

**Об использовании географических координат
при поиске библиографической информации**

Geographical References in Bibliographic Information Retrieval

**Про використання географічних координат
при пошуку бібліографічної інформації**

О. Л. Жижимов

*Институт вычислительных технологий Сибирского Отделения РАН,
Новосибирск, Россия*

Н. А. Мазов

*Институт нефтегазовой геологии и геофизики им. Академика А. А. Трофимука
Сибирского Отделения РАН, Новосибирск, Россия*

Oleg Zhizhimov

*Institute of Computational Technologies of the Siberian Branch
of the Russian Academy of Sciences, Novosibirsk, Russia*

Nikolay Mazov

*Ac. A. A. Trofimuk Institute of Oil and Gas Geology and Geophysics
of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia*

О. Л. Жижимов

*Институт обчислювальних технологій Сибірського відділення РАН,
Новосибірськ, Росія*

М. О. Мазов

*Институт нефтегазовой геологии та геофизики им. Академика А. А. Трофимука
Сибірського відділення РАН, Новосибірськ, Росія*

Рассматриваются некоторые аспекты, связанные с поиском информации по критериям, задаваемыми географическими координатами. Обсуждаются возможности различных систем метаданных (MARC21, RUSmarc, GILS, GEO), способы построения поисковых запросов и их представление в стандартизованном виде. В качестве иллюстрации рассматриваются интерфейсы шлюза Z39.50 – WWW ZooPARK.

Several aspects of information retrieval by geographically tagged criteria are discussed. Possibilities of various metadata systems (MARC21, RUSmarc, GILS, GEO), methods of query formulation and their standardized representation are examined. Interfaces of Z39.50 – WWW ZooPARK gateways are pointed to as an illustration.

Розглянуто деякі аспекти, пов'язані з пошуком інформації за критеріями заданих географічних координат. Обговорюються можливості різноманітних систем метаданих (MARC21, Rusmarc, GILS, GEO), способи побудови пошукових запитів та їх презентація у стандартизованому вигляді. У якості ілюстрації розглядаються інтерфейси шлюзу Z39.50 – WWW Zoopark.

Сегодня географические координаты объектов могут присутствовать в метаданных (описаниях объектов) в соответствии с различными стандартами описания объектов. Наиболее известными схемами данных (схемами описания объектов), содержащими поля с географическими координатами, являются:

1. USmarc, Marc21 – американский стандарт на библиографические описания различных объектов, ориентированный на структуру стандарта ISO-2709. В табл. 1 представлен фрагмент описаний полей с примерами представления географических координат в формате MRAC21.

Таблица 1

Поле	Подполе	Описание	Примеры
034	\$d – западная долгота \$e – восточная долгота \$f – северная широта \$g – южная широта	Coded Cartographic Mathematical Data. Подполя \$d, \$e, \$f, и \$g всегда присутствуют вместе. Координаты могут быть записаны в форме <i>hddmmss</i> (полушарие-градусы-минуты-секунды), однако другие формы записи также допустимы. Выравненные вправо подэлементы заполняются лидирующими нулями. Односимвольный код полушария: W(-) = запад, E(+) = восток, N(+) = север, S(-) = юг. Возможны обозначения +/--. Лидирующий символ + может быть опущен.	<i>hddmmss</i> : 1#SaaBb220000000\$dW1800000\$eE1800000\$fN0840000\$gS0700000 <i>hddd. dddddd</i> : 1#Saa\$dE079.533265\$eE086.216635\$fS012.583377\$gS020.419532 <i>+-.ddd. dddddd</i> : 1#Saa\$d+079.533265\$e+086.216635\$f-012.583377\$g-020.419532 Без лидирующего + 1#Saa\$d079.533265\$e086.216635\$f-012.583377\$g-020.419532 <i>hddmm. mmmm</i> : 1#Saa\$dE07932.5332\$eE08607.4478\$fS01235.5421\$gS02028.9704 <i>hddmmss. sss</i> : 1#Saa\$dE0793235.575\$eE0860727.350\$fS0123536.895\$gS0202858.125
255	\$c – сведения о координатах.	Cartographic Mathematical Data	##\$aScale 1:41,849,600\$c(W 180°--E 180°/N 90°--S 90°).

