



Date : 03/07/2008

## Канадана – национальная библиография Канады – в цифровую эпоху

**Лиз Маккин**

Директор,

Отдел описания ресурсов,

Отделение опубликованного наследия,

Библиотека и Архивы Канады

*Translated by: Alexandra Teplitskaya,  
Russian State Library, Moscow, Russian Federation*

**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>

### Введение

В докладе проанализированы изменения, произошедшие в цифровую эпоху, в концепции и практике национальной библиографии. Я остановлюсь на опыте Канады. Мы сейчас переосмысливаем возможности и характер национальной библиографии, мы осознали, что надо искать партнеров для совместного создания национальной библиографии.

### Цифровые публикации в Библиотеке и Архивах Канады (LAC): предыстория

#### Обязательный экземпляр электронных публикаций

В Канаде электронные публикации вошли в состав обязательного экземпляра в январе 2007 г. на основе закона об обязательном экземпляре. Это было важно, но не было чем-то совершенно новым, т.к. национальная библиотека Канады, одна из предшественниц Библиотеки и Архивов Канады, собирала электронные публикации с 1994 г. Наша небольшая, но интересная, коллекция насчитывает приблизительно 30 тысяч названий или 405 гигабайтов данных. Часть названий находится под защитой авторского права, часть нет. Они доступны для всех на нашем веб-сайте. При их каталогизации были использованы Англо-американские правила каталогизации и другие стандарты. Все электронные публикации отражены в национальной библиографии.

Теоретически по закону в состав ОЭ входят веб-сайты, базы данных, картографические данные (geospatial data) и многие другие цифровые материалы, попадающие под определение «публикация». Но на практике мы ограничили состав ОЭ только цифровыми публикациями,

аналогичными с традиционными, которые имеют форму и внешнее сходство с печатными книгами и сериальными изданиями. По закону издатели должны присылать в LAC какие бы то ни было метаданные, необходимые для описания их публикаций.

### Сетевой сбор (web harvesting)

В 2005 г. LAC, используя программу web crawler Heretrix, предприняла эксперимент по сбору в интернете как можно большего количества сайтов. Во всяком случае, был сделан моментальный снимок каждого сайта, который был доступен в день проведения эксперимента. В 2006 г. мы начали обследовать (crawling) веб-сайты отделов канадского федерального правительства. Эту процедуру мы проделываем два раза в год, и теперь мы имеем полный перечень выявленных (собранных) веб-сайтов федерального правительства. Мы также обрабатываем веб-сайты министерств десяти канадских провинций и трех территориальных правительств. Мы планируем в дальнейшем увеличить объем работы и улучшить ее качество.

“Сетевой архив правительства Канады” стал доступен общественности в интернете в ноябре 2007 г. Интерфейс устроен таким образом, что пользователь сразу понимает, что перед ним архивная копия, а не живой сайт; мы не хотели, чтобы налогоплательщики, просмотрев сайт канадского агентства по доходам, например, отправились платить налоги.

«Сетевой архив правительства Канады» содержит большое количество повторов или снимков приблизительно 1 500 веб-сайтов. Каково содержание этих веб-сайтов? Наш первый сбор (crawl) домена .gc.ca насчитывал около 40 млн. цифровых объектов (2 терабайта данных), что составило приблизительно 900 000 PDF файлов. С каждым новым обследованием мы собираем все больше и больше цифровых объектов. Последний раз, обследовав веб-сайты федерального правительства, мы собрали свыше 70 млн. цифровых объектов и 3 терабайта данных. Это является публикациями? Это выпуски сериальных изданий? Является ли это формой получения гражданами правительственной информации? А монографии, которые мы копируем каждый раз, когда делаем мгновенный снимок, что это? Это новые публикации?

