

编目工作的挑战及应对措施

王松林



图书馆的地位与作用随着信息资源普遍可获得性的程度提高在不断下降。但就像20世纪40年代末，收音机、电影和电视曾作为当时的“高新技术”对人们的阅读习惯和图书馆造成巨大的挑战一样，以上种种数字化和网络化，也不会挑战到置图书馆于死地的地步。而可能置于死地倒是图书馆内部的某些机构或人员，如图书馆的编目部和参考咨询人员。目前在我国，这一点尤为明显。

报告内容

- 1 编目工作面临的挑战
- 1.1 编目业务外包
- 1.2 电子在版编目与中文图书标识
- 2 编目人员的应对措施
- 2.1 树立信心
- 2.2 技术创新

1 编目工作面临的挑战

分类编目历来是图书馆的核心业务；今天的国际文献联合会，1895年创建时的名称为“国际书目协会”；而1929年改为现名的国际图联，1927年成立时的原名为“国际图书馆与目录委员会”。

We are Google without garbage.

我国图书馆编目工作所面临的挑战，我认为现阶段主要来自于编目业务的外包，但今后还将来自于电子在版编目等的出现。

1.1 编目业务外包

源于企业的业务外包，是指企业为了将有限的资源和精力集中在能使企业保持竞争优势并能为客户提供独到的产品和服务的核心业务上，将企业内部那些无法胜任、经济效益低或不占优势的业务剥离出去，并让外部专业厂商承包经营的模式。

1.1.1 国外图书馆编目业务外包的发展及其特点

- ① 国外图书馆编目业务外包的发展
- (1) 萌芽阶段（20世纪初-1970年代）
- (2) 形成阶段（20世纪80年代-90年代初）
- (3) 发展阶段（20世纪90年代初-90年代中）
- (4) 普及阶段（20世纪90年代中-）

2) 国外图书馆编目业务外包发展的特点

- (1) 编目业务外包已成为一种发展趋势
- (2) 编目业务外包的内容多元化
- (3) “即可上架”的编目业务外包
- (4) 优化图书馆的管理模式

1.1.2 我国图书馆编目业务外包的动因分析

1) 研究视角分析

对采编业务外包的探讨，国内几乎所有的论者都是从图书馆的自身发展，尤其是从技术环境的角度而不是从组织环境或制度环境的角度来看采编业务外包问题。其实，图书馆作为一个公共组织，外包也是一种典型的组织行为，但我们目前的研究恰恰没从组织研究的视角来探讨图书馆的业务外包，这不能不说是一种缺憾。

2) 适应型动因分析

我国高校图书馆采编业务外包的深层动因，我们认为包括以下三个方面或者说可归纳为以下三种因子：第一种因子是为适应扩招和教育部本科教学评估的适应性因子；第二种因子是降低作业成本，加快服务结构转化的效率性因子；第三种因子是响应反商业贿赂，整肃图书采购行为的法治性因子。从徐月华对广州高校图书馆采编业务外包动因的调查可以看出，各外包馆的外包动因基本都是这三种因子的不同程度的复合。

1.2 电子在版编目与中文图书标识

根据原始文献的输入方式，编目专家系统可以分成以下两种类型：一是将普通纸质文献（主要是文献题名页）通过阅读装置输入系统，一是与电子出版系统相连接，文献直接以电子形式输入系统。

相比较而言，在计算机和网络环境下，直接与电子出版系统相连接的编目专家系统更易实现图书的自动编目。

1.2.1 中美两国ECIP的发展

1) 美国ECIP的发展

```
<tp>beginning of title page
</tp>end of title page
<sp>beginning of series page
</sp>end of series page
<cp>beginning of copyright page
</cp>end of copyright page
<toc>beginning of table of contents
</toc>end of table of contents
<ch1>beginning chapter 1
<ch2>beginning chapter 2
etc.
</ch>end of last chapter
```

2) 国内ECIP的探索

《中文图书ECIP与自动编目手册》一书出版后，国内在对ECIP与自动编目深入探索中，认识到只有在图书排版过程中同时实现印刷版与数字版的制作，ECIP与自动编目才能得到有效处理，从而逐步形成数字复合出版的概念。而《国家“十一五”时期文化发展规划纲要》和《新闻出版业“十一五”发展规划》的《国家数字复合出版系统工程》产生，为ECIP与自动编目提供了实施的条件。