Ниже в табл. 2 представлен пример описания картографического материала в формате MARC21. Запись иллюстрирует заполнение полей, относящихся к картографическим материалам: поле 034 (Coded Cartographic Mathematical Data) и поле 255 (Cartographic Mathematical Data).

Таблица 2

Поле	Содержание
Leader/00-23	*****cem##22*****#a#4500
001	<control number>
003	<control number identifier>
005	19920506101053. 0
007/00-05	dc#cen
008/00-39	850203c19841980dk#g#####d#####1###eng##
034	1#SaaBb41849600\$dW1800000\$eE1800000\$fN0900000\$gS0900000
040	##\$a[organization code]\$c[organization code]
052	##\$a3170
100	1#\$aHarig, Karl-F.
245	10\$aReader's Digest world antique spot globe /\$cartography by Karl-F. Harig.
246	14\$aReader's Digest world antique spot globe: encyclopedic handbook.
255	##\$aScale 1:41,849,600\$c(W 180°--E 180°/N 90°--S 90°).
260	##\$a[Copenhagen] Denmark:\$bScan-Globe, \$cc1984, c1980.

Поле	Содержание
300	##\$a1 globe:\$bcol., plastic, mounted on plastic base; \$c31 cm. in diam.
500	##\$aRelief shown by spot heights.
500	##\$aCover title of accompanying text: Reader's Digest world antique spot globe:encyclopedic handbook.
500	##\$aGlobe lights up from inside by means of detachable electric cord and interior light bulb.
500	##\$aWhen the globe is lit, any place on earth can be pinpointed if the latitude and longitude are known. When these two facts are given, it is possible to adjust the two scales to the given figures of latitude and longitude. It causes a pinpoint of light to indicate where the place lies.
500	##\$aAccompanied by text and index: Spot globe 2000 edited by Steen B. Bocher and Henrik B. Hoffmeyer. 17th ed., 1983. 168 p.: ill., maps; 15 x 21 cm.
650	#0\$aGlobes.
700	1#\$aBocher, Steen Bugge, \$d1906-
700	1#\$aHoffmeyer, Henrik B.
710	2#\$aReader's Digest Association.
710	2#\$aScan-Globe A/S.
730	01\$aSpot globe 2000.

2. RUSmarc – российский вариант схемы описания библиографических данных UNIMARC, ориентированный на структуру ISO-2709, учитывающий национальные правила каталогизации. Географические координаты могут присутствовать в полях 123 и 206. В табл. 3 представлены описание и примеры заполнения соответствующих полей.

Таблица 3

Поле	Подполе	Описание	Примеры
123	\$d – западная долгота \$e – восточная долгота \$f – северная широта \$g – южная широта	Координаты для планетарных и земных картографических материалов. Обязательное для картографических материалов. Повторяется, если материал содержит данные в различных масштабах и с разными координатами. Каждое подполе координат (\$d-\$g) имеет фиксированную длину в 8 символов и не повторяется. Каждое подполе содержит следующие данные: – Позиция символа 0. Стороны света. Односимвольный код: w = запад, e = восток, n = север, s = юг – Позиции символов 1-3. Градусы. 3 цифровых символа, выравниваемые вправо заполнителями – нулями. – Позиции символов 4-5. Минуты. 2 цифровых символа, выравниваемые вправо заполнителями – нулями – Позиции символов 6-7. Секунды. 2 цифровых символа, выравниваемые вправо заполнителями – нулями	123 2#\$aa\$b150000\$b25000\$de0150000\$ee0173045\$fn0013012\$gs0023035 Часть карты Заира с линейными масштабами 1:150000 и 1:25000, долгота – от 15°В до 17°30'45» В, широта – от 1°30'12» С до 2°30'35» Ю.
206	\$d – сведения о координатах. Факультативное, не повторяется	Обязательное для картографических материалов. Индикатор 1: Определяет, структурированы ли данные в поле # – Данные представлены в неструктурированном виде 0 – Данные представлены в структурированном виде Индикатор 2: # (не определен)	Неструктурированное представление 206 ##\$aScale 1: 6336 000 (W 170° -W 50°/N 80° -N 40°) Структурированное представление 206 0#\$bScale 1: 6336 000\$d(W 170° -W 50°/N 80° -N 40°)