### Массовая оцифровка фонда

В течение нескольких лет LAC оцифровала отдельные коллекции произведений, карт и т.д. В 2007 г. мы приступили к реализации программы по систематической массовой оцифровке своих фондов. Эта программа еще только набирает силу, разрабатывается общая стратегия, но уже насчитывается большое количество проектов по оцифровке, многие из которых реализуются на кооперативной основе, например, с Canada Gazette (официальный печатный орган правительства), Hansard (официальный вестник парламента), Canada Yearbook (издается статическим органом Канады), плюс множество проектов по переписи, которые вызывают интерес у специалистов по генеалогии. Проблем, возникающих при оцифровке, много. Это и финансирование, и управление, и выбор материала носителя, и сотрудничество с другими организациями, которые занимаются оцифровкой, и соблюдение авторских прав, и консолидация внутренних и внешних ресурсов, нацеленных на повышение эффективности массовой программы. Основное требование к оцифровке – это точное соответствие метаданных с оцифрованными объектами, т.к. пользователи должны найти то, что они ищут.

## Canadiana in the Digital Age

### Digital Publications at LAC

- e-Publications and selected Websites
  - Legal deposit of e-publications, January 2007
  - e-collection of 30,000 titles, 405 gigabytes
- Web harvesting
  - 1,500 Canadian federal department websites
  - Canadian provincial and territorial websites
- Mass Digitization of LAC collection

### Цифровое издательское дело

Вывод: массовое издание в цифровом виде в Канаде все возрастает. Издание в таком виде получило широкое распространение и отражает менталитет и творческий потенциал нации, и является важной формой фиксации памяти нации. Цифровое издательское дело все больше вытесняет традиционные формы. И это нельзя не учитывать.

## Canadiana in the Digital Age

### Digital Publishing

- Enormous quantities
- Current
- Reflects & records the nation's thought
- May supplant traditional publishing formats over time

## Каталогизация цифровых публикаций

Традиционная библиотечная каталогизация это достаточно дорогой процесс, который нелегко реализовать в цифровом виде.

В 2005 г. LAC провела “День описания”, в котором участвовало около 100 ее сотрудников, работающих и в библиотеке и в архивах. Разговор шел о нововведениях в описании. В одной из спонтанно образовавшейся группе активно обсуждались проблемы каталогизации сети. Эта группа быстро пришла к заключению, что из-за огромного количества объектов их нельзя каталогизировать традиционным способом. В 2006 г. LAC выработала цели каталогизационной политики, которые были достаточно радикальными. Они устанавливали баланс между огромным количеством имеющихся цифровых публикаций и принципами доступа. “Механизм описания цифровых публикаций” доступен на веб-сайте LAC.

### Механизм описания цифровых публикаций: политика

Учитывая огромное количество названий цифровых объектов и прагматическую необходимость осуществления их описания на корпоративной основе, было выработано много подходов к решению данной проблемы.

## Canadiana in the Digital Age

### Resource Description for Digital Publications: Policy

LAC policy on cataloguing of digital publications (2006)

Partially implemented 2008; future implementation depends on systems functionality

Available on LAC Website

1. Главный – Основной доступ ко всем цифровым публикациям осуществляется на основе полнотекстового индексирования и поиска. В идеале этот вид поиска должен быть доступен пользователям через общественный интерфейс, отличительным признаком которого является высококачественный, сложный механизм поиска, ранжирующий результаты так, как это надо пользователям.

## Canadiana in the Digital Age

### Resource Description for Digital Publications: Policy

- Basic access : full-text search

2. Дополнительный – Должен быть обеспечен дополнительный доступ через метаданные. Создание метаданных должно быть рентабельным:

2.1. Первый вариант: использование метаданных, созданных другими: издателями, библиотеками, институтами, авторами, студентами и т.д. Описания можно использовать в той или иной степени до тех пор, пока они соответствуют формату MARC.

2.2. Второй вариант: использование метаданных созданных автоматически или извлеченных непосредственно из самих цифровых публикаций. Эта форма искусственного интеллекта пока еще полностью не исследована; может потребоваться время для ее дальнейшего изучения.

2.3. Третий вариант: использовать описания, созданные сотрудниками LAC в процессе комплектования. Эти записи можно использовать для идентификации, но они обычно не имеют таких точек доступа как предметные рубрики, классификационные индексы, связанных с авторитетными файлами имен. Записи, создаваемые в процессе комплектования, могут быть составлены с использованием метаданных, созданных другими или автоматически.

2.4. Последний вариант: для создания полного библиографического описания с увеличенным количеством точек доступа, таких как двуязычные предметные рубрики, индексы классификаций Дьюи и Библиотеки конгресса, авторитетные заголовки, все возможные примечания и т.д., надо использовать международные стандарты (MARC, AACR, ISBD).