1.2.2 《中文图书标识规则》的主要内容

1) 《中文图书标识规则》的引言

在“引言”部分,《中文图书标识规则》说:

本标准的目的在于推进编辑出版工作的标准化。在图书生产过程中,创建适应不同需求的图书元数据,以满足出版、发行、收藏利用、信息服务等机构对中文图书标识数据的制作和管理需要,实现图书生产、流通和使用环节对图书元数据的共享。

2) 《中文图书标识规则》的范围

在“范围”部分,《中文图书标识规则》说:

本标准规定了图书产品需要标识的数据内容和相应的标识规范。

本标准适用于依法经国家新闻出版行政管理部门批准设立的出版者在图书生产过程中对数据元素的标识。

3) 《中文图书标识规则》的术语和定义

在“术语和定义”部分,《中文图书标识规则》对中国机读目录格式、CNMARC 字段、CNMARC 子字段、DC 元数据、DC 元素、DC 修饰词、图书标识、数据元素等8个术语进行了定义。

4) 《中文图书标识规则》的标识数据的结构与描述

《中文图书标识规则》的“标识数据的结构与描述”部分共由以下三部分组成:

- (1) 标识数据的构成
- (2) 标识数据的描述结构
- (3) 数据元素的描述

5) 《中文图书标识规则》的标识数据

在“标识数据”部分,《中文图书标识规则》对中文图书中的每一基础数据元素和补充数据元素,按上4)的(3)数据元素的描述格式进行了描述。如正书名:

所在位置:扉页

CNMARC 标识名称: 200\$a

DC 元数据标识名称: T,Proper

数据类型: 字符型

使用说明: 必选, 可重复

注: 图书的主要书名,包括交替书名。同一作者有多个书名且无总书名时,重复以本数据元素标识。

6) 《中文图书标识规则》的附录A

《中文图书标识规则》的附录A“中文图书标识数据使用 CNMARC 字段、子字段一览表”是一资料性附录,主要栏目有“CNMARC 字段名称”、“字段指示符”、“子字段代码”、“数据元素名称”、“可否重复”、“DC 元数据”和“修饰词”。

1.2.3 书目数据的产生与利用

1) 书目数据的产生

前述扩充修饰词的DC元数据结构化标签，通过与CNMARC字段、子字段的对照，在图书未付印前就可生成ISO 2709文本上报给CIP中心和《新华书目报》编辑部，并用于本社印前业务管理；而图书付印时则可生成正式书目送交给CBIP网站发布，并据以转换为二维码印刷在封底，当然同时也可用于本社的图书与信息的管理。

2) 书目数据的利用

过去，中文图书的在版编目数据主要通过《中国图书在版编目快报》及其光盘版获取。而现在，ECIP中的书目信息不仅可以通过《新华书目报》和CBIP网站发布，而且转换为二维码的书目信息，图书馆只需一个解码器就可避免重复编目，即图书馆只需在二维码提供的图书信息基础上添加其有特殊要求的内容就可完成本馆的目录编制工作。

2 编目人员的应对措施

图书馆职业在历史上一直承担着人类知识有序化的工作，只是由于知识生产的批量化和工业化，缺乏必要的工具，在速度和深度上无法根据知识的内容进行整序，而只能根据载体形态进行组织加工；数字图书馆技术的发展将有可能借助语义技术和许多新的工具，使图书馆行业数千年来积累起来的知识组织经验，在万维网时代发扬光大。

2.1 树立信心

面对信息技术以及全文检索技术的迅速发展，Michael Gorman对书目控制的未来发展提出如下三点建议：①加强书目控制的地位。图书馆读者服务部门的馆员与编目馆员要相互了解与尊重。书目控制应为图书馆学图书馆教育的核心课程，以此来引导所有的图书馆员了解书目控制的架构与目的；②提倡合作编目。推导全球书目控制的观念，鼓励编目资源的全球共享；③图书馆编目员的工作范围应扩展到电子资源。即对电子资源予以适当的分类与编目，也图书馆员的任务。