Следует заметить, что поле 123 содержит те же данные о масштабе и координатах, которые записываются в поле 206, но в кодированной форме. Следует также обратить внимание, что представление координат в поле 123 в RUSMARC совместимо с представлением MARC21 (поле 034). Однако обратной совместимости нет! Возможности представления координат в MARC21 существенно шире и не ограничены жестким форматом.

3. Схема GILS – Government Internet Locator Service [2].

Элемент схемы	Содержание
(4,71)	spatialDomain
(4,71)/(4,91)	boundingCoordinates
(4,71)/(4,91)/(4,9)	westBoundingCoordinate
(4,71)/(4,91)/(4,10)	eastBoundingCoordinate
(4,71)/(4,91)/(4,11)	northBoundingCoordinate
(4,71)/(4,91)/(4,12)	southBoundingCoordinate

Форма представления координат в схеме GILS соответствует форме MARC21.

4. Схема CIMI – Computer Interchange of Museum Information

Элемент схемы	Содержание
(4,4)	objectInfo
(4,4)(4,14)	digitalObject
(4,4)(4,14)(4,29)	actualDo
(4,4)(4,14)(4,29)(5,58)	xCoordinateInSpatialReferencingSystem
(4,4)(4,14)(4,29)(5,59)	yCoordinateInSpatialReferencingSystem

5. GEO – профиль Z39.50 [1,3], соответствующий стандарту представления пространственных метаданных FGDC Content Standards. В таблице приведено соотношение полей, содержащих граничные географические координаты для профиля Z39.50 GEO, стандарта FGDC Content Standards и MARC21.

Элемент схемы	Содержание	FGDC	MARC21
idinfo/spdom/bounding/westbc	West Bounding Coordinate	1. 5. 1. 1	034 d
idinfo/spdom/bounding/eastbc	East Bounding Coordinate	1. 5. 1. 2	034 e
idinfo/spdom/bounding/northbc	North Bounding Coordinate	1. 5. 1. 3	034 f
idinfo/spdom/bounding/southbc	South Bounding Coordinate	1. 5. 1. 4	034 g

Форма представления координат для профиля GEO соответствует форме MARC21.

Потенциальная возможность присутствия полей координат в записях не означает их действительного наличия. Наполнение записей показывает отсутствие необходимых полей в большинстве библиографических баз данных. Исключения составляют

- база данных GeoRef, содержащая описания литературных источников за большой временной период (более 100 лет);
- базы данных FGDC [5-7], изначально ориентированные на хранение описательной информации по объектам с географической привязкой.

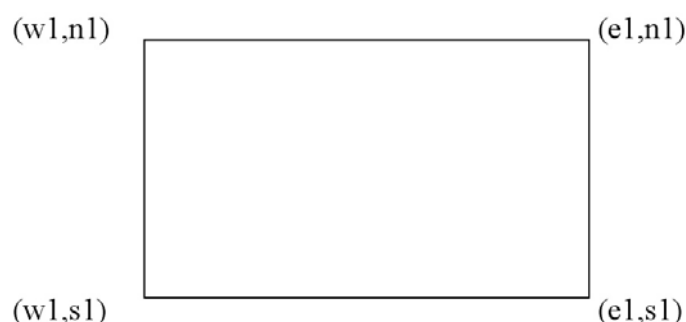
Для работы с географической составляющей этих баз данных необходим специальный подход, применяемый сегодня в ГИС-системах. При этом одним из основных аспектов этого подхода является создание подсистем поиска информации по географическим координатам методом формирования в пользовательских интерфейсах специальных поисковых примитивов (геометрических структур, фиксирующих границы поискового запроса) и реализации алгоритмов конвертирования пользовательских геометрических примитивов в логические операнды, составляющие поисковый запрос к базам данных.