Этот вариант – самый дорогой, и его можно использовать применительно только к тем названиям, которые были тщательно отобраны из всего массива цифровых публикаций. Если на документ составляется полное библиографическое описание, то этим самым LAC будет подчеркивать его значимость и важность.

## Canadiana in the Digital Age

### Resource Description for Digital Publications: Policy

- Supplementary access, in approximate order of preference:
  - Metadata supplied by others
  - Metadata automatically generated from digital object
  - Metadata created for other purposes (acquisitions)
  - Full, standard bibliographic records

В каких же случаях мы будем составлять «полную библиографическую запись»? Мы предлагаем для обсуждения следующие критерии отбора названий:

- Названия тех документов, в которых отражены оригинальные или актуальные события, эпизоды, инциденты, имеющие отношение к Канаде (например: выборы, природные катаклизмы и т.д.).
- Названия, предназначенные для специальной или приоритетной коллекции LAC, если это определено Политикой развития фонда.
- Названия, необходимые для ведения исследовательской и справочной деятельности.
- Названия, на которые, как считают сотрудники LAC, должны быть составлены полные библиографические записи.
- Названия, которые подпадают под специальные библиографические программы, такие как Каталогизация в издании.

Выше перечисленные критерии детально еще не проработаны, т.к. не накоплен достаточный опыт работы.

## Canadiana in the Digital Age

### Resource Description for Digital Publications: Policy

How will we select digital titles for full cataloguing?

1. Titles that document seminal or topical events, experiences in Canada (elections, SARS crisis)
2. Titles that are part of an LAC special or priority collection

## Canadiana in the Digital Age

### Resource Description for Digital Publications: Policy

How to select digital titles for full cataloguing?

3. Titles that are research or reference tools
4. Titles for which cataloguing is requested by LAC staff
5. Titles falling under special bibliographic programs (Cataloguing in Publication)

3. Оцифровка – В большинстве случаев оцифровываются названия из фонда ЛАС с целью репродукции. Часто высококачественные библиографические записи на оригиналы изданий уже есть. В «Механизме описания цифровых публикаций: политика» предлагается присоединять оцифрованную версию к существующей записи на оригинал. Такой подход – существование единой записи – не отвечает требованиям создания отдельных

библиографических записей на каждый формат, как это оговорено в стандартах, правилах каталогизации, национальной библиографии. Но LAC считает, что рентабельнее создавать упрощенную библиографическую запись, что позволяет пользователю рационально осуществлять поиск, при котором ему выдается обобщенная информация.

## Canadiana in the Digital Age

### Resource Description for Digital Publications: Policy

- Digitized titles – digitized versions of originals in the LAC collection will be added to the bibliographic record for the original
- Single record approach

#### Воздействие на национальную библиографию

#### Национальная библиография:

Концепция универсального библиографического учета, четко сформулированная ИФЛА в 1970-е годы, строится на том, что каждая национальная библиотека или национальное библиографическое агентство несет ответственность за учет в национальной библиографии всех публикаций своей страны. Теоретически, если все страны будут сотрудничать, полный учет мировой литературы возможен. Это предоставляет огромные возможности для развития мировой науки, выработки всеобщего знания, распространения идей, обогащения культур.

Исчерпывающая ретроспективная и текущая национальная библиография, создаваемая в каждой стране, это окно в государство; национальная библиография в сжатом виде отражает национальную культуру, мысли, которые будут доступны последующим поколениям. Текущая национальная библиография позволяет исследователям и просто читателям найти и получить доступ к публикациям данной страны.



## Canadiana in the Digital Age

### National Bibliography

- Universal Bibliographic Control
- Comprehensive listing of the world's literature, country by country
- National bibliography: a window on a nation
- Enhanced access through comprehensive listing, unambiguous identification of titles

В Канаде текущая национальная библиография берет свое начало с 1953 г., когда стали выходить ежемесячные выпуски Канадианы. За прошедшие годы охват литературы и качество самих выпусков значительно улучшились. Одновременно сотрудники занимались созданием ретроспективной национальной библиографии, начиная с 1775 г. – время появления печатной продукции в Канаде. Это на столетие раньше, чем образовалось само государство Канада. Сегодня библиографические записи Канадианы доступны в режиме он-лайн в AMICUS (нашей сетевой библиографической системе) и в виде ежегодно издающегося CD-ROM, который содержит около 2 миллионов записей.

