

Особенности разработки базы данных терминологического словаря по СИБИД для системы автоматизации библиотек ИРБИС64

Challenges of the SIBID terminological database design for LAS IRBIS64

Ю. В. Смирнов

*Государственная публичная научно-техническая библиотека России,
Москва, Россия*

Yury Smirnov

*Russian National Public Library for Science and Technology,
Moscow, Russia*

В статье рассматривается возможность переноса данных из электронного терминологического словаря по системе стандартов по информации, библиотечному и издательскому делу (СИБИД), изданного в ГПНТБ России в 2017 году, в базу данных системы автоматизации библиотек ИРБИС64. Автор приводит основные особенности рассматриваемого терминологического словаря, которые необходимо перенести в версию для САБ ИРБИС64. Далее описываются ключевые особенности этой системы, благодаря которым данный перенос оказался успешным. Достаточно подробно описывается структура рабочих листов, типы поиска, а также реализация ссылочного аппарата. В конце статьи автор приходит к выводу, что, поскольку перенос данных был осуществлен в САБ ИРБИС64 практически без потери функциональности оригинального терминологического словаря, то можно считать данный переход успешным и рекомендовать его дальнейшее использование для других изданий такого рода.

Ключевые слова: терминологический словарь, электронный словарь, САБ ИРБИС64, база данных.

The author discusses the transfer of the data terminological dictionary of the system of standards on information (SIBID), librarianship and publishing, which was published by Russian National Public Library for Science and Technology in 2017, into the LAS IRBIS64 database. The key system characteristics that enabled the migration process, are specified. The worksheet structure, search types and reference tools are described in detail. The author concludes that the data migration was accomplished almost without any losses of functionality and recommends to use the technology for other publications of this kind.

Keywords: Terminological dictionary, electronic dictionary, LAS IRBIS64, database.

В середине 2017 года ГПНТБ России выпустила электронный терминологический словарь по СИБИД [1] (Системе стандартов по информации, библиотечному и издательскому делу). После презентации этого издания встал вопрос о целесообразности разработки версии для системы автоматизации библиотек ИРБИС64 (САБ ИРБИС64).

Характеристики этой системы позволяют создавать не только библиотечные и библиографические базы данных (БД), но и, применительно к терминологическому словарю, лингвистические БД. Такого рода базы присутствуют в стандартной поставке ИРБИС64, но обычному пользователю (не администратору) они обычно не видны.

При разработке БД терминов по СИБИД для ИРБИС64 нам хотелось сохранить максимум особенностей электронного терминологического словаря по СИБИД [1]:

- 1) БД словаря должна содержать информацию не только о самом термине, но и о его источнике, т. е. быть не только словарной БД, но и библиографической;
- 2) каждый термин должен быть снабжен библиографическим описанием источника с указанием страницы его расположения;
- 3) если термин встречается в нескольких стандартах, то все его определения должны присутствовать в БД;
- 4) возможность осуществлять поиск стандартов по значимым полям базы данных, например: по номеру стандарта, по заглавию, по переводу заглавия, по ключевым словам, по индексам Универсальной десятичной классификации (УДК), по индексу Общероссийской классификации стандартов (ОКС) и др.;
- 5) возможность отбора терминов по выбранному стандарту;
- 6) возможность навигации по истории просмотров;

7) использование HTML-разметки, которая предоставляет возможность гипертекстовой навигации по словарю, вместо использования пометок «см.» или «см. также» с последующим поиском указанной ссылки;

8) поддержка одновременной работы с несколькими БД словарей, в том числе осуществление поиска терминов одновременно по всем включенным базам данных;

9) поддержка работы БД на операционных системах (ОС) семейства Microsoft Windows и GNU/Linux.

Большинство из вышеперечисленных особенностей удалось успешно интегрировать в БД терминов ИРБИС64, кроме:

1) возможности навигации по истории просмотров, поскольку САБ ИРБИС64 не предполагает использование данного режима, однако в системе присутствует возможность навигации по истории поисковых запросов, которая зачастую выполняет эту же задачу;

2) поддержки одновременной работы с несколькими выбранными БД словарей, однако это решается объединением или загрузкой всех необходимых словарей в одну базу данных;

3) поддержки работы на ОС семейства GNU/Linux, поскольку сама система ИРБИС64 пока работает только под Microsoft Windows, однако мы надеемся, что рано или поздно версия для ОС семейства GNU/Linux увидит свет (стоит отметить, что САБ ИРБИС64 запускается в ОС семейства GNU/Linux через ПО Wine, предназначенное для запуска Windows-приложений, и большинство функций работают, однако не все).

