

图书馆集成系统

沈 艺 (南京师范大学)

一、引言

图书馆集成系统(Integrated Library System)是图书馆自动化发展的方向。它以能适应多文种标准数据的输入、输出和处理为主要目标,具备书刊管理和情报检索等功能,并具有用户界面友好、通用性强、实用性强、可靠性好和可护性好等特点。

早在六十年代,计算机对于大多数图书馆来说还相当昂贵,美国一些图书馆就开始了图书馆自动化开发、研究工作,其中机读目录格式 MARC(Machine-Readable Catalogue)的设计和联机式计算机图书馆中心 OCLC(Online Computer Library Center)的确立,形成了采编协调、馆际互借、读者服务的综合自动化网络,为图书馆集成系统的发展打下了坚实的基础。

与此相比,我国的图书馆自动化和计算机情报检索起步较晚,进入八十年代才有了较大的发展。目前已拥有各种档次的中、小型机和微机,设计出上百种图书情报管理系统,引进了国外文献数据库近百种,自建中西文数据库近 300 种,并和十多个大型国际联机检索系统建立了联机检索业务。但从总体上看,图书馆计算机管理系统大多是为满足各自的业务工作需要设计,故开放性和通用性差,难以实现信息资源的共享,造成大量的人力物力的浪费,系统也难达到预期的目的。从实践来看,图书馆集成系统不论规模大小,其立足点应该是机读目录和联机信息资源的共享,这样才能保证集成系统全面发挥效能。本文就图书馆集成系统的功能、机读目录及其开放性要求等问题进行讨论和分析。

二、系统功能

集成系统应用计算机的模式有三种:①中小型机多用户集成系统;②微机多用户集成系统;④微机局域网集成系统。中小型机的速度快,容量大,一次性投资较多,技术条件要求高;微机局域网投资灵活,管理方便;微机

多用户介于两者之间。

构成集成系统有三个基本要素:①一个公用的数据库;②覆盖图书馆的主要功能;③联机系统。从系统论的角度来看,集成系统并不是几个单一功能系统的相加,而是有机组合而成的,并通过一个共用数据库互相联系、互相影响的统一整体。

图书馆集成系统的处理对象有两类,一类是图书、连续出版物(如期刊等)及其内容,即信息资源;另一类是信息资源的使用者。从系统功能看,有采访子系统、编目子系统、流通管理子系统、连续出版物管理子系统、检索子系统等。其功能模块如下图所示:

各子系统的主要功能如下:

1.采访子系统

采访子系统是计算机参与处理的图书馆采访业务,采访功能划分为五大部分:订购管理、验收登记、经费管理、统计及报表生成、赠送交换功能。

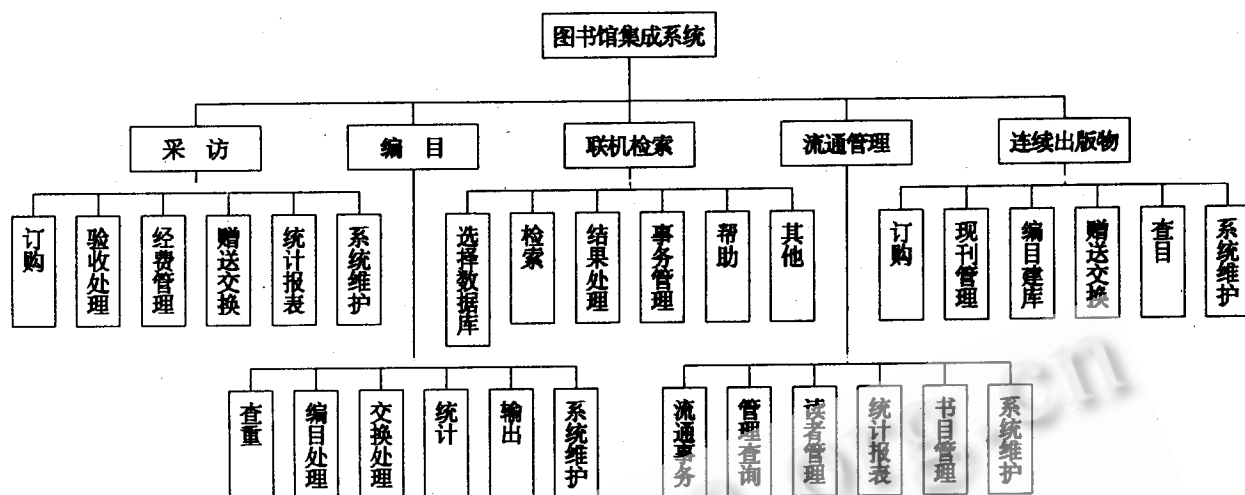
(1)订购管理:此功能主要进行馆藏和正在订购图书的查重、订购数据库的建立等,其中订购数据的来源有两种途径:一是由标准数据源传进采购库,然后加上采购信息数据,这类标准数据有 ANY-BOOK 光盘、USMARC 光盘、CNMARC 磁盘等;二是手工输入;

