



Date : 26/06/2008

第74届国际图联大会  
加拿大魁北克，2008年8月10-14日  
书目组公开会场

## 数字时代的加拿大国家书目

作者: **Liz McKeen**

资源描述部  
出版遗产科  
加拿大图书/档案馆

中文翻译: 韩玲 (中国国家图书馆)

Chinese translator: HAN Ling (National Library of China)

**Meeting:** 162. Bibliography

**Simultaneous Interpretation:** English, Arabic, Chinese, French, German, Russian and Spanish

### WORLD LIBRARY AND INFORMATION CONGRESS: 74TH IFLA GENERAL CONFERENCE AND COUNCIL

10-14 August 2008, Québec, Canada

<http://www.ifla.org/IV/ifla74/index.htm>

## 导言

早上好，我很高兴能在这里代表加拿大国家图书/档案馆讨论数字时代的加拿大国家书目。

我将和你们一起探讨在数字时代，国家书目在概念和实践上发生的变化。我将着重讨论加拿大的经验，我们加拿大人正走向对国家书目从范围到本质上的重新认识，我们也日益认识到与他人合作在制作国家书目过程中的重要性，并且我们也正在为这种合作创造机会。

## 数字出版物在加拿大图书/档案馆(LAC): 迄今为止的故事

### 电子出版物的文献呈缴制度

在加拿大，文献呈缴制度从2007年元月起就涵盖了电子出版物。这对我们来说意义重大，但却也不完全陌生。加拿大国家图书馆——加拿大图书/档案馆的前身之一，从1994年起就开始收集电子出版物，我们已经收集的电子出版物不多但很有趣，它们被编成大约3万个书目或者说405GB的数据。这些书目大多数都是没有版权或者是皇家版权的资料，我们可以把这些资料放在我们的网页上，向

全世界开放。他们被适时地按照《英美编目条例》或其他的标准加以编目，并列于国家书目之中。

虽然理论上这个法律在对出版物的定义上包含了网站、数据库、地理空间数据和很多其他类型的电子资料，但是出于实际的考虑，目前我们将呈缴制度的应用范围缩小到“类似于”传统出版物的数字出版物上，即他们都采用了印刷版专著和连续出版物的类型和外观。呈缴制度也要求出版者为LAC提供可以描述其出版物的元数据。

## 网页收割

2005年，LAC开始试验使用网络爬虫Heretrix从整个网站世界中获取资料，至少从每个网站中截取一幅快照，这些快照保持了网页从网络中被截取出来时的形态。我们从加拿大联邦政府机构网站入手，2006年对它们有了一个初步的“爬行”，现在有一个固定时间来“爬行”联邦政府网站——一年两次。我们已经完成了对加拿大10个省和三个准州政府部门的爬行，并准备在定期的基础上做得更多。

2007年11月，“加拿大政府网络档案”诞生，并于网上向公众开放。我们强调以这样一种方式设计界面——清楚而明确地告诉用户，这是一个存档的版本，而不是一个可用的版本。比如说，我们不希望加拿大的纳税人在加拿大联邦政府的存档网页上交他们的个人所得税。

加拿大政府网络档案包含约1500个网站的很多更新和快照。这些网站的内容是什么？我们的第一次爬行以.gc.ca为域名的网站就网罗了4000个数字目标

（2TB数据），约90万个PDF文件。后来的每一次爬行都网罗了更多的数字目标。在对联邦政府网站的最后一次爬行中，我们捕捉了7000万数字目标和3TB数据。他们是出版物吗？他们是连续出版物？他们是公民为获得政府服务而填写的表格？他们是每被截取一次快照都是一次复制的专著？他们是新的出版物？

## 馆藏的海量数字化

虽然LAC曾对个别馆藏项目进行过数字化，如地图等。若干年后的2007年，LAC承担了一个系统地将其所藏进行海量数字化的项目。这个项目目前仍处于起步阶段，在设计一个全面的战略计划的同时，我们还承担着许多数字化项目，这些项目中的很大一部分要与其他机构合作完成。这些机构包括加拿大公报（加拿大官方政府出版物机构）、国会议事录（国会事务的文字记录）、加拿大年鉴（由加拿大统计机构出版）、和许多关键的人口普查项目（谱系学专家们对此很感兴趣）。

数字化面临很多挑战，包括资金、管理、如何挑选进行数字化的资料、与数字化专业人员的协作、对版权的尊重、设计一个有效的实际操作方法应用于内部和外部的重大项目。数字化的一个主要要求是：将有效的元数据与所有的数字化书目联系起来，以使用户能够找到他们想要的东西。

## 数字出版物

总之，加拿大数字出版物的数量是巨大的，并且这一数量在不断地增长。这种出版物最为紧跟时代脚步，它们反映了国家的精神状况和创造力，是加拿大人思想的重要记录。数字出版物正日益取代更多的传统物。它不能被忽视。

### 数字出版物的编目

图书馆传统的、一条一条地对出版物进行编目是一个昂贵的程序，这种编目方法很难承受处理数字出版物这一重荷。

2005年，LAC举办了一个“描述日”，大约100名来自图书馆和档案机构的工作人员参加了这个活动。其中一个自然组成的组，侧重于研究怎样“对网络进行编目”。这个组很快得出结论：由于包含的数目巨大，它不能用传统方法解决。2006年，LAC提出了一个编目政策建议，这个提议我们觉得基本足以包含巨量的新式数字出版物，同时它在访问原则上立场坚定。这个“数字出版物的资源描述”可以在LAC网站上获得。

### 数字出版物的资源描述：政策

简要地说，认识到数字书目数量的巨大，和对合作与协作的实际需要，这个政策提出了一系列解决方法。

1 基础——通过全文索引和搜索向所有数字出版物提供基础检索。

2 补充——通过提供的元数据提供补充检索。但是元数据的制造通过一系列可行的方法在成本有效的方式下进行，

2.1 第一个选择：使用他人提供的元数据：出版商、其他图书馆、其他机构、作者、学生，等等……基本按原样描述，只要像MARC的样子就可以了。

2.2 第2个选择：使用电子出版物自己自动生成或提炼出的元数据。这种人工智能的方式还没有完全被开发出来，可能还需要进一步的研究。

3.3 第三种选择：使用LAC员工为采访目而创建的描述。这些记录通常足够确认一条书目，但是并不经常包括检索优化，例如主题、分类号、相关人名规范记录。采访记录可以使用他人或自动生成的元数据来创建。

4.4 最后一种选择：使用国际标准(MARC, AACR, ISBD)，通过优化检索来创建一个完整的书目记录，例如双语主题标目、杜威和LC分类法、规范标目、所有可用的附注，等等。

这一选择是最为昂贵的选项，它将被留给很少一部分从海量数字书目中挑选出来特殊对待的书目。通过提供一种标准的书目描述，LAC可以突出这一文件，让它更为显著和易于检索。

我们什么时候使用“完整书目记录”这一选项？我们决定哪个出版物获得完整检索处理的标准如下：

- 记录了加拿大重大的或主题性的事件、趣闻、事故、经历（例如：选举、SARS危机、自然灾害等等）的文献
- 馆藏发展方针中所定义的LAC的特藏或重点馆藏的文献。
- 作研究或参考之用的文献。

- LAC员工请求制作完整书目记录的文献。
- 被归入图书再版编目等特殊书目计划的文献。

这些标准仍然比较粗略，因为它们还没有在实践中被细化。

### 3 数字化了的（文献）

数字化了的文献大多数情况下被从LAC的馆藏中挑选出来用于复制。在很多情况下，一条优质的书目记录为原始文献而存在。这种数字出版物的资源描述政策建议数字化了的版本简单地加入到已经存在的原始文献的记录中。这种单一的记录方法与标准著录规则所要求的、每种格式要求一个单独的记录并不一致，与国家书目通常的要求也不一致。但是，LAC看到，作为一种性价比高的方法，它通过创造简化了的书目记录和最新的“一站式”搜索体验为终端用户带来了显而易见的便利。

## 对国家书目的影响

### 国家书目

20世纪70年代，由国际图联大会提出的世界书目控制概念是以这样一个观念为前提的，即每个国家图书馆或国家书目机构都有义务将本国的所有出版物登记在一个国家书目中。所有的国家一起努力，一个全面的世界文献书目就将成为可能，起码在理论上是可能的。世界学术、知识共享、观念传播、文化的丰富和人类的进步都将获得巨大的成果。

此外，一个完整的国家总书目（包含回溯的和现行的）在每个国家中不断增长，国家总书目是国家的一个窗口，浓缩而又代表性地反映了过去这个国家的文化、表达与思想。这样一个现行的国家总书目将大大提高研究人员和普通读者发现和得到这个国家的著作的能力。