Для реализации в БД записей как словарного типа, так и библиографического, была задействована возможность работы в ИРБИС64 с разными рабочими листами.

Структура рабочего листа терминов представляет собой следующие поля:

- 1: Термин;
- 2: Синоним (повторяющееся поле);
- 3: Сокращение (повторяющееся поле);
- 4: Перевод (повторяющееся поле);
- 5: Определение;
- 6: Примечание;
- 7: Номер стандарта;
- 8: Страница размещения определения;
- 9: Ссылка (повторяющееся поле);
- 920: Код рабочего листа.

Как видно из представленной структуры, некоторые поля являются повторяющимися, поскольку одному термину могут соответствовать несколько синонимов, сокращений и/или переводов (в том числе на разные языки), а также даваться несколько ссылок.

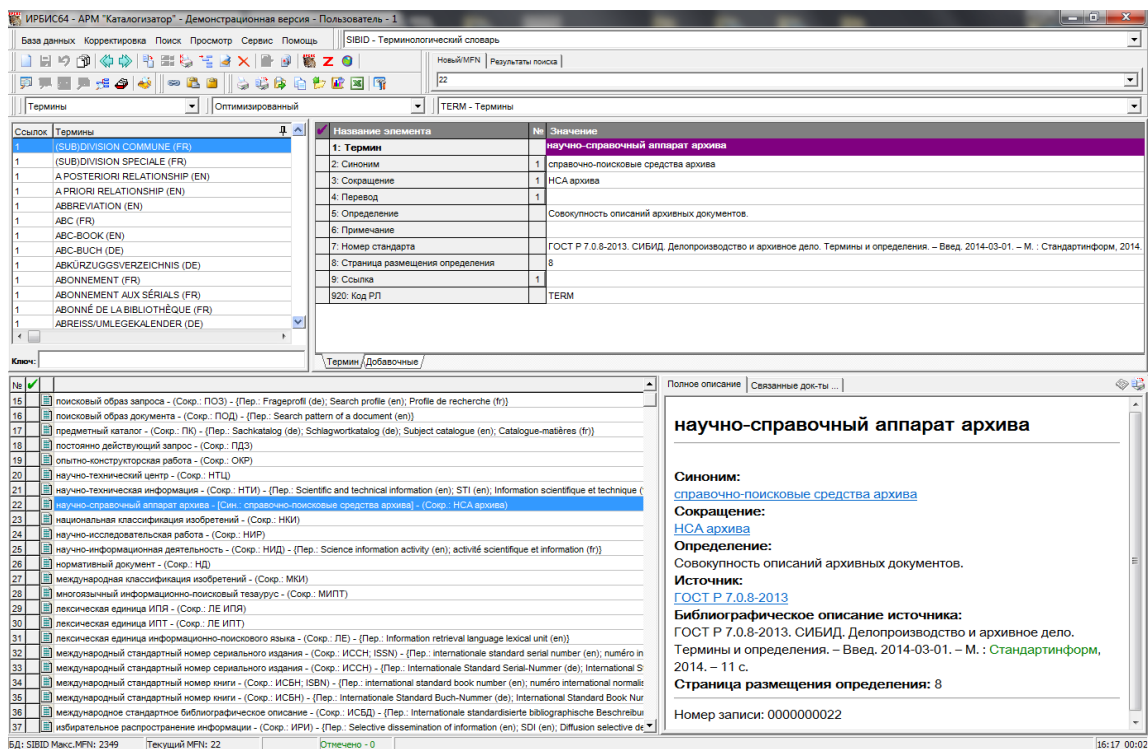


Рис. 1 – Рабочий лист для терминов

Структура рабочего листа стандартов состоит из большого количества полей:

