

Национальный файл географических названий — новый проект РГБ

National File of Geographic Names — a New Project of Russian State Library

Національний файл географічних назв — новий проект РДБ

Лавренова О. А.

Российская государственная библиотека, Москва, Россия

Olga A. Lavrenova

Russian State Library, Moscow, Russia

Лавренова О. О.

Російська державна бібліотека, Москва, Росія

Доклад представляет собой презентацию проекта национального файла нормативных/авторитетных записей для географических названий, разработка которого начата РГБ в 2003 году по договору с Министерством культуры Российской Федерации в рамках подпрограммы «Развитие культуры и сохранение культурного наследия» федеральной целевой программы «Культура России (2001-2005 годы)». Файл войдет в состав комплекса файлов нормативных записей для ЛИБНЕТ и призван обеспечить использование в библиотечной практике (при каталогизации документов и поиске информации) нормализованных наименований российских географических объектов, установленных официально на федеральном уровне. Российский национальный файл географических названий получил международный MARC-код «rugeo» для использования его в библиографических записях. Разработан технический проект, определены источники данных, подготовлены программы-конверторы, сформирован экспериментальный файл.

This paper presents a project of the national file of regulatory/authority entries for geographic names that was started in RSL in 2003 under the agreement with RF Ministry of Culture in the frame of sub-programme «Development of Culture and Preservation of Cultural Heritage» in the Federal Dedicated Programme «Culture of Russia (2001-2005)». The file will make a part of a folder of regulatory entries for LIBNET and is aimed to ensure the application of standardised Russian geographic names, as officially approved on the federal level (for the cataloguing of documents and retrieval of information). The Russian national file of geographic names got an international MARC code «rugeo» that should be used in bibliographic entries. A technical project is developed, sources of data are identified, converting software is prepared and a trial file is made.

Доповідь є презентацію проекту національного файлу нормативних/авторитетних записів для географічних назв, розробку якого почала РДБ у 2003 році за угодою з Міністерством культури РФ у межах підпрограми «Розвиток культури та збереження культурної спадщини» федеральної цільової програми «Культура Росії (2001-2005 роки)». Файл стане частиною комплексу файлів нормативних записів для ЛІБНЕТ та забезпечить використання у бібліотечній практиці (при каталогізації документів та пошуку інформації) нормалізованих найменувань російських географічних об'єктів, які впроваджено офіційно на федеральному рівні. Російський національний файл географічних назв отримав міжнародний MARC-код «rugeo» для використання його в бібліографічних записах. Створено технічний проект, визначено джерела даних, підготовано програми-конвертори, сформовано експериментальний файл.

1. Основные положения

Проект национального файла географических названий разрабатывается в России впервые в интересах всего библиотечного сообщества по инициативе и силами РГБ при поддержке Министерства культуры РФ (есть надежда и на поддержку нового министерства). Автор благодарит рабочую группу сотрудников РГБ, с которой ведение проекта оказалось столь успешным и интересным.

Национальный файл нормативных/ авторитетных записей для географических названий — организованный массив нормативных/ авторитетных записей для названий географических объектов на территории Российской Федерации. Сокращенное наименование — национальный файл географических названий (или файл географических названий ниже в тексте). Используемая аббревиатура — НФГН.

Нормативная запись для географического названия — машиночитаемая запись нормативного описания, составленного для географического названия.

В НФГН включаются нормативные записи для «наименований географических объектов Российской Федерации, континентального шельфа и исключительной экономической зоны Российской Федерации, а также географических объектов, открытых или выделенных российскими исследователями в пределах открытого моря и Антарктики, если иное не предусмотрено международными договорами Российской Федерации» (Российская Федерация. **Федеральный закон** от 18 декабря 1997 г. № 152-ФЗ. «**О наименованиях географических объектов**», ст.3, п.1). Данный закон устанавливает **правовые основы деятельности** в области присвоения наименований географическим объектам и переименования географических объектов, а также нормализации, употребления, регистрации, учета и сохранения наименований географических объектов как составной части исторического и культурного наследия народов Российской Федерации.