(2)验收登记:此功能包括验收、个别登记、打印财产帐、总括登记查询、打印总括登记表等功能;

(3)经费管理:有预付款管理、实付款管理、经费使用报告等功能,主要实现控制、管理经费使用情况及有关单据(清款单、发票等),以及与发行者之间财务往来,对所有经费支出作有效记载;

(4)统计与报表生成:包括预订统计、到书统计、接收统计、赠送统计并产生上述各种统计报表,报表内容中有分类统计、文种统计、版式统计、文献类型统计等;

(5)赠送交换:管理交换往来与各种无偿赠送的图书的登记、移交,签定交换关系,并进行管理。



2. 编目子系统

编目子系统的依照机读目录标准及有关规范,建立起中央书目数据库,提供与此过程有关查重、数据输入等功能环境。

(1) 编目查重: 在分编库、总书目库以及其它转换库中, 核查某一本书是新书还是复本;

(2) 建编目库: 在建编目库时, 系统提供两种书目数据输入编辑方式, 一种为全屏幕编辑方式, 即在输入和修改过程中, 信息窗自动显示当前字段的提示信息; 另一种为对话方式, 输入或修改过程中, 屏幕逐个显示字段, 等待输入或修改。

产生标准编目数据的方式亦有两种: 一种是自动编目, 它是编目人员通过各种途径从标准书目数据源(如 USMARC、CNMARC 等) 获得自己需要的记录, 经过修改或增加字段, 添加馆藏信息, 成为本馆的馆藏书目记录; 另一种是人工手原始编目, 凡标准书目源中没有的需要进行原始编目, 即填写原始编目工作单, 校对后输入到计算机;

(3) 编目查询: 提供书目编号、ISBN、分类号、题名、责任者、主题词、索取号等多种查询途径, 以全屏幕方式显示查询结果, 所查数据库包括分编库、总书目库等;

(4) 输出产品有: 卡片目录、书标、书本式目录、CNMARC。

(5) 统计: 种、册统计, 类别统计, 语种统计, 新种统计;

(6) 交接处理。

3. 连续出版物管理子系统

期刊与图书管理有所不同, 期刊订购周期性强, 除新刊订购外, 其经常性的工作是续订、周期性验收和登记, 而其编目工作稳定, 一次编目可以长年使用。但刊名、出版日期等也常有变化, 还常有增刊, 补篇之类的附属信息并有装订等信息。通常对期刊管理采用一条龙管理, 即采、编、登记、流通等业务管理在一个子系统中实现, 故有小图书馆之称。该子系统包括了从订购、催购、装订、编目入藏、检查到流通整个连续出版物的管理过程。

(1) 订购: 包括的内容有查重订购、机读型订购、打印订单、订购统计、订购库备份、订购备份查询等;

(2) 现刊管理: 包括记到、催询、装订、打印装订清单;

(3) 期刊编目: 期刊著录输入、打印著录卡片、装订期刊入藏、馆藏记到、打印期刊登记簿、打印期刊注销簿、统计;