- 101: Статус;
- 102: Номер стандарта;
- 103: Система (серия);
- 104: Основное заглавие;
- 105: Заглавие оригинала;
- 106: Перевод заглавия;
- 107: Код по классификатору (повторяющееся поле);
- 108: Рубрика (и) по классификатору (повторяющееся поле);
- 109: Индекс(ы) УДК (повторяющееся поле);
- 110: Рубрика(и) по УДК (повторяющееся поле);
- 111: Ключевые слова (повторяющееся поле);
- 112: Дата введения;
- 113: Дата утверждения;
- 114: Дата ограничения;
- 115: Дата продления;
- 116: Дата перенесения;
- 117: Взамен стандарта;
- 118: Библиографическое описание;
- 920: Код рабочего листа.

В этом рабочем листе также были использованы повторяющиеся поля, там где это необходимо.

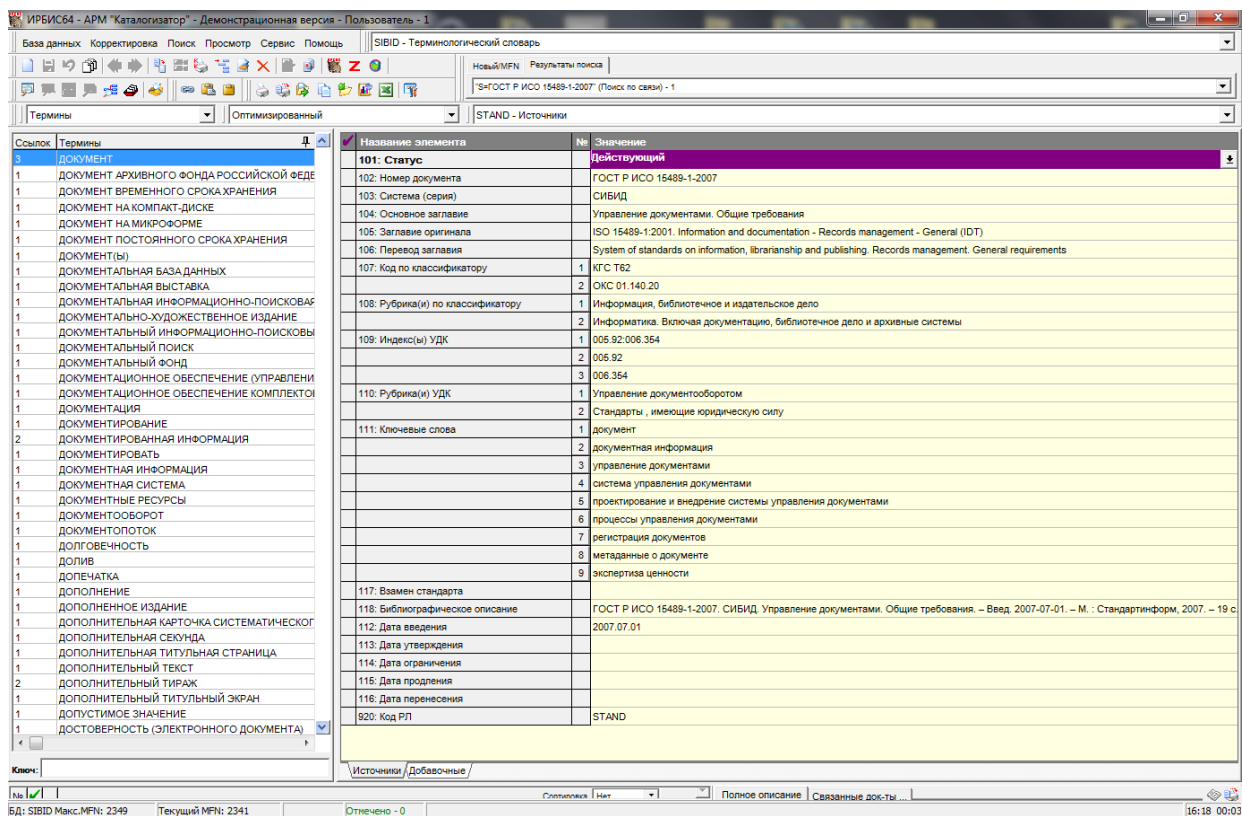


Рис. 2 – Рабочий лист для стандартов

Конечно, ни одна база данных не может существовать без поисковых возможностей, и система ИРБИС64 предлагает обширный и мощный инструментарий для этих целей. Для БД терминологического словаря были сформированы следующие виды поиска:

