

[首页 >> 学术争鸣](#)

对当下数字人文的一些困惑与思考

苗怀明

笔者个人是数字人文的受益者，在近年来的教学科研活动中经常利用数字人文的成果，自己也相继创办并主持中国古代小说网网站和古代小说网微信公众号，深知其中的甘苦，对那些致力于数字人文研发的同道，一直怀有深深的敬意。

数字人文为我们的教学科研工作带来了极大的便利，这是毋庸置疑的，这是未来人文学科学术研究的一个发展方向，且不可逆转，这也是毫无疑问的。但同时还要看到，在数字人文的研发、使用过程中出现了一些让人感到困惑和忧虑的问题，需要引起关注。

以下从不同角度着重谈谈这个方面，为数字人文的研发者提供参考和借鉴。

一、如何克服和纠正数字人文成果使用过程中出现的一些弊端。

年轻学人是数字人文成果使用的主要群体，但在实际使用过程中时常发现，一些年轻学人特别是硕士生、博士生在撰写毕业论文的时候有不当使用数据库的现象，比如他们没有将大量搜索而来的资料进行很好的消化吸收，对文本及背景缺少足够的熟悉和理解，没有经过一个融会贯通的过程，就进行简单的复制和粘贴，论文看似旁征博引，但水准并没有实质的提升，形成新的剪刀加浆糊现象。打个比方，这样的论文有一股浓浓的电脑味。

有不少年轻学人对数字人文成果严重依赖，离开数据库的检索已经无法写论文，而且需要说明的是，这已经并不是个别现象，应引起足够的重视。

二、如何处理数字人文成果与必要的学术基础训练之间的关系。

有了数据库和数字化技术，确实给学术研究带来了很大便利，可以代替研究者做一些耗时耗力的机械劳动，但同时也出现一个问题，那就是如何在大数据时代进行人文学科学术训练，有些看似机械劳动的基本技能比如分段、加标点、编目录索引、校勘等，对初学者来说都是必须亲手去做的，这是进行学术研究的童子功，不可省略，否则达不到进行学术研究所需要的素养和功底，缺少判断力。年轻学子从一开始就使用人文数字化成果，可能会给他们带来一些不好的影响，这是需要引起警惕的。

从目前的情况来看，我们还没有做好面对人文数字化的充分准备，对数字人文带来的改变缺少足够、深入的认识，缺少必要的训练和引导，因此需要从学科角度引起重视，比如进行专门研究，开设课程，编写教材，对年轻学子进行基本的数字人文素养与技能的训练，正确处理学术训练与数字人文的关系。

三、如今人文数字化成果的应用比较普及，但由此也带来一些新的问题，或者说是烦恼，比如信息的庞杂和冗余问题。

众所周知，数据库越来越大，固然意味着获得的信息量越来越大，但其中也有一个让人烦恼的问题，那就是我们通过搜索，得到了很多无效的信息或者说是冗余信息，也就是说存在信息过剩的现象，检索出来以后还要花费大量的时间和精力去筛选。这个问题该如何解决，目前似乎还没有见到特别有效的解决办法。究其根源，是不少数据库的设计存在缺陷，对用户需求考虑不够。

收藏文章



周访问排行	月访问排行	总访问排行
杜甫《望岳》赏析	杜甫《春夜喜雨》赏析	阳光下的罪恶 (3)
柳永《雨霖铃·寒蝉凄切》赏析	杜甫《春望》赏析	杜甫《客至》赏析
宋代家族与文学【第二章】	辛弃疾《青玉案·元夕》赏析	

四、近年来，人文数字化的开发及研究逐渐成为一个学术增长点，但是我们也看到一些让人忧虑的形式主义和严重浪费现象。

时下申报研究项目特别是各类国家及省部级重大科研项目，有一个时髦的做法，那就是在申报书中往往会提到建设数据库，甚至作为一个子课题。但是到了结项时，有不少要么根本就没有建成，要么是比葫芦画瓢建成了，但既没有进行测试，也没有公开发布，成了有头无尾的“烂尾楼”，最后不了了之。这样就造成着数据库的重复开发，造成浪费。