(4) 交换赠送: 包括交换赠送刊查重、交换赠送刊统计等功能;

(5) 检索途径包括: 记录控制号、ISSN 号、统一刊号、订购刊号、分类号、题名、责任者等。

4. 流通子系统

该子系统处理图书外借流能业务, 现都是以条形码作为图书与读者的识别符号, 以便能快速处理借、还等流能业务。其它功能还包括: 事务管理、管理查询、读者管理、统计报表生成与打印、书目管理、系统维护等功能。

(1)借书、续借书:包括对读者的资格审查,凡有过期书、欠罚款,超借书量等情况,系统自动拒借。发现过期或有罚款的,转入相应的处理功能;

(2)还书:还要能处理预约图书;

(3)预约:对已借出的书,其他读者可以预约该书;

(4)查询:包括预约查询、图书借阅查询、读者借阅查询、日志查询等;

(5)读者和简书目库的维护:包括增、删、改等功能;

(6)统计:按日、月、年统计借书量,按类法大类统计,按读者类型统计;

(7)打印:可每天批处理打印过期、罚款、催还和预约到书通知单。

5.检索子系统

检索子系统提供多种有效的联机查询手段,既可查询整个集成系统中的书日期刊数据库,又可查询多种专业的情报数据库。它应提供灵活多样的检索方式,支持用户索引检索、限制检索、组配检索和二次检索,用户可以随意输出检索结果和保存检索结果。检索既可以按菜单提示进行,又可用检索命令方式。命令方式使得检索过程更加灵活。

三、机读目录

机读目录是以代码形式和特定结构记录在计算机存储载体上的,并能够被计算机识别和编辑输出书目信息的目录形式。它是以计算机可读形式同其它国家书目机构以及国内图书情报部门进行书目信息的交换。机读目录是计算机编目后的输出产品。它的一个主要特征,就是它有一套完整的标识系统,计算机依靠这套标识系统来完成书刊数据的处理和转换工作。

机读目录不仅有国际标准(ISO—2709),各个国家也有相应的国家标准。我国的《中国机读目录通讯格式》(China MARC Format)就规定了专著、连续出版物机读形式书目记录的字段标识符、指示符和子字段代码,以及记载在磁带、软盘等载体上的书目记录和它的内容标识符的逻辑和物理的格式。并可逐步扩充到地图、乐谱、视听等其他资料类型。不同的机读目录标准稍有差异,但总体结构是相同的,它们都有一套完整的标识系统,计算机依靠这套标识系统来识别著录字段。下面以《中国机读目录通讯格式》(简称 CNMARC)说明其格式和处理

过程。

1.字段设置

为了使计算机准确地辨识每个著录项目的内容,CNMARC 对每个著录项目设置了一个对应的字段,字段分为两级,第一级是字段,相当于一个著录项目;第二级是子字段,相当于著录项目中的分项。字段的设置,实际上是机读目录数据的安排,表明了机读目录中包含了那些内容。其字段设置数量多,顺序从 001—999,即字段数最多可有 999 个。

2.标识符号

CNMARC 的标识符号用于清楚地标明记录、字段、子字段的含义,以确保计算机辨认机读目录数据的绝对准确性。

字段标识符:用三个数字表示,从 001—999,其中大部分规定了其所代表的著录项目,如 200 字段代表的是题名与责任者项。

字段指示符:与变长字段相关的字符(数据和字母)它提供有关可变长字段内容、记录中不同字段的关系及某些数据处理过程中所需操作的附加信息。

子字段标识符:由两个字符组成的代码,用以识别可变长字段中不同子字段。第一个字符为 ISO-2709 中规定的专用符号 IS1(ISO 646(1/15));第二个字符为字母或数字。如 200 字段中的子字段标识符 \$a 标识的子字段内容为正题名,\$A 标识正题名汉语拼音,\$f 标识第一责任者,\$F 标识第一责任者汉语拼音。

字段分隔符:在每个可变长字段的结尾处用以分隔字段的控制符。它也用于目次区的结尾。该字符是 ISO 2709 中规定的专用字符 IS2(ISO 646(1/14))。