- 1) «Термины» – это поиск не только по основным терминам, но и по синонимам, сокращениям и переводам;
- 2) «Переводы терминов» – это поиск исключительно по переводам терминов;
- 3) «Термины по номеру стандарта» – с помощью этого вида поиска можно найти все термины, которые представлены в выбранном стандарте;
- 4) «Номер стандарта»;
- 5) «Заголовок стандарта по ключевым словам» – поиск по каждому слову, представленному в полях заголовка («104: Основное заглавие», «105: Заглавие оригинала» и «106: Перевод заглавия») стандарта;
- 6) «Полный заголовок стандарта» – в отличие от предыдущего вида поиска, здесь поиск осуществляется целиком по заголовку стандарта, без разбиения его на ключевые слова;
- 7) «Код по классификатору»;
- 8) «Рубрика по классификатору»;
- 9) «Индекс УДК»;
- 10) «Рубрика УДК»;
- 11) «Ключевые слова».

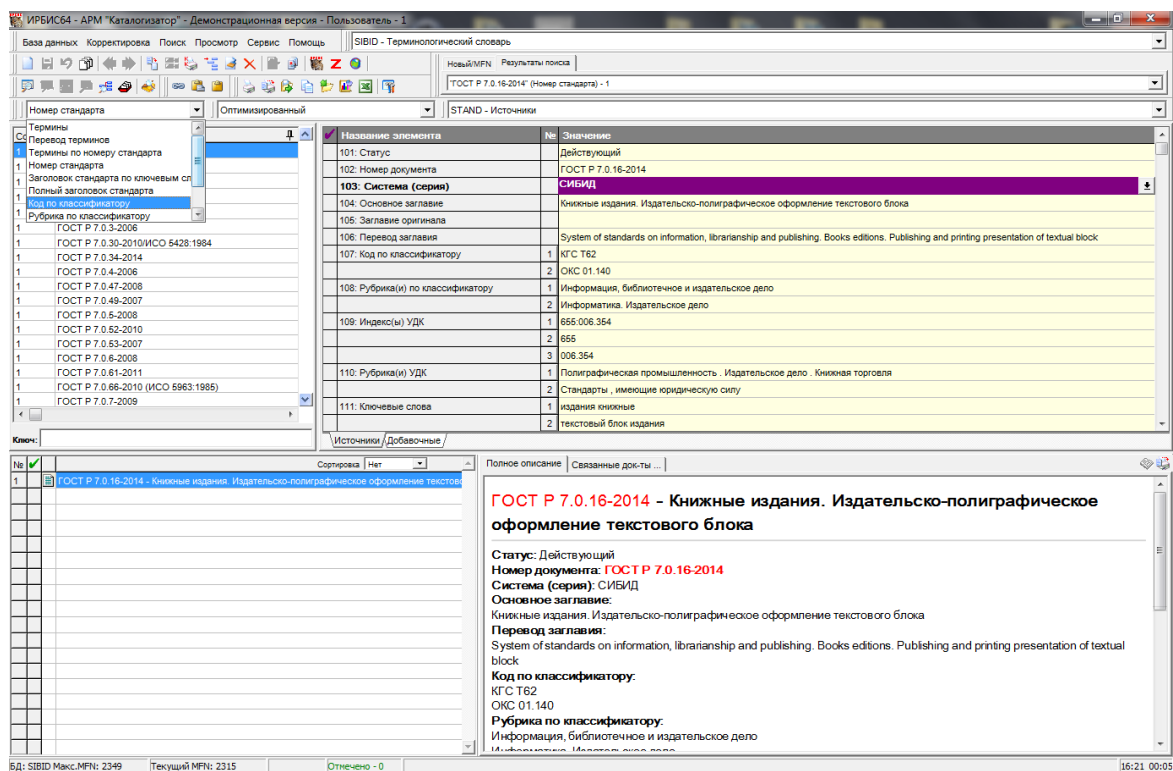


Рис. 3 – Виды поиска

Ссылочный аппарат для этой БД реализован с помощью гиперссылок, таких же, как и на любой веб-странице, что упрощает и ускоряет процесс навигации по БД.

