

Из опыта создания базы данных серой литературы в Латвийской академической библиотеке

Серая литература – что это?

Есть различные определения серой литературы, но чаще всего цитируется определение 1920 г.: «Серая литература» – это материалы, которые невозможно купить в книготорговой системе и поэтому их трудно выявить и приобрести». К серой литературе относятся научно-технические отчеты, диссертации, материалы конференций, препринты, статистические сборники, библиографические издания.

Преимