

国家图书馆古籍信息整理与开发的思考

富平

(全国文化信息共建共享工程办公室)

代表着中华文化底蕴的汉语文古籍是中华民族悠久历史和古代文明的象征，也是国家图书馆宏富收藏中引以为荣的骄傲。但在这种荣耀之中，我们考虑更多的是如何利用知识宝藏造福于广大民众，这也是摆在国家图书馆面前的重要任务。在信息社会，文化信息需求迅速增长的情况下，使我们多了一层思考，那就是更好地整理和开发、传播利用国家图书馆的古籍文献信息。

近几年来，国家图书馆在中文古籍文献数字资源开发、加工、揭示、传播方面作了大量的工作。馆藏善本古籍、普通古籍、金石拓本、舆图等书目数字化工作基本完成，并通过网络向公众进行全面的报道；宋元图书全文数据、敦煌文献全文数据加工工作正在进行；“中国典籍与文化”系列讲座的开办等，都取得较好的社会效益和影响。但随着我国文化市场需求的不断扩大，进一步深层次开发古籍信息，还有更多的事情要做。

一、远期规划与近期目标

国家图书馆馆藏古籍藏品的数字化，是一个长期性任务，应该制定长期规划。但在近期应该作什么事情，要根据资金、人力的实际情况，分阶段去实现。在数字化的过程中，怎样与用户需求结合，把利用率高，用户需要的信息先期加工，通过一定数量的多媒体数字信息吸引力广大用户，是古籍信息开发在起步时，应该把握好的问题，对持续发展很重要。例如：中国有丰富的旅游信息，而且很有特色，国家图书馆可以与各省图书馆合作，开发较丰富地旅游资源。最近我有机会到四川省图书馆考察文化信息共享工程工作，感到四川省图书馆在这一方面做的很有成效。例如：他们与四川省旅游局合作，把四川省内著名的旅游景点的有关信息开发出来，其中包括历史、人物、风景的信息，专门编写出脚本，然后分别撰写文字、拍摄图片等，由四川省图书馆统一制作数字资源库，完成的数字信息在四川省图书馆的网站上供用户使用，同时，把各个景点的数字多媒体信息加工制作成光盘，作为景点门票，使游人在亲身游览景点后，离开后可以进一步了解和回味景点的历史和自己感兴趣的问题。这样做不但满足了用户的需求，也解决了省图书馆前期制作数字资源的经费，同时，也促进旅游事业的发展，为后续发展奠定了基础。四川省图书馆的做法给我们启示是：第一，选好专题，适合用户需求和市场需要；第二，找到合作伙伴，注意与其它行业合作；第三，为持续发展找到立足点。

二、创作精品与市场需求

古籍信息开发与现代信息需求的结合，是古籍信息开发的目标。古籍信息开发相对投入比较大，开发难度大，需要对信息进行深层次加工，对专业人员的要求比较高，因此，在古籍信息开发过程中一定要结合市场需求，满足人们对文化信息需求，这一点要向商家学习。最近，我出差在火车上看到一个幼儿玩具，对我启发很大。一个大约四岁的女孩手里拿了一个手机娃娃，外形是一个洋娃娃造型，但功能与真的手机几乎相同。按其中的几个号码后，玩具发出不同声音，有大人打电话常用的问候语、礼貌用语，让儿童从小知道手机的作用，同时受到文明用语的熏陶；另外按几个号码，玩具发出不同动物声音，可以让儿童辨认，提高儿童对玩具的兴趣。商家结合现实生活的需要，把现代通讯工具与儿童教育紧密结合，受到儿童和家长的欢迎。笔者认为，图书馆工作人员有固定的思维模式，习惯把自己馆在图书馆内考虑问题，不去了解市场。因此，有时花很大力气制作了许多信息产品，但利用率不高，

或者不能得到用户的喜欢。要捕捉市场需求，首先，开发需求型的信息产品，适合用户的口味，为用户定制裁衣；其次，生产引导型信息产品，通过挖掘与整合古籍信息的精华，向公众介绍优秀文化，让公众在潜移默化中接受优秀文化。古籍信息开发的战略，应该是精品战略，任何产品开发，首先要给自己定位，应该是经过认真研究策划，创作为精品，尤其不能粗制滥造。

三、加工方法与遵循标准

传统信息的加工和揭示方法，是以信息载体为单元。而在网络环境下，数字信息的加工与传播是以知识信息为单元。因此，古籍信息的开发，要选择好知识内容，在对古籍信息与数据的处理过程中，要结合现代技术，采用标准的元数据和对象数据的格式与标准。笔者认为，国家图书馆在古籍信息开发加工中应该注意的问题有以下几个方面：

1、古籍书目数据的制作，应该严格遵循中文文献编目规则和 MARC 格式，注意与现代文献的结合，不能一味强调古籍文献的特殊性，否则在网络环境下，将无法满足不同途径的检索的需要。

- (1) MARC 数据中代码字段的选择，一定要准确和标准，不能省略位数，它将对今后数据深层次开发和选择奠定基础；
- (2) MARC 数据中客观描述字段应准确到位，其它题名信息描述应该严格区分，不能相互混淆。
- (3) MARC 数据中的版本项与附加版本说明，应该继承体现传统版本学的学术价值，而不是淡化之。

2、古籍信息的规范问题，是一个值得重视的问题。它除了书目中题名规范、著者规范、地名规范外，更重要的是汉字的规范。揭示古籍信息应该说比现代信息本来难度就大，再加上繁体字与简体字的对照，异体字的问题，使得人们不愿意看。因此，在揭示古籍信息的时候，要从用户的立场出发，加强规范建设。尤其是在网上古籍数字信息的制作中，要有规范控制，帮助用户更好地检索所需信息。国家图书馆应该化较大的力量，加快规范数据库的建设，并解快应用到书目系统和全文检索系统中。

3、古籍数字资源库的制作，在对古籍数字信息重组过程中，要让更多的专业人员参加数据的提炼和标引工作，提高资源库的质量。因为在数字资源库建设中，信息类型包括：文字、图像、视频、音频；内容类型包括：人物、地名、作品、景观、时间、事件、目次资源、相关资源。这些都要求数据加工人员，有较高的专业水平，把握住知识点之间的关系。也就是说在数字资源标引过程中专业知识含量较高，不同专业背景和水平的人员，制作的数据库差异较大。数字资源库加工对传统图书馆而言是一种新的工作流程和技能，要培养新型的加工人才。

4、古籍专题数据库的制作，在文字的表述上，必须要规范的简化汉字版本，同时可以有繁体字版本。有人说，古籍文献和数据库就是为懂古籍的人看的，看不懂人也不会查阅古籍文献和信息，我认为这种想法是片面的，应该让更多的中国人和外国人了解中国优秀的古代文明和古籍文献。另外在内容的揭示上，除了专题较强的研究性内容外，在内容的揭示上，尽可能符合现代人的需要。例如：关于敦煌文献，在揭示的过程中，让人们更多的了解敦煌文献、敦煌艺术、敦煌历史、中国佛教、中国美术等多方面的知识。不可能让所有的人去读敦煌写经。

总之，我们国人肩负着历史的重任，采用先进技术整理与开发古籍信息，传播先进文化，让中国优秀文化为广大的公众所享用。