Также, благодаря отличной работе системы ИРБИС64 с языком разметки html, все информационные сведения как о терминах, так и о стандартах представлены в структурированном виде.

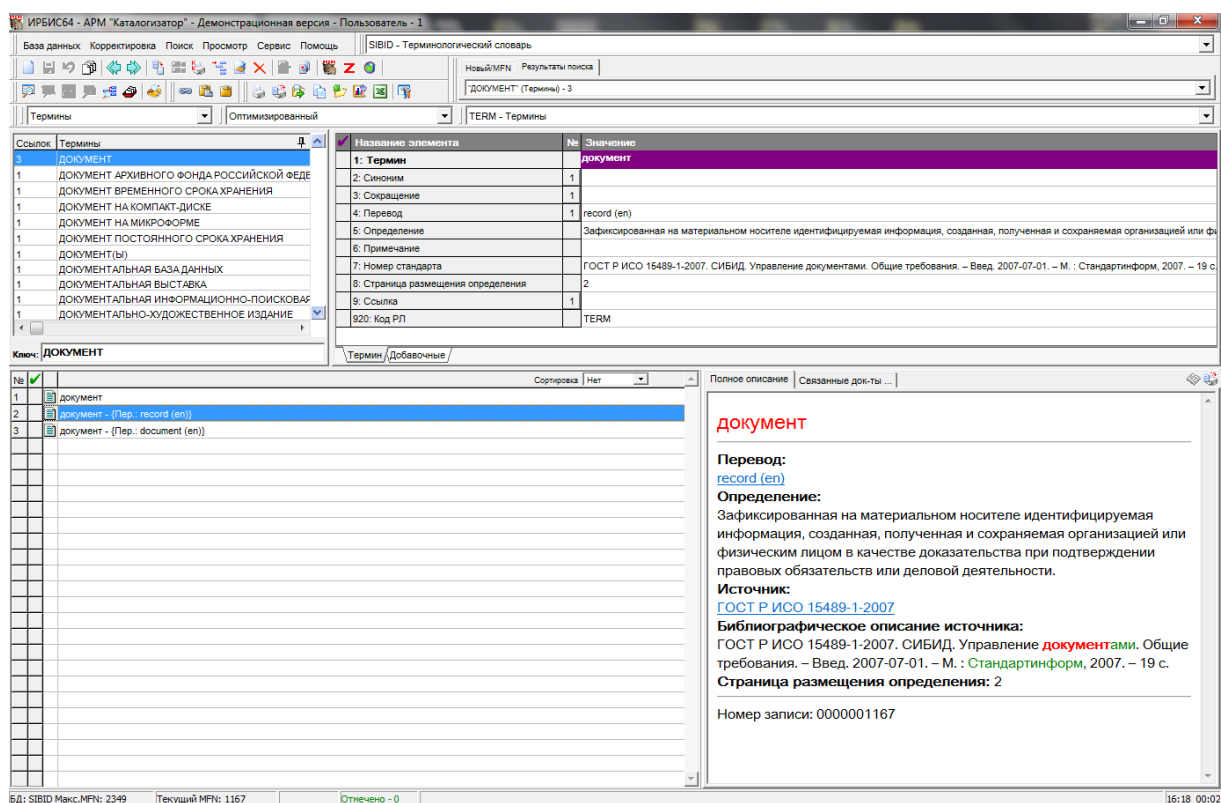


Рис. 4 – Гиперссылки и язык разметки html в ИРБИС64

Как видно из описанного выше, нам удалось, практически без потери функциональности оригинального терминологического словаря, перенести его в БД ИРБИС64. А поскольку отсутствующие функции не являются критическими и не влияют на работу БД, то можно считать данный переход полностью успешным и рекомендовать его использование для других изданий такого рода.

Список литературы

1. Электронный терминологический словарь по СИБИД. А – Я [Электронный ресурс] : [справ. пособие] / Гос. публ. науч.-техн. б-ка России ; [сост. Ю. В. Смирнов ; ред. Е. М. Зайцева]. –М. : ГПНТБ России, 2017. – 1 электрон. опт. диск (DVD-ROM).