

内蒙古自治区图书馆古籍保护和古籍工作报告

何远景（内蒙古自治区图书馆）

我区古籍工作开展较早，上世纪八十年代末九十年代初，全区古籍联合目录的编纂工作就开始了，此项工作历时十五年，至二零零四年末，该目录方得以出版。此书收录了全区五十余家古籍收藏单位的古籍五十余万册。著录款目两万余种。

考虑到电脑排版的需要，以及能让更多读者方便利用《内蒙古自治区线装古籍联合目录》，本人自行开发了古籍编目软件。使用这一软件，我们将全区古籍录入电脑，建成了内蒙古自治区线装古籍数据库。

二零零七年末至二零零八年初，我馆改扩建工程开始，所有藏书全部密集堆放，库房封闭；少数珍贵古籍转移至内蒙古考古所的文物库房内保存。由于无法正常提书，古籍工作基本上处于停滞状态。在这种情况下，我只能报告一下我本人所做的工作。

《充分利用已有成果，做好普查工作》

在内蒙古线装古籍联合目录的编纂过程中，我们曾走过不少弯路，做过不少无用工作，浪费了不少时间，几经摸索之后才走上编纂工作的正轨。根据馆领导的指示，我总结了自己在内蒙联目编纂中积累的经验，撰写了《充分利用现在成果，做好古籍保护工作》一文，对于古籍编目工作提出了自己的一得之见。（拙文附后）

学习 ORACLE

内蒙古线装古籍联合目录数据库虽已建成，它是基于 ACCESS 和 EXCEL。目前全国的古籍数据平台业已建成，它的设计基于 ORACLE。鉴于目前图书馆工作的实际状况，馆方不可能为古籍部专门配制一名 DBA，为全面深入地掌握全国古籍数据平台，在日后的古籍工作中更好地使用它。本人特意学习了 ORACLE 数据库技术。

古籍著录总汇

古籍编目过程中经常要参考前人的著录成果。现在通行的古籍目录很多，检索起来很不方便，为解决这一问题，我制作了古籍目录总汇数据库，该数据库将内蒙古线装古籍联合目录，四库全书、续修四库全书、四库存目丛书、四库未收丛书、四库禁毁丛书，丛书集成新编、贩书偶记、明人文集、清人文集目录所收目录汇为一编，共有十四万条数据，使用它可以轻松地检索这十四万数据的著录情况，为古籍检索提供一条便捷方式。

中国丛书综录数据库

《中国丛书综录》是古籍界最常用的工具书。古籍著录总汇虽然也收录了一些丛书子目，但仅为内蒙古线装古籍联合目录中的丛书子目，为方便这一经典古籍目录的检索，近日我又将中国丛书综录中二千七百部中的二千六百种余丛书的子目做成了数据库，

共得古籍数据六万四千余条。古籍著录总汇与中国丛书综录相合可得二十万古籍数据。有了这些数据在日后的古籍编目工作中可大大提高编目的速度和质量。今后如有闲暇，古籍总汇还可以大大扩充，再把尚未收录的古籍目录，如中国善本总目、中国地方志联合目录数据以及中医联合目录数据也补充进去。

难字检索工具

电脑是古籍编目的利器，但它也有与之俱来的缺憾。在古籍编目过程中常常会遇到一些难字不知如何录入的问题。遇到生僻字，使用对于使用以字音为编码基础的输入法，不知读音，无法录入；使用以字形为编码基础的输入法，字形不易准确掌握，亦难以录入。本人在内蒙古线装古籍联合目录的录入中，常常为生僻字的录入大伤脑筋。为解决这一难题，编写了自己独创的难字检索软件，使用它可对 GBK 字库从部首、笔数、拼音、四角号码多个角度进行检索。这一软件从编程角度看尚属草率之作，但在编目工作中起了很大作用，我馆的编目员在遇到难字时常用它来查字。不过这个软件也存在很多问题，主要有几下缺憾：一、四角号码会用者很少，且四角编码有数套方案不尽一致，这就造成取码不准的问题。二、用笔数查字，相同笔数的字太多，对于生僻字确定笔数亦非易事。三、用偏旁找字，对于某些字很确定它的偏旁，不同的工具书有不同的入部规则，在具

体字的入部上，又有很强的随意性，查找不便。为改进难字检索问题，我一直在寻找一种对 GBK 字库较为便捷的检索方式。因为汉字多为合体字，于是就想到把所有 GBK 汉字全部按部件分开做成一个汉字部件数据库，这样不但可对汉字从形傍检索，也可从声旁检索。从 GBK 中把每一个汉字的组成部件都找出来，工作量极大。对 GBK 中 20902 个汉字每个字的每个组成部分都在其中找一遍，其工作量=20902*20902*N，（代表汉字部件的数量的平均值）。这将是一个天文数字。因为工作量大，空有设想无法付诸实施。近日我找到了一种快速查找汉字部件的方法，并为此设计好了工作软件，现已将一万多个汉字的部件提取出来。如果把全部汉字的部件都提取出来，就可以对这种查字方法进行实验。如果实验成功，我相信会做出一个可供各位同仁查找生僻字的工具。

除上述古籍编目工具外，我还做其他些编目常用工具，这些工具主要有：

古籍普查十六表制作工具

古籍普查十六表检索工具

年号纪年、干支纪年与公元纪年换算工具

千字文查找工具

历代职官数据库

这些工具很多古籍同行都在使用或曾经使用，在此不再一一介绍。