В соответствии с этим Федеральным законом «О наименованиях географических объектов» (ст.8, п.2): «В документах, картографических, иных изданиях на русском языке или на других языках народов Российской Федерации употребляются **нормализованные наименования географических объектов**». В данном законе, подчеркивается необходимость «установления и использования единых стандартизованных названий географических объектов как в официальных документах, так и в картографических и иных изданиях».

В связи с этим принимается следующее **основное требование к файлу** нормативных записей для географических названий: нормализованные (принятые) заголовки нормативных записей должны соответствовать нормализованным заголовкам, установленным на федеральном уровне и приведенным в «**Государственном каталоге географических названий России (ГКГНР)**».

В проекте решено использовать определения основных понятий, данные в Федеральном законе:

- **«географические объекты** — существующие или существовавшие относительно устойчивые, характеризующиеся определенным местоположением целостные образования Земли: материки, океаны, моря, заливы, проливы, острова, горы, реки, озера, ледники, пустыни и иные природные объекты; республики, края, области, города федерального значения, автономная область, автономные округа; города и другие поселения, районы, волости, железнодорожные станции, морские порты, аэропорты и подобные им объекты;
- **наименования географических объектов** — географические названия, которые присваиваются географическим объектам и служат для их отличия и распознавания;
- **установление наименований географических объектов** — выявление существующих наименований географических объектов, присвоение наименований географическим объектам и переименование географических объектов;
- **нормализация наименования географического объекта** — выбор наиболее употребляемого наименования географического объекта и определение написания данного наименования на языке, на котором оно употребляется» (Федеральный закон, ст.1).

Уместно отметить, что в определенной степени задачи нормализации наименований географических объектов в области присвоения этих наименований и переименования объектов вполне сопоставимы с задачей нормализации используемых в сфере библиотечной деятельности элементов библиографических данных. Это обстоятельство, кстати, представляет собой дополнительное доказательство правильности выбора в качестве перевода английских терминов «authority file» и «authority record» терминов «нормативный файл» и «нормативная запись».

Сопоставимость задач составляет основу весьма продуктивного процесса использования данных из ГКГНР, в особенности в его электронном варианте, для создания нормативных записей. По сути, различия заключаются в форматах представления данных и, в некоторой степени, в составе их элементов.

2. Функции файла географических названий

Файл географических названий создается в интересах пользователей электронных каталогов, электронных библиотек и других автоматизированных информационных систем (АИС), в частности — автоматизированных информационно-библиотечных систем (АИБС).

Файл географических названий выполняет следующие функции:

- обеспечивает **использование нормализованных наименований российских географических объектов (российских географических названий)** в географических заголовках библиографических записей (БЗ), предметных рубриках, классификациях, рубрикаторах и других типах метаданных для автоматического контроля их использования и поиска документов; поддерживает процессы каталогизации, а именно: формирование определенных полей БЗ, представляя нормативные и другие формулировки соответствующих элементов библиографических записей, подлежащих унификации, а также дополнительную (справочную) информацию;
- создает условия для централизованного хранения данных о географических названиях, а также возможность получения в теледоступе надежных данных о принятых (используемых) и не принятых (например, сокращенных) географических названиях, их местоположении и изменении;
- фиксирует отношения синонимии, иерархические и ассоциативные связи между географическими названиями (между прежними и последующими названиями географического объекта, отношения типа «шире — уже», между принятыми и непринятыми названиями), расширяя сферу поиска данных;
- обеспечивает возможность обмена данными о российских географических названиях с отечественными и зарубежными библиотеками и другими организациями, библиотечными сетями и объединениями.