在加拿大，现行的国家书目——*Canadiana*开始于1953年，每月出版一期，数年后*Canadiana*在范围和覆盖面上变得越来越完整。同时，员工们花了很多时间编辑回溯性的国家书目，在时间上这一书目一直回顾到1755年，就是印刷术被介绍到加拿大的那一年，这一年的一个多世纪前加拿大这个国家才刚刚形成。今天的*Canadiana*被书目记录所涵盖，位于在线的AMICUS——我们以网络为基础的书目系统——之中，并且每年以CD的形式出版，包含大约200万条记录。

### 国家书目和数字出版物

当我们考虑到自然和纯粹的数字出版物规模，我们可以很快地认识到，一个完整的国家书目的传统概念——将一个指定国家的所有出版物罗列其上——可能是不可能的并且是不需要的。

在数量上，很多原因造成数字出版物比传统出版物多很多并将继续多下去。比如，网上出版物成本低，风险小，并且大大减少了编辑工作的基础设施，因而它的读者群更大。以上特点还导致数字出版物中能出现更加民主的出版物（比如，维基百科）、更加多样的出版物、更加短促的出版物（如博客和电子邮件）、以及随之而来的、越来越多样的作者和其他创造者。

在民族性上，我们很难将网站分派给一个特殊的地理政治学实体或者一个民族。当然，这种信息在网站上通常并不明确。网站天生对全世界平等地开放。

## Canadiana: 数字出版物资源描述政策的影响

在加拿大，LAC对数字出版物资源描述政策的采纳将通过很多方式潜在地影响国家书目。

**全面性：**直到目前为止Canadiana一直将成为一个完整的书目列表作为自己的目标，这个列表包括过去和当下加拿大所有的出版物，同时也包括加拿大作者在海外出版的书目、研究加拿大的书目和与加拿大主题有很大关系的书目。在新政策下，以我们之前提到的选择标准为基础，国家书目中一个与新政策相关联的、较小的数字出版物集合将被确认。其它的加拿大数字出版物将不被列入其中。所建立的规范人名标目也将少于完全版的。

**明确标识：**Canadiana为每一条书目的版本和出版形式提供单独的列表。依照新的政策，数字化了的版本将被添加到原始文献的书目记录中，因而创造一个反映两种出版物形式的记录。

**权威性：**尽管Canadiana仍然是一个权威的参照源，但它可能将不再是所有加拿大出版物唯一的、权威的来源，就像它自己将不再完整和明确一样。

## Canadiana的意义

LAC基于如下的理由绘制了蓝图：

1) **平衡：**LAC认可了数字出版物的价值——加拿大思想和创造力的传达工具。当下的编目政策更加看重的是传统出版物而非数字出版物。这种偏见可能不会被纠正并且随着时间的流逝变得越来越不具有可行性，因为数字出版物正变得越来越重要。在数字出版物源描述政策下，LAC致力于取得一个在加拿大书目覆盖面上的平衡，而不管它是以何种形式存在。

2) **杆杠技术：**LAC正在寻求使用技术手段以便通过全文索引、描述的自动或半自动化生成、数字出版物本身的新特征等手段来为数字出版物提供检索。这种方法性价比高，并且有潜力提供非常好的检索。

## 合作

世界图书馆之间的合作模式拥有漫长而光辉的历史。传统上，LAC与其他图书馆一样，让创建的书目记录在使用上达到最大化，并且尽可能地与其他图书馆共享。LAC是加拿大出版物和其它加拿大书目的主要创建者和分享者。LAC也支持由国会图书馆协调的合作编目计划，为高质量的书目和相关数据、ISSN网络和其它合作的最大化共享提供框架。

基于以各种形式出现的数字出版物的范围和规模，LAC和很多其它图书馆正寻求合作模式来达到合理的数字覆盖面。但是，数字时代的合作模式虽变得更重要却更无组织。LAC对寻求获得、利用各种形式的元数据的方法越来越感兴趣，这些元数据不仅是图书馆创建的，也可以是其它人用不同的方法创建的，例如出版商、作者、学生、图书从业者、对此感兴趣的个人和各个领域的专家等等。