记录分隔符:置于生每个记录的结尾,用来区分记录的控制符,为 ISO 2709 中规定的专用字符 IS3(ISO 646(1/13))。

3.CNMARC 的构成

每条机读目录记录由以下几个部分组成:记录头标区、地址目次区、数据字段区和记录分隔符。

(1)记录头标区固定 24 个字符长。为记录处理提供某些基本参数,如:记录长度、指示符长度、子字段标识符长度、数据基地址等;

(2)地址目次区由若干个目次项组成。有多少个数据字段就有多少个目次项,每个目次项固定为 12 个字符

长,其中字段标识符(3个字符),字段长度(4个字符)和字段起始字符位置(5个字符)。各目次项按字段标识符排序。目次区最后由一个字段分隔符(IS2)结尾。数据字段区由一些可变长数据字段组成。除001字段和005字段由数据和一个字段分隔符组成外,其余每个字段都有两个指示符,后接若干子字段。每个子字段以一个子字段标识符(IS1和一个小写字母或数字构成)开始,后接可变长数据。每个字段都以一个字段分隔符(IS2)结尾。每个记录都以一个记录结束符(IS3)结尾。

每条机读目录所著录的字段数不尽相同,字段的取舍,取决于转换为机读形式的具体记录。记录的数据内容,由编目条例和负责建立记录的机构实际执行的规范确定,即数据元素的有无,不仅由格式的要求确定,而且由国家编目条例和实际执行的规范确定,但以下几个字段是必备的:

001—记录标识号

100—一般处理数据(仅为某些数据元素)

101—作品语种(当作品有语种时)

200—题名与责任者

801—记录来源字段

4.机读目录的应用

机读目录使信息资源的共享成为可能,从理论上讲,每一本书只要有一个图书馆估和了机读目录,其它所有的图书馆就可以享用此标准数据,从而避免了大量的重复劳动。机读目录主要有以下几方面的应用:

(1)选书。以机读目录记录为依据,选出适合自己馆藏的图书资料,编制成“待入藏文档”供选书使用;

(2)建立图书订购文档。在国家图书馆享有的“在版编目”优先权得到保证后,其MARC的发行要比一般的图书要早,那么使用MARC的单位可利用MARC记录建立本馆采购文档。如果MARC的某一条记录合乎需要,就可以把这条记录的内容加入到订购文档,而不必重复编制;

(3)编目。将MARC记录的格式转换成适合自己计算机系统使用的格式,成为自己的编目成果。这可有两种方式实现,一是通过记录有机读目录的磁带或磁盘;二是借助联机协作编目系统,如OCLC就有此功能。利用磁带或磁盘更方便快捷;

(4)开展定题书目情报服务。

5.MARC处理过程

MARC处理应该作为“图书馆集成系统”的一个支持环境,其处理过程主要是由创制机读目录的一系列技术和操作方法来构成,从功能上看,MARC处理过程由以下四个方面加以实现:

(1)MARC记录的输入。收集MARC数据,填写工作单,经校订后输入到计算机内;

(2)MARC记录的维护。进行MARC记录的存贮管理、追加、修改、删除;

(3)MARC记录的检索。利用计算机检索,把使用者的提问与数据库中的资料进行逻辑区配运算,得到所需要的结果;

(4)MARC记录的输出。即输出各种专题目录、馆藏书目目录、目录卡片、MARC记录的磁带磁盘数据等。

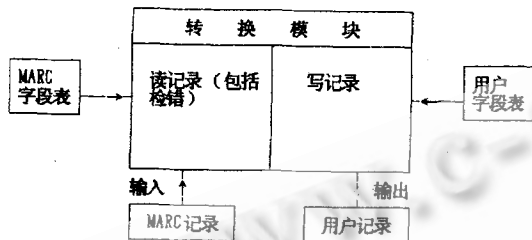
四、开放性要求

一个系统只有是开放的,才能实现“资源共享”。图书馆“资源共享”的重点不在于硬件资源共享,而应在于信息资源共享。MARC仅定义了书刊编目数据的一个标准格式,大家在数据交换时必须依照这个标准格式,而各个图书馆的集成系统在实现书刊管理和情报检索诸多功能时,不一定都采用这个变长的数据格式,可以根据自己的实际需要设计各个内部处理数据库的格式。但是,作为对图书馆集成系统的开放性要求,系统中应设计有将MARC数据与集成系统内部处理的数据格式相互转换的接口。这个接口的设计应有两个层次:一是通过磁盘或磁带处理MARC数据;另一个是通过网络(包括局域网和广域网)处理MARC数据。