其实在项目申请者中有相当多的人对数据库建设一窍不通，根本不具备这样的知识和能力，他们之所以提出建设数据库是因为这样比较容易拿到项目，说白了，这不过是个敲门砖。至于那些评委们，他们同样对数据库建设所知甚少，也是外行，于是大家心照不宣地走过场，一本正经地相互糊弄。

内行人都知道，数据库的建设需要相当大的资金投入，需要专门的人才，设计、开发、运行及维护等每个环节需要大量人力、资金成本，需要专门的平台，即便项目申请到手，将所有经费都用来建设数据库，也是杯水车薪。

遗憾的是，到目前为止，我们还缺少包括数据库内在的数字人文的准入、审核和评价机制，缺少兼具学术与技术的复合型人才。这种情况应该引起主管部门的高度重视，认真对待数据库建设，避免项目申请过程中以此做敲门砖的形式主义。

还有一些问题，是应该进行深入思考的，数字人文目前还处在探索阶段，很多事情需要有长远的规划和思考，打提前量。笔者认为如下几个问题应引起研究者的关注：

一、人文数字化的广泛使用会不会剥夺研究者进行学术研究的乐趣及情怀。

众所周知，学术研讨的一个重要特点，就是具有鲜明的个性，富有人文情怀，通过研究，研究者获得发现的乐趣，获得成就感。如果我们的研究工作由冷冰冰的搜索及智能技术来完成的话，会不会由此失去研究的乐趣和情怀？

二、当下有不少人从事人文数字化研发工作，数字人文越来越智能化，许多简单的归纳和统计已经可以用机器来独立完成，代替很多人工，这样我们就不能不思考如下一些问题：

人文数字化的终极目标是什么样的？它给人文研究最终会带来哪些改变？这一工作高度智能化之后是否会替代专业学者的研究？对那些研发人员来说，他们现在各自从事的人文数字化的理想状态应该是什么样的？有哪些工作是人文数字化无法代替的，人文数字化目前还存在哪些局限和困难？就现在的研发情况而言，已经实现了哪些目标，还有哪些目标没有实现？

不管情愿不情愿，转眼之间，我们已经猝不及防地走到数字时代，生活工作中的各个方面深受数字人文的影响，但我们并没有做好迎接数字人文的准备，没有形成与之相应的思考方式和生活方式，对其产生的深远影响不管是益处和弊端，都还没有引起足够的重视。

三、就现有的人文数字化成果来看，质量良莠不齐，标准不统一。

当下数字人文的研发，主要以商业性为主，牟利的驱动缺少必要的监管，势必会出现一些乱象。比如逐渐出现的行业垄断、无序竞争现象；不少数据库内容重复，导致用户重复购买；人文数字化成果相互之间不兼容，往往使用一个数据库就要下载一个专门的浏览器，对使用者来说，很是不方便。类似的问题还有不少。

就总体情况而言，目前国内的数字人文化开发和研究还处在各自为政的分散状态，缺少统一的行业标准和管理，各类人文数字化成果之间缺少兼容共享机制，缺少必要的沟通，商业化过于严重。

这就需要政府相关部门高度重视，做顶层设计，制定标准，将人文数字的研发提高到国家文化战略的高度，作为和高速公路、高铁、5G等同等重要的国家基础设施来进行规划和建设，因为数字人文是造福子孙的重要基础工作，关

乎整个民族文化的传承与创新，相当于用数字手段对整个民族的传统文献进行一次全新的保存，意义重大。

总的来说，数字人文不能完全采取商业化的方式运作，应该由国家相关部门出面，进行长远规划，建设一批国家级的数据库，实现全民共享。现在各方面的条件已经成熟，应当有这种紧迫感。

转自“古代小说网”微信公众号

[关于我们](#) | [联系方式](#) | [意见反馈](#) | [投稿指南](#) | [法律声明](#) | [招聘英才](#) | [欢迎加盟](#) | [软件下载](#)

永久域名: literature.cass.cn E-Mail: wenxue@cass.org.cn

版权所有：中国社会科学院文学研究所 京ICP备05084176号 