#### Национальная библиография и цифровые публикации:

Осознавая характер и объемы цифрового издательского дела, мы понимаем, что традиционная концепция всеобъемлющей национальной библиографии, учитывающей все публикации страны, становится нереальной или даже нежелательной.

Цифровое издательское дело по разным причинам количественно значительно превосходит традиционное книгопечатание, и в дальнейшем этот разрыв будет только увеличиваться. Например, публикации в интернете доступны большему количеству людей, их себестоимость ниже, рисков меньше и практически нет затрат на издательскую инфраструктуру. В результате издательское дело становится более демократичным (например, wikis), появляется больше разнообразных качественных публикаций, больше однодневок (например, блоги и электронная почта) и все больше и больше различных авторов и других создателей.

По национальному признаку трудно отнести тот или иной веб-сайт к определенной геополитической принадлежности или национальности. Веб-сайты по своей природе одинаково доступны всем, независимо от того, в какой части света человек находится.

## Canadiana in the Digital Age

### National Bibliography & Digital Publications

- Quantity of digital publications (electronic serials & monographs, Websites, email, wikis, blogs, Facebook, etc.)
- Nationality of digital publications unclear

#### Канадиана: влияние «Механизма описания цифровых публикаций: политика»

В Канаде внедрение LAC «Механизма описания цифровых публикаций: политика» оказало большое влияние на национальную библиографию.

Полнота: До сегодняшнего дня *Канадиана* стремилась к исчерпывающему учету (ретроспективному и текущему) всей канадской печатной продукции, также как и работ канадских авторов, опубликованных за пределами страны; выборочно учитывались работы, посвященные Канаде или представляющие интерес для Канады. С учетом новой политики только небольшое количество цифровых публикаций как таковых будут учитываться в национальной библиографии. Это связано с критериями отбора, о которых говорилось выше. Другие канадские цифровые публикации не будут учитываться в национальной библиографии. Авторитетный файл имен также будет менее полным.

Четкая идентификация: В *Канадиане* ведется отдельный учет всех изданий и форматов каждого названия. В соответствии с новой политикой сведения об оцифрованных версиях будут отражаться в основной библиографической записи оригинала. Таким образом, будут создаваться записи, в которых указаны все формы реализации публикации.

Авторитетность: *Канадиана* может стать авторитетным справочным источником (но не единственным) для всех канадских публикаций, т.к. больше нет так долго существующего полного и четкого источника.

## Canadiana in the Digital Age

### Resource Description for Digital Publications Policy: Implications

- **Less able to approach aim of comprehensiveness**
- **Unambiguous identification of titles (expressions, manifestations) not always possible**
- **Authoritativeness as an information source could be compromised**

#### Выводы по Канадиане

LAC выбрала этот курс по ряду причин:

- 1) **Баланс** – LAC признает ценность цифровых публикаций, как средство распространения мыслей и творчества канадцев. Политика текущей каталогизации больше ориентирована на традиционные публикации, чем на цифровые. Такой вектор развития нельзя одобрить. Через некоторое время, когда значимость цифровых публикаций еще больше возрастет, этот путь приведет в тупик. «Механизм описания цифровых публикаций: политика» LAC ориентирован на достижение баланса охвата канадских публикаций, независимо от их формата.
- 2) **Прогрессивная технология** – LAC стремится использовать технологические возможности для обеспечения доступа к цифровым публикациям, посредством полнотекстового индексирования, автоматического или полуавтоматического составления описаний, а также особенностей самих цифровых публикаций. Такой подход рентабелен и может обеспечить хороший доступ.