3. Источники данных

При наличии форматов представления нормативных данных и соответствующих руководств для российских пользователей наиболее сложным и важным оказался вопрос определения источников данных для формирования файла географических названий, в качестве которых выбраны:

- автоматизированный (электронный) Государственный каталог географических названий России (АГКГН) — фрагменты в виде файлов;
- картотека Государственного каталога географических названий России (ГКГН) — справки в традиционной форме;
- «Общесоюзный классификатор административно-территориальных образований» (ОКАТО);
- БЗ, формируемые при каталогизации документов в РГБ, а затем и в других библиотеках;
- нормативные записи из электронных каталогов (ЭК) РГБ, РНБ и других библиотек;
- справочный аппарат каталогов РГБ;
- карточный каталог КГР (отдела картографических изданий РГБ);
- таблицы ББК (ТТД — территориальные типовые деления);
- справочные издания.

Отечественные библиотеки в редких случаях системно подходят к проблеме формирования нормативных файлов географических названий и практически не опираются на официальные источники для постоянного поддержания данных в актуальном состоянии; данные из нормативных файлов библиотек (например, из РНБ) планируется использовать фрагментарно.

АГКГН, который ведет **Центральный институт геодезии, аэросъемки и картографии (ЦНИИГАиК)**, выбран в качестве основного источника данных, так как он представляет собой первоисточник сведений и абсолютно надежен, но поскольку он только создается, еще несколько лет придется пользоваться также справками из традиционной картотеки **ГКГН**. Чтобы ситуация была более понятной, необходимо сказать несколько слов об этих каталогах.

Каждое российское географическое название должно проходить **государственную регистрацию** и получить **постоянный регистрационный номер**, который не меняется при изменении самого названия данного географического объекта. Таким образом, **регистрируется сочетание «объект — название»** под постоянным номером. В **ГКГН** учитывается **более 200 родов объектов**. Вместе с названием регистрируются **30 показателей**, среди которых различаются обязательные и дополнительные (факультативные). Таким образом, для **файла нормативных записей** можно

получить (приобрести) **практически все основные данные**, а именно: нормализованное название, источник установления нормализованного названия, дата установления нормализованного названия, род географического объекта, признак существования объекта, источник установления рода географического объекта, административный статус населенного пункта; сведения об административно-территориальной привязке географического объекта (наименование субъекта РФ, наименование административного района); географические координаты объекта (широта, долгота, привязка к другим географическим объектам); варианты названий, источники установления варианта названия, дату установления варианта названия, дату изменения нормализованного названия и т. д. Например, для реки фиксируются сведения, по каким административным районам она протекает. В структуре данных в АГКГН отражается принятое в настоящее время название географического объекта, название, которое было ранее; предусматривается позиция для названия, которое может появиться со временем. Регистрируется нормализованное название объекта, **полученное обязательно из официального источника**, который также регистрируется с указанием даты регистрации.

АГКГН создан в 2000 г. и находится в состоянии накопления базы данных. Планируется открыть каталог в сети Интернет не ранее, чем через 5-7 лет, когда будет сформирована полная база данных для карт масштаба 1х500000. Затем планируется формировать данные по картам масштаба 1х100000, что означает наличие в базе данных сведений о населенных пунктах вплоть до деревень. В целом, на картах РФ имеется около 2,5 млн. географических названий. В настоящее время в АГКГН есть полные данные по «миллионной» карте России (масштаб 1х1000000), более подробные данные по картам 1х500000 формируются по регионам. В электронной форме выпущен **бюллетень «Изменения 2001-2002 гг.»**, где указываются соответствия «старое — новое», что весьма полезно для нашего файла.

Разумеется, такое положение вещей осложняет нашу задачу. Сейчас разрабатывается тактика работы с ЦНИИГАиК, что требует целого ряда переговоров и финансирования договорных отношений по передаче данных в нормативный файл географических названий.

Другой источник данных в машиночитаемой форме — **ОКАТО**, который выпускает в печатном виде **Госкомстат**. Этот классификатор существует и в виде базы данных. Он содержит данные об административной структуре РФ.

В составе базы данных ОКАТО можно выделить следующие **элементы данных**, которые целесообразно преобразовывать в элементы данных, включаемые в нормативные записи файла географических названий: наименование географического объекта, род географического объекта, административное подчинение, код географического объекта (код ОКАТО). Сведения в этом классификаторе представлены либо в виде текста, либо закодированы в структуре кода географического объекта. Данных недостаточно для нормативного файла, но зато в базе данных ОКАТО описано около 300 тыс. административно-территориальных образований.

4. Структура файла географических названий

Файл географических названий состоит из нормативных/ авторитетных записей в формате **РУСМАРК** для представления нормативных данных в **ЛИБНЕТ**.

Для поддержки файла географических названий в **ЭК РГБ** на основе программного обеспечения **АЛЕФ** нормативные записи представляются в соответствующем формате серии **MARC21**.

При наличии соответствующих данных **нормативная запись включает следующие сведения:**

- нормализованное наименование географического объекта;
- непринятые наименования географических объектов (сокращенные и другие варианты названий);
- географические названия на других национальных языках субъектов РФ и иностранных языках; названия, представленные в другой графике;
- географические координаты;
- регистрационный номер географического названия в Государственном каталоге географических названий, который сохраняется даже при изменении названия;
- источник установления нормализованного названия;
- дата (или только год) установления нормализованного названия;
- род географического объекта;

- признак существования объекта;
- административный статус населенного пункта;
- административно-территориальная привязка географического объекта (наименование субъекта РФ, наименование административного района);
- примечания различного рода.

Предусматривается возможность ввода информации об изменении названия, т. е. ввод в качестве нормализованного нового названия, которое может появиться со временем, сохранение данных о прежнем названии и установление связи между ними. **При изменении принятого географического названия** для нового наименования географического объекта формируется новая нормативная запись, заголовком которой становится новый принятый (нормализованный) заголовок. Записи связываются между собой с помощью полей трассировки ссылок «см. также».

Каждой нормативной записи для географического названия подразделение, которому поручается поддержка файла в РГБ, будет присваивать **национальный регистрационный номер**. Структура регистрационного номера разрабатывается и согласуется на этапе подготовки техно-рабочего проекта в 2004 г.

Пример (1) нормативной записи в формате РУСМАРК (конвертирование из АГКГН):

```
#00100\аАГКГН0000003
#00300\аRuMoRGB
#00500\а20031211163136.0
#100\а20031110crusy0102 са
#152\аPSBO\brugeo
#160\ае-ru\а75 (АГКГН)
#215\аАцинский, хребет\z(31.12.1976 -)
#3001\аАсиновский (31.12.1966 — 31.12.1976), Асинский (31.12.1938 — 31.12.1966)
#3301\ахребет\аЧитинская область
#3301\алевобережье р. Чикой
#3301\а50 град.8 мин. С. Ш., 109 град.30 мин. В. Д.
#415\аАсиновский, хребет\5а
#415\аАсинский, хребет\5а
#801 0\арu\bRuMoRGB
#801 2\арu\bRuMoRGB\с20031211
#810\аАГКГН
```

Пример (2) нормативной записи в формате РУСМАРК (конвертирование из ОКАТО):

```
#00100\fОКАТО0000529
#00300\аRuMoRGB
#00500\а20031211163719.0
#100\а20031110crusy0102 са
#152\аPSBO\brugeo
#160\ае-ru\а27 (ОКАТО)
#215\аКалининградская, область (Российская Федерация)
#3301\аобласть\асубъект федерации\ацентр — г. Калининград
#801 0\арu\bRuMoRGB
#801 2\арu\bRuMoRGB\с20031211
#810\аОКАТО
```

5. Технология формирования и поддержки файла географических названий

Общая схема технологии формирования файла изображена на *рис.1*. Более подробно технологические процессы и вопросы взаимодействия с ЛИБНЕТ будут проработаны в 2004 г.

