和推行汉语拼音方案以及汉语规范化，推动了语言学诸领域的研究。修订完善后的《汉语方言调查表》成为使用至今的标准方案，推广普通话和拼音促进语音研究；确立普通话词汇规范，促成经典《现代汉语词典》的诞生，带动了汉语词汇学的发展；简化汉字促成了现代汉字学的产生。

以王力的《汉语史稿》为标志，历史语言学观念使原先静态的古汉语研究，发展成揭示几千年发展规律的汉语史新学科；音韵学从难懂的“绝学”，成为可传授、可论证、可验证的现代学问；甲骨文金文为代表的古汉字研究拓展到词汇和语法研究。

改革开放对语言文字学术最直接的推动，



## 热盼统筹 无远弗届

所有研究者都渴盼的支持检索的全文字识别，需要花费的精力和金钱都很巨大，每本书都是一个小工程。

清华大学图书馆副馆长窦天芳提出，高校图书馆各自将其保存的古籍数字化，是否有重复投入？能否国家统筹？我们需要的不仅是数字化资源而且是数据化资源，文本可检索，能自由提取人物、地点、家谱、地理信息、当时的经济状况等信息，这种方式将带来新的研究方法。说到底，就是要将古籍资源库变为古籍知识库，让社会大众很容易地识别、了解、利用。

北京师范大学教授周少川建议，在目录学书目著录上，应考虑增加新的分类，收录一批流通广泛、使用频繁的优秀古籍数据库，如“中国基本古籍库”“瀚堂典藏”“国学宝典”等属于多功能综合古籍电子丛书；而像敦煌、家谱、方志等特色数据库，则应属于专类古籍电子丛书。数字古籍这一新品种，也应纳入古籍普查、保护的范畴，这也是国家文化安全的需要。如果联合发布后，能形成全国上线数字化古籍的联网，甚至可以链接到全球中华古籍的信息，那就更好了。

古籍数据库用户艾俊川说，数字化是对古籍高效、永久的保存。古籍化身千百，不仅造福学者，也让业余文史爱好者方便使用，功德无量。



# 语言文字学的70年历程

刘丹青

是开放健康、全面均衡的学术生态，这是学术可持续发展的保证。

高考制度恢复，为语言学提供人才保障。《中国语文》等两三种期刊发展为约20种核心期刊的方阵。中国语言学会于1980年成立，文字、音韵、训诂、修辞、方言、世界汉语教学、辞书等各分支学科纷纷成立学会及下属分会，广大学人恢复了自主的学术生活。

理论方法的引进和运用，使国际前沿的功能-认知语言学、语法化学说、语言类型学、语言接触等的研究范式成为汉语界的显学。

形式语言学得到深入研究，历史语言学、地理语言学使汉语史研究和方言调查研究更具系统性和多维性。音韵、训诂通过与历史语言学融合，成为探究汉语发展大势的现代学科体系；文字学从具体汉字解析上升到系统的汉字构形学理论；社会语言学、心理语言学等交叉学科成为研究热点。信息技术催生了计算语言学的形成发展，同时反哺语言学本体，促进了语言研究中统计方法的运用和大数据语料库的建设。

1985年国家语委成立后，中国语文政策更加追求和谐合理的语言文字生活。语言规划、

《史记的读法》（广西师范大学出版社），书名老老实实在地陈述了我写这本书的用意，希望大家关注一种在今天似乎变得愈来愈少见的读书法，那就是透过书去接近、体会一颗伟大的心灵。

了解司马迁的心灵结构对于解读《史记》来说是相当重要的。《史记的读法》认真用心地读司马迁留下来的《史记》这本书，不是为了研究汉代的历史，不是为了享受好看的故事，更不是为了考据书中内容的对错，而是为了认识司马迁这个人——他的思想、他的人格，他对于什么是“历史”所提出来的观念，他对于史学家任务与角色的执着态度。

司马迁有着独特的人格，他在生命的关键时刻做出了独特的选择。当所有人都大骂李陵时，他会坚持说出自己所相信的是非；被处以最侮辱的宫刑，他会决定为了将《史记》写完而忍辱活下来。