## Canadiana in the Digital Age

### Resource Description for Digital Publications Policy: Implications

- Will achieve better balance in coverage of traditional and digital publications
- Leverages technology to assist in providing access
- Affordable

### Сотрудничество

Сотрудничество библиотек мира имеет давнюю и прекрасную историю. Традиционно LAC, подобно всем библиотекам, по возможности максимально полно использует библиографические записи, созданные другими библиотеками. LAC является главным создателем и участником совместной каталогизации канадской печатной продукции и других изданий, отражаемых в Канадиане. LAC также участвует в Программе совместной каталогизации, координируемой Библиотекой конгресса, которая преследует цель максимально полно использовать возможности совместной каталогизации для создания высококачественных библиографических и вспомогательных данных, сети ISSN и других продуктов.

Учитывая сферу влияния и масштаб цифрового издательского дела во всех его формах, LAC, как и многие другие библиотеки, пришла к выводу о необходимости объединить усилия для отражения цифрового пространства (в разумных пределах). В эпоху цифровых технологий создание модели сотрудничества становится просто необходимо, но в то же время в организационном плане это мало выполнимо. LAC все больше и больше заинтересована в нахождении и использовании различных форм метаданных, созданных не только библиотеками, но и другими: издателями, авторами, студентами, книжными посредниками, частными лицами, работающими в разных областях и имеющих разный опыт.

## Метаданные от студентов

### Дипломные работы и диссертации

В течение нескольких лет мы получаем рефераты на дипломные работы, написанные студентами, и присоединяем их к библиографическим записям; на основе реферата можно осуществлять полнотекстовый поиск; реферат также повышает качество предметного поиска. В последние годы для сбора дипломных работ в электронном виде, защищенных в некоторых канадских университетах, и связанных с ними метаданных, создаваемых самими студентами, мы используем OAI (Инициатива открытых архивов). Эти метаданные автоматически собираются и добавляются в AMICUS – нашу библиографическую систему в режиме он-лайн – в виде записи MARC, которая не совсем отвечает стандартным требованиям.

## Метаданные от издателей

### Доверительный (trusted) цифровой репозиторий

LAC развивает Виртуальный загрузочный док, который является частью Доверительного цифрового репозитория, позволяющий канадским издателям цифровых публикаций организовать их хранение совместно с LAC и передавать их метаданные по сети. Затем эти метаданные будут переданы в репозиторий метаданных для обеспечения доступа к этим публикациям.

На основе системного пакета TDR LAC также планирует развитие технологии, позволяющей автоматическое создание метаданных непосредственно самим цифровым объектом (если это получится). Пока что этого нет.

## Метаданные ISBN

LAC развивает автоматическую систему ISBN, которая обеспечивает импорт и экспорт библиографических данных в формате ONIX, с возможностью конвертировать ONIX в записи MARC, для включения их в библиографический ряд. Записи ONIX будут передаваться в LAC для Программы Каталогизация в издании (CIP), и каталогизаторы смогут их использовать в качестве базовой записи, улучшать и доводить до уровня действующих стандартов.

## Canadiana in the Digital Age

### Collaboration:

#### Compiling the National Bibliography by Using Metadata Supplied by Others

- Students' metadata on their thesis or dissertation
- Publishers' metadata
  - When depositing legal deposit e-publications in Trusted Digital Repository (Virtual Loading Dock)
  - With ISBN request

#### Данные от частных лиц

LAC осуществляет несколько проектов по аккумулярованию знания индивидуальных пользователей. Данные, поставляемые частными лицами, не являются авторитетными. Несмотря на то, что эти метаданные не отвечают стандартам, но они имеют уникальную ценность и ориентированы на читателей, которые могут высказать свое мнение к выработанным нами критериям отбора материала.

Проект «Лики войны»: В проекте «Лики войны» посетителей веб-сайта попросили прокомментировать переведенные в цифровую форму фотографии и по-возможности идентифицировать людей, запечатленных на фотографиях. Этот проект может стать очень личным, т.к. посетители, например, могут опознать своего дедушку и его молодых друзей, живших в другую эпоху. Комментарии быстро просматриваются сотрудником, затем выставляются на общественное обозрение, при этом делается ссылка, что это комментарий частного лица.