仅把MARC的磁盘或磁带作为数据交换的手段,其时效性差、效益低,而利用计算机与通信技术相结合的计算机网络系统,为MARC的交换提供了极大的方便。这可以在联合采编的基础上建立公共的中心数据库和地区网络,逐步创造条件实现各馆之间的远程通信和数据共享。

该接口的第一层是将MARC磁盘或磁带数据转换成本馆集成系统所能直接处理的格式,或将本馆的编目数据转换成MARC的磁盘或磁带格式。具体地说是MARC记录所具备的字段或子字段与本馆集成系统中

编目数据库所处理的字段间的相互转换。即如下图所示:



通过网络这一层实现资源共享,对于那些正准备购置计算机设备的图书馆,则应选择符合开放系统要求的计算机系统;对已有小型机的图书馆,为了实现合作网,可考虑应用“综合计算环境”;对于拥有高档微机的图书馆,因其已具备较好的开放基础,可考虑与本地区有条件的馆互连成广域网。

在实现网络化的过程中,一个十分重要的问题是网络体系结构标准化。计算机网络在 70 年代得到迅速发展,许多计算机厂商都相继推出它们的网络产品,提出了各自的网络体系结构。例如,IBM 公司的“系统网络体系结构”SNA(System Network Architecture);DEC 公司的“分布式网络体系结构”DNA(Distributed Network Architecture)等等。随着各种网络不断涌现,迫切需要解决异种机异型网络互连,以达到相互信息交换、资源共享、分布处理和分布应用。国际标准化组织 ISO 经过多年的努力,于 1984 年 10 月提出了“开放系统互连基本参考模型”——ISO/OSI(Information Processing System—Open System Interconnection Basic Reference Model),尔后又陆续制定出各层次的标准化协议。

图书情报服务网格在遵守标准化协议的基础上采用

以下模式:①单机直接互联。主机及终端采用单一规程,通过同步或异步(终端)方式直接与分组交换网互联。如以 PC 为终端、IBM4381 为主机,互联时,主机或分组终端通过 CCITT 的 X.25 协议与分组交换网的集中器或节点机直接互联,异步终端通过 X.28 协议进网,既可以专线接入,也可以通过电话线拨号接入。这种方式是国外八十年代初、中期普遍采用的方式,我国现阶段也主要采用这种方式。在这种方式下, MARC 数据的交换通过单机完成,再通过第一层的 MARC 转换接口,实现与本馆工作数据的互换。②通过局域网(LAN)与公用数据网互联。这种方式避免了每个单机都解决与共用数据网的接口问题。③通过程控电话交换机(PABX)与公用数据网联网。与现代现代办公自动化潮流中的局域网相比,采用 PABX 构成局部范围的集中型网络系统,实现声音、数据综合通信,更加经济方便,目前的 PABX 不仅可以和各种主机、终端、局域网互联,还可以通过 X.25 接口与分组交换网互连。

五、结束语

从实践来看,通过编程形成图书馆集成系统各种处理个功能并不是十分困难的,难的是如何使系统功能发挥作用,并实现图书馆的全面管理。因为单凭一个馆或几个馆是无法将全部书刊数据快速、正确、全面地录入到计算机中,这些标准化的数据是集成系统的基础。因此在建设图书馆集成系统时,不能只考虑到一个馆简单管理功能的实现,而要考虑到如何适应图书馆的开放环境。以上是对几个影响较为深远的问题发表了一些个人的见解,供参考。