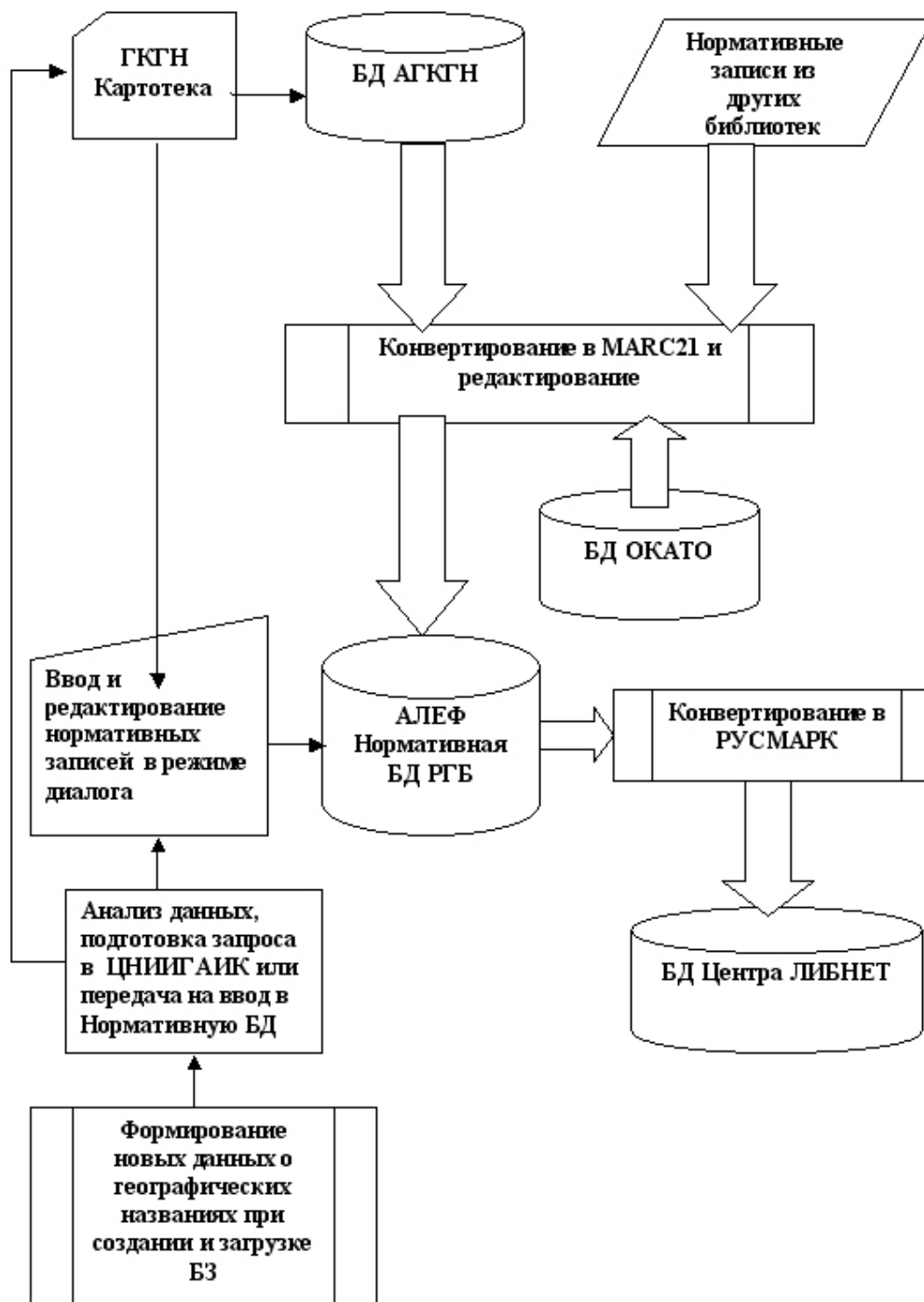


Рис.1. Схема формирования национального файла географических названий

6. Обеспечение взаимодействия файла географических названий и ТТД

Файл территориальных типовых делений (ТТД) таблиц ББК формируется в рамках проекта создания машиночитаемых таблиц ББК в классификационном формате серии РУСМАРК для ЛИБНЕТ и в формате MARC21 для нормативных записей в ЭК РГБ.