## Canadiana in the Digital Age

### Collaboration:

- Metadata from Individuals
  - Faces of War <http://www.collectionscanada.gc.ca/faces-of-war/index-e.html>
- Example of Comment on Photograph:

My father Abe G..... is the second from the right. He arrived in france July 6 /1944. Dad married a british gal named Joan Mary W..... They had 4 children , Frances, Gary , Wayne and myself Derek..Dad rarely talked of the war but I know it changed him in some ways . however if he had not been there , I would not be here. War has many strange twists .  
Dad passed away 1981.My deepest thanks for having this picture for me to find online . Thank you.  
Derek

Проект «Назови имя»: Коллекция фотографий LAC с изображением людей севера Канады насчитывает тысячи фотографий, датированных концом 1800-х годов – серединой 20-го столетия. Однако, очень немного людей Inuit, изображенных на этих фотографиях, были опознаны за то время, как фотографии были выставлены. Люди Nunavut (одна из трех канадских территорий, расположенных на далеком севере) никогда не имели возможности помочь в опознании, т.к. эта коллекция фотографий находится очень далеко от их общины, а до оцифровки коллекции не было средств доставить фотографии в Nunavut. Веб-сайт с оцифрованными фотографиями был создан LAC в 2001 г., и мы попросили старшее поколение и молодежь Inuit идентифицировать людей, изображенных на фотографиях. Процесс опознания тогда еще анонимных людей был очень эмоциональным. Люди старшего поколения могли стать последними, кто смог бы опознать людей из прошлого, их имена могли исчезнуть навсегда. Этот проект осуществляется LAC в сотрудничестве с Отделом культуры, языков, пожилых людей и молодежи Nunavut Sivuniksavut и Nunavut. С его помощью предпринимается попытка вовлечь молодежь в дела общины.

Нитрат (?) фотографических негативов – Nitrate Photographic Negatives (предполагаемый проект): LAC располагает большим собранием старых нитратов негативов, на которые нет никаких метаданных. Предполагается оцифровать эти исторические фотографии, разработать сетевой интерфейс и попросить пользователей прокомментировать их. Нам нужна информация об изображенных людях, местах, событиях, объектах. Эти метаданные будут автоматически компилироваться в генерируемую пользователями базу данных, с соответствующими комментариями, которые другие пользователи смогут откорректировать, исправить (как это делается в Wikipedia).

## Canadiana in the Digital Age

- Metadata from Individuals
  - Project Naming  
<http://www.collectionscanada.gc.ca/inuit/index.html>
  - Nitrate negatives (project under development to obtain user-generated metadata)

### Заключение

В докладе проанализирован канадский опыт организации национальной библиографии в цифровую эпоху, обозначены изменения, коснувшиеся ее концепции и практики. Накопленный нами опыт пока невелик, полученные результаты носят предварительный характер, но это заставляет нас задуматься о переосмыслении возможностей и характера национальной библиографии. Мы все больше осознаем необходимость поиска путей сотрудничества с другими организациями, частными лицами по созданию национальной библиографии. Мы хотим сотрудничать на новых условиях, в т.ч. использовать знание частных лиц, которое может быть научным, личным или носить анекдотический характер. Через наш «Механизм описания цифровых публикаций: политика», который по мере наших возможностей набирает оборот, мы пытаемся осознать и приспособиться к реалиям цифрового издательского дела и цифрового доступа. Я не сомневаюсь, что спустя некоторое время, этот первоначальный опыт будет значительно расширен и углублен и приведет национальное библиографическое агентство к более полному пониманию своих возможностей и обязательств, чем это было в начале 21-го столетия.



## Canadiana in the Digital Age

### Conclusion

- Redefining scope & nature of national bibliography
- Seeking collaboration with new partners
- Seeking to obtain new types of metadata
- Adapt to realities & opportunities of digital

Изданную печатную продукцию нации можно рассматривать как квинтэссенцию мыслей, отношений, культуры и творчества граждан страны. И в будущем национальные библиографические агентства будут продолжать играть важную роль в обеспечении правдивого и полного отражения этой продукции, гарантированного доступа всего мира к публикациям нации. В то время как цель остается постоянной, средства ее достижения будут обогащаться совместным опытом, приобретаемым в ходе реализации возможностей, предоставляемых цифровым будущим.

## Canadiana in the Digital Age

### Conclusion

- Publications are distilled thought of a nation
- Continuing role of national bibliography– to provide a true and full reflection of a nation's published output, and to ensure access to it for current and future generations