Michael Gorman认为，根据电子资源本身的价值，其书目控制的方法有以下四种级次可供选择：

- ①完全根据国家和国际标准编制成完整的MARC记录；
- ②增进型DC“编目”；
- ③最低限度的DC记录；
- ④让搜索引擎做。

除此之外，关于图书馆编目工作的发展，一是可以参考本人去年在《中国图书馆学报》上发表的《图书馆实体信息资源组织的两大发展路径》一文。该文提出图书馆扬搜索引擎之长、避搜索引擎之短的实体信息资源的组织理论与方法，即扬搜索引擎之长，使实体信息资源组织章节化；避搜索引擎之短，使实体信息资源组织FRBR化。真正做到“与狼共舞”。

二是可以参考2008年1月9日由美国国会图书馆发布的《书目控制的未来报告》（Report on the Future of Bibliographic Control）。该报告的任务是要考察在不断发展的信息和技术环境下，书目控制和其他描述实践能有效地支持图书馆资料的管理与检索、建议图书馆界集体向目标前进的方法，以及向美国国会图书馆建议其所发挥的作用和工作重点。

2.2 技术创新

2006年1月，美国北卡罗莱纳州立大学图书馆发布的新版OPAC，虽然仍然使用原来的书目记录，但其检索界面已经改用Endeca公司具有引导导航（guided navigation）能力的ProFind作为检索平台，并同时提供检索与浏览功能。检索返回结果类似于搜索引擎，缺省方式是按相关性排序，另可选择按出版日期、题名字顺、作者字顺、索书号排序，尤为特别的是可以按流通数据获取的、受欢迎程度的排序。依据检索结果，在左端一栏列出用作限定检索的相关主题，如流派、载体、地区、时代、语种、作者及馆藏地等。

在国内，2006年5月25日，“Web2.0与信息服务”会议在上海图书馆召开。此次会议为国内图书馆2.0关注者的第一次聚会，同时也是图书馆2.0热潮中极为重要的一次学术研讨。会前会后形成的图书馆2.0热潮几乎吸引了图书馆界各层人士的关注。该会议主要涉及“Web2.0与信息服务”、“Web2.0与电子商务”、“新型OPAC”等。其中编目精灵所做的关于OPAC改进的报告尤其让与会者对图书馆2.0的未来充满希望与信心。

读秀知识库简介

刘炜发出这样的呼吁：掌握利器、决胜未来。这些利器包括：

①将传统编目中属于“主题编目”的一类技能和成果，也就是目前统称为“知识组织系统”（包括分类法、叙词表等的KOS）的规范体系，转化为Web可以读懂和处理的格式，因为传统书目系统或情报检索系统采用这些概念体系都是给人用的，不具有严格的机器编码，机器对这些词表和分类法并不可读。

②为进行更多的主题编目或开发主题编目算法，编目员要研制各类本体，以及研发各种支持SKOS（简单知识组织系统）和OWL（网络本体语言）的KOS编目工具，并从事资源的主题标注、批处理或互操作开发工作。

③可能被称之为“元数据编目员”的工作可能更多的是进行元数据标准规范（领域应用的元数据规范一般称之为“元数据应用纲要”）及著录规则的制定、流程设计、培训推广、质量控制等，需要支持自动或批量的元数据编目。另外由于所有的标准规范或方法工具最终都将体现于元数据编码中，所以制订、审核并应用各类元数据编码方案（主要是基于RDF/XML的编码方案）也应该是这个环节中的重要内容。

在数字图书馆谈论有多少个TB时，世界数字信息已进入PB、EB、ZB时代。美国趣味科学网站举例说，一首MP3歌曲的存储空间约为5MB。如果以每分钟1MB的速度不间断播放，1ZB能够存储的歌曲足以让人听上19亿年。而据“数字世界”调查预测，未来10年，全球信息总量又将是现在的44倍。而推动这一增长的强劲动力，仍是与图书馆相关的媒介形态的持续变革，即电视、广播、印刷行业将不断实现模拟信号向数字信号的转变。

谢谢！