Поскольку единственным стандартом, регламентирующим унифицированный доступ к гетерогенным базам данных сегодня является Z39.50 и его специальные клоны, например, SIP для доступа к данным по объектам с географической привязкой, естественной основой для построения соответствующих поисковых подсистем является методика отображения графических поисковых запросов в логику RPN-запросов Z39.50, оперирующей поисковыми атрибутами, объединенными в большие стандартизированные группы – наборы поисковых атрибутов.

Поисковые атрибуты USE для различных наборов поисковых атрибутов, имеющие отношение к географическому поиску, перечислены представленной ниже таблице.

Bib-1	GILS	GEO	Название
1118	2038	2038	West Bounding Coordinate
1119	2039	2039	East Bounding Coordinate
1120	2040	2040	North Bounding Coordinate
1121	2041	2041	South Bounding Coordinate

Ниже приведен фрагмент запроса RPN на поиск информации в прямоугольнике с заданными граничными координатами



выглядит следующим образом (нотация PQF, принятая в продуктах IndexData и сервере ZooPARK)

```
@and @and @and @attr bib1 1=1118 @attr 2=4 @attr 4=109 w1
@attr bib1 1=1119 @attr 2=2 @attr 4=109 e1
@attr bib1 1=1120 @attr 2=2 @attr 4=109 n1
@attr bib1 1=1121 @attr 2=4 @attr 4=109 s1
```

В качестве примера реализации подсистемы географического поиска библиографической информации можно привести ссылку на шлюз Z39.50-HTTP сервера ZooPARK Сибирского отделения

РАН (<http://z3950.nsc.ru:210>). Поиск производится по координатам, задаваемым вершинами прямоугольника на масштабируемой карте. В настоящей реализации используется GoogleMap.

В заключение, отметим, что представленная работа актуальна в связи с тем, что в настоящее время все большее значение приобретают различные геоинформационные системы, интегрирующие различную пространственную атрибутивную информацию. Поэтому, использование библиографических данных в качестве дополнительных атрибутов подобных систем, на наш взгляд, является перспективным. Несомненно это должно находить отражение в стандартах и рекомендациях на подготовку представления библиографической информации.

Литература

1. ANSI/NISO Z39.50-1995. Information Retrieval (Z39.50): Application Service Definition and Protocol Specification / Z39.50 Maintenance Agency Official Text for Z39.50-1995. – July 1995.
2. Application Profile for the Government Information Locator Service (GILS), Version 2, November 24, 1997. (http://www.gils.net/prof_v2.html).
3. Douglas D. Nebert. Z39.50 Application Profile for Geospatial Metadata or «GEO» / Version 2. 2 / U. S. Federal Geographic Data Committee (<http://www.blueangeltech.com/Standards/GeoProfile/geo22.htm>)
4. Content Standard for Digital Geospatial Metadata (<http://www.fgdc.gov/metadata/constan.html>)
5. FGDC – Federal Geographic Data Committee (<http://www.fgdc.gov>)
6. Crosswalk: USMARC to FGDC Content Standards for Digital Geospatial Metadata. // [электронный ресурс] <http://www.alexandria.ucsba.edu/public-documents/metadata/marc2fgdc.html>
7. Crosswalk: FGDC Content Standards for Digital Geospatial Metadata to USMARC // [электронный ресурс] <http://www.alexandria.ucsba.edu/public-documents/metadata/fgdc2marc.html>