书写历史并不是简单地罗列事件。他给予自己的任务，是全面整理当时所知道的人类经验，给予这庞大的材料一种可以掌握的秩序，进而能彰显出其中的道理。那道理绝对不是成王败寇，而是人活着有各种不同的价值追求。不同的人以不同方式去追求，因为目的与手段间的差距、个人意志与时代环境的互动，得到或成功或失败的结果。作为一位历史学家，司马迁清楚他该用什么标准和逻辑来架构通篇，才能“成一家之言”。

我们想了解历史的根本原因，是我们想要了解人的行为是受什么动机、心理和观念影响，而这些行为在历史上产生了什么样的结果。司马迁用本纪来断代并展开该时代的主要政治事件，只是因为方便，书和表在《史记》中的地位都是非常重要的。只读几篇本纪和列传，就无法看到司马迁搭建的整个架构。故事只能传递非常狭窄的历史经验，光听故事不能从历史里学到真正的智慧，被整理的中国传统，要跟我们的生命发生关系。

我深信这些内容对今天的读者很有帮助。司马迁想了我们想不到的，像如何判断一个社会中值得被关注、被记录的人？在帝王将相之外还有什么人是重要的？为什么？刺客重要吗？游侠重要吗？还有，做买卖甚至算命的人重要吗？

什么叫“究天人之际”？“天”是巨大的、集体的和外在的力量，是个人没有办法转移的。一件事情的成功，哪些部分属于“天”的，哪些部分才归于人的努力？他甚至坚持历史内部的一种正义感，甚至可以说是对抗“天”或命运的。在“天”那里没有真正的善有善报恶有恶报，好人常常被误会被冤枉，坏人却被阿谀被美化。如果光依靠“天”，我们要如何坚持做好事当好人？所以，正义必须有人努力。做一个史家，要替那些被误

# 读懂司马迁

杨照



会的人平反，要拨开谎言，更要将许多大家忽略遗忘的善行美德保留下来。

依靠家世的积累，大汉帝国所提供的建设基础，年少时展开的壮游，还有担任太史令能得到的文献资料，司马迁获得了成为一位史家的难得准备，他自豪、他珍惜、他确定自己必须对得起如此难得的机会。那不是个人意义上的难得，而是全体人类经验上的难得。

容我不夸张地强调，古往今来只有这么一个司马迁。而何其幸运，2000年之后，我们竟然可以深切认识这个司马迁，因为他留下了一部内容极其丰富的《史记》。仔细阅读《史记》的完整内容，仔细比对不同篇章的安排与写法，司马迁的思想、观念、人格、执着就穿越时空显影浮现了。

《史记的读法》试图告诉读者有这样一种读法，将作品视为引领我们接近作者心灵的线索，认真看待每一条线索。《史记的读法》试图传递给读者这样一种态度，有时候藏在书里的作者比书本身更重要更伟大。

当然，这只是关于《史记》众多不同读法中的一种。这种读法或许也可以运用在《史记》以外的其他书籍上。（作者为著名作家、文学评论家）

语言战略、语言资源等领域得到重视，2000年通过了《国家通用语言文字法》。

进入21世纪，语言文字学诸多领域与国际前沿同步推进。

语法化、语言接触及地理语言学、构式语法、语义地图模型等都有优秀成果涌现，普通语言学意识大为提升。

语言学界探求语言规则背后的心理机制，重视儿童语言获得规律；心理学界关注语言文字产出和理解机制，重视儿童语文能力的发展过程，二者共同提升了心理语言学的学术地位。脑神经科学，包括核磁共振成像、眼动仪等新技术引入，语言文字的认知心理机制得以进一步还原为生理机制，以不同方式参与人工智能等前沿领域的建设。语言文字研究日益倚重语料库的量化优势，以基因技术结合大数据语料库探求久远的语言谱系关系。

近年来，“语言库藏类型学”“高级版本的意合语法”“新描写主义”等理论探索，已引起国际关注，有望成为中国学者对普通语言学的贡献。（作者为中国社科院语言研究所所长、《新中国语言文字研究70年》主编）



## 《超预期》 超幸福