## 来自于学生的元数据

### 硕士论文和博士论文

我们从学生那里得到论文摘要，并将它们添加到书目记录中去，这一过程持续了若干年。摘要可以被全文检索，并且可以通过主题提供进一步的检索。最近几年，我们使用OAI从特定的加拿大大学中收割电子论文，连同论文一起的是学生们制作的、与之相联的元数据。这些元数据自动地传送和添加到AMICUS——我们的在线编目系统中，成为一条包含着一些非标准内容的MARC记录。

## 出版者的元数据

### 可靠的数字库

LAC正在开发一种虚拟码头，作为可靠数字库的组成部分，它允许加拿大数字出版物的出版商存储这些（以数字形式出版的）文献，并且通过网络形式提供这些书目的元数据。这些元数据将被传送到一个元数据库中作提供检索用。

LAC也希望TDR系统包可以提供一种技术方法从数字对象中自动生成元数据。这种功能还没有被开发出来。

## ISBN元数据

LAC已经开发出一种自动化的ISBN系统提供给ONIX形式下书目数据的输入和输出，这种系统可以为编目列表把ONIX转化成MARC记录。ONIX记录将被传送到LAC的公共编目系统中，从而记录员可以将此作为基础添加数据和标准。

## 来自个人的数据

LAC 曾用几个为获取个人用户知识而设计的项目做过试验。个人提供的数据需要被确认，因为原则上来说这并不权威。尽管这些数据很不标准，但对于那些提供它的读者来说仍具有独一无二的价值，那些提供者可以为我们对馆藏资料的理解提供一个非常个人的侧面。

## 战争中的脸

在一个叫“战争中的脸”的计划中，网站浏览者被要求对数字化照片进行评论，并且如果可能的话，对照片中的个人进行辨认。这个计划可以变得非常个人，因为浏览者可以从另一个年代中辨认他们的祖父或年轻的朋友。这些评论经员工们简单校阅后公开发表，但注明它们是个人的意见。

计划命名：LAC收藏的描述加拿大北部人的图片，包括了从19世纪晚期到20世纪中期的数千幅照片。但是，在拍摄时，照片中的因纽特人中只有很少一部分人的身份得到了确认。努纳武特（Nunavut，加拿大最北边的三块领土之一）人从来没有机会确认照片中的这些人，因为馆藏地离他们的社区很远，在数字化之前，没有任何方法可以轻松地将这些照片传到努纳武特。LAC在2001年创造了包含数字化照片的网站，并且让因纽特的长者和年轻人来辨认照片中描绘的人。确认这些仍然匿名的人是一件很有时间紧迫感的事情。今天的长者可能是最后一个

可以辨认他们的人，那些照中人的名字可能仍将永远被丢失。这一计划由LAC联合努纳武特Sivuniksavut以及努纳武特文化、语言、老年和青少年部门开展，并且还包括了青少年参观社区这一项目。

硝酸盐照片底片（仍在开发的工程）：LAC有一个具备相当规模的硝酸盐底片馆藏，但它们没有任何相应的元数据存在。这一尚在研发中的计划意在将这些历史照片数字化，并且开发出一个网络界面让用户可以给这些底片添加确认信息。我们将从网站浏览者那里寻求有关照片中人物、地点、时间、描述物的信息。这些元数据将被自动编辑成一个由用户生成的“寻求帮助”，在给以适当警告的情况下，这些数据将被其他用户下载、修正和改进，就像维基百科那样。

## 结论

我和你们通过加拿大经验，共同探讨了数字时代国家书目在概念和实践上产生的一些变化。加拿大的经验尽管很新、还带有实验性，但仍将带我们走向对国家书目在范围和性质上的重新定义。我们正日益认识到与他人合作在制作国家书目过程中的重要性，并且正在为这种合作寻找机会。我们希望以新的方式合作，包括发掘加拿大个体的知识，无论它是专家的见解还是个人的、趣闻轶事般的贡献。通过我们的数字出版物资源描述政策（如果机会允许我们将加大其贯彻力度），我们正尝试认识和适应数字出版和数字检索带来的现实情况。我毫不怀疑，在未来的岁月中，这些早期的经验将被大大扩充、深化，形成在这个世纪早期对国家书目机构的潜在价值和义务更加丰富的理解。

一个国家的出版物可以被认为是公民思想、态度、文化和创造力的浓缩。真实完整地反映一国的出版现状，并且保证全世界可以得到这一国的出版物，仍是国家书目机构的重要使命。结尾还是那样：创造这种反映（国家出版物现状）并将其呈现在世界面前的方法将是一个共享的经验，并在未来被数字时代塑形。