В условиях одновременного использования в ЭК файла ТТД для российских географических объектов в рамках машиночитаемых таблиц ББК и файла географических названий уста-

навливаются связи между записями для одного и того же географического объекта как в рамках таблиц ТТД и в пределах НФГН, так и между записями, входящими в разные файлы. Для этого используются средства, принятые в MARC — форматах. Кроме того, и запись в файле ТТД, и нормативная запись для одного и того же географического названия содержат один и тот же индекс ББК (ТТД).

При конвертировании данных из АГКГН и ОКАТО в MARC — формат географические названия будут в будущем программно проверяться на вхождение в файл ТТД. При обнаружении названия в заголовке записи для ТТД индекс будет автоматически копироваться в соответствующее поле нормативной записи в файле географических названий.

Записи в НФГН будут корректироваться оперативно по мере получения данных из каталога ЦНИИГАиК подразделением РГБ, ответственным за поддержку файла (КГР). Сведения об изменениях географических названий или других изменениях в данных о них затем должны быть переданы в НИЦ ББК. После принятия редакционным советом соответствующих решений **изменения будут вносены в файл ТТД** и связанные с ним записи машиночитаемых таблиц ББК, а также в связи между ними.

7. Основные результаты и ближайшие задачи

В 2003 г. проведено предпроектное исследование задачи, разработан технический проект национального файла географических названий, разработаны программы-конверторы для преобразования файлов данных о географических объектах из различных форматов сторонних источников в форматы MARC21 и RUSMARC для нормативных данных; подготовлен экспериментальный массив записей в этих форматах, полученных путем конвертирования и доработки экспериментальных файлов из АГКГН (150 записей) и «Общесоюзного классификатора административно-территориальных образований (ОКАТО)» (100 000 записей). Результаты работ 2003 г. обеспечивают основу для дальнейшего развития проекта и внедрения его результатов в сети ЛИБНЕТ, поскольку удалось решить принципиальные вопросы.

Далее потребуются более широкое обсуждение и согласование проектных решений за пределами РГБ, проведение дополнительного изучения нормативных файлов российских географических названий, которые ведутся в РНБ и некоторых других отечественных библиотеках, организовать взаимодействие с ними, возможно, в форме договоров. Планируется также сформировать договорные отношения и технологическое взаимодействие РГБ и Национального информационно-библиографического центра ЛИБНЕТ в плане передачи исходных файлов в формате РУСМАРК и постоянной актуализации данных; совместно отладить технологию и программное обеспечение.

Технологические структуры, создаваемые на 2-ом этапе, должны обеспечить стабильную работу и взаимодействие подразделений РГБ и организаций-участниц для промышленного внедрения проекта. Необходимо также рассмотрение вопросов обмена файлами географических названий с зарубежными каталогизирующими организациями, включая проблему представления российских географических названий на других языках.

Литература

1. Российская Федерация. Федеральный закон от 18 декабря 1997 г. № 152-ФЗ. «О наименованиях географических объектов»
2. Российский коммуникативный формат представления авторитетных/ нормативных записей (Российская версия UNIMARC/Authorities) / М-во культуры Рос. Федерации. Рос. библио. ассоц. — СПб.: Изд-во РНБ, 1998
3. Формат MARC21 для нормативных записей
4. Общероссийский классификатор объектов административно-территориального деления: ОКАТО 019-95. Принят и введен в действие 31.07.95. Т 1-3. — М.: Издательство стандартов, 1997.