

Представление тезауруса MeSH в формате RUSMARC посредством протокола Z39.50

О. С. Колобов

Научно-техническая библиотека Томского политехнического университета

Н. А. Мешечак

*Научно-медицинская библиотека Сибирского государственного медицинского университета
(г.Томск)*

А. С. Карауш

Муниципальная информационная библиотечная система г.Томска

Oleg S. Kolobov

Research Library of Tomsk Polytechnic University, Tomsk, Russia

Natalya A. Meshechak

Scientific-Medical Library of the Siberian State Medical University, Tomsk, Russia

Alexander S. Karaush

Municipal Information Library System, Tomsk, Russia

*Аннотация

1. n+1 причина использования MeSH как Z39.50 базы данных

Medical Subject Headings (MeSH) контролируемый словарь используется для индексирования, каталогизации и поиска информации и документов в предметных областях биологии, медицины и здравоохранения. Существует множество реализаций программных приложений использующих MeSH, и в каждом предусмотрен сервис поиска в MeSH.

Сочетание MeSH и Z39.50 на сегодняшний день ассоциируется с тем, что данные из MeSH доступны через Z39.50 протокол. И это позволяет использовать стандартные решения для организации поиска и удаленной предметизации с применением внешнего тезауруса MeSH. Преимущества такого подхода, очевидны для той категории пользователей, которые повседневно используют системы, построенные на основе Z39.50.

Важным результатом в работе по представлению MeSH в Z39.50 среде является то, что предлагается не конкретная реализация системы, а технология использования MeSH без ограничений по языку.

К сожалению, при широком использовании и сопровождении разработчиков (NLM) MeSH в англоязычном мире, работа по локализации MeSH для других языков не ведется. Поэтому важно реализовать возможность использования MeSH русскоязычными пользователями, с учетом всех изменений, которые ежегодно проводит NLM с MeSH (например, добавление/удаление дескрипторов и квалификаторов)».

Для адаптирования MeSH в русскоязычной среде выполнен полуавтоматический перевод главных рубрик и подрубрик с последующим конвертированием записи в формат RUSMARC Authority. При конвертировании сохранена структура MeSH, что позволяет использовать оригинальные методы навигации внутри словаря. Таким образом, генерируется вариант MeSH, рассчитанный на русскоязычных пользователей или, другими словами, частично локализованный MeSH, но вполне справляющийся с основными функциями MeSH, как тезауруса.

2. MeSH и Z39.50

Если рассматривать MeSH как тезаурус - семантическую иерархию терминов, как это описано в ISO 2788 [International Organization for Standardization. ISO 2788: Guidelines for the establishment and development of monolingual thesauri, 2nd ed. Geneva: ISO, 1986. For some inexplicable and inexcusable reason, ISO standards are not generally available on-line.], то на сегодняшний день существует Z39.50 профиль Zthes [Zthes: a Z39.50 Profile for Thesauri Navigation см. <http://zthes.z3950.org/profile>], описывающий эту модель данных. Zthes-профиль определяет не только модель данных и навигацию для тезауруса, но и формат и протокол передачи данных.

Для представления локализованной версии MeSH в Z39.50 не был использован данный подход из-за дополнительных затрат на развитие серверного программного обеспечения. В работе были использованы уже отработанные решения, как с технической, так и технологической точек зрения.

MeSH исторически создавался и сопровождался во внутреннем формате NLM MeSH204 (NLM's internal MeSH vocabulary file). И на сегодняшний день NLM использует открытые коммуникативные форматы для передачи MeSH, как XML документ, размеченный ASCII файл и как USMARC Authority файл. Для всех перечисленных форматов существуют положения о преобразовании из внутреннего формата MeSH204 (например для USMARC Authority [Conversion Specification for MeSH to USMARC Authority Format <http://www.nlm.nih.gov/mesh/marc99.pdf>]). Данный факт позволяет лучше понимать организацию MeSH через «призму» другого, знакомого и открытого формата.

По условиям проекта конечным форматом представления локализованной версии MeSH должен быть RUSMARC Authority. Из доступных форматов - формат представления оригинального MeSH USMARC Authority является наиболее привлекательным. Поскольку, по опыту и результату работы над предыдущими проектами ОЭБ [] было разработано программное обеспечение для обработки библиографических записей представленных в коммуникативных форматах.

Задача представления локализованной версии MeSH в Z39.50 состояла в том, что необходимо было дополнить профиль индексирования RUSMARC записей для авторитетных файлов и добавить поддержку точки доступа MeSH-subjects из стандартного набора артибутов Bib-1 [].

3. Организация доступа к MeSH

MeSH в среде Z39.50 представляется как традиционная база данных. Записи могут быть найдены и извлечены из базы данных через стандартный сервис поиска и представления Z39.50 в синтаксисе RUSMARC. Преимущество такого подхода состоит в том, что MeSH может быть использован для решения разнообразных задач, например: для расширения пользовательского запроса нормативными терминами MeSH; для задачи удаленной предметизации с применением MeSH.

В качестве пользовательского интерфейса для работы с MeSH используется Z39.50 шлюз Научной библиотеки СибГМУ [], который позволяет формировать запрос к Электронному каталогу библиотеки.

4. Заключение

Использование MeSH в среде Z39.50 открывает возможность интеграции тезауруса в существующие или во вновь создаваемые информационно-поисковые системы.

Zthes: a Z39.50 Profile for Thesauri Navigation см.

[<http://zthes.z3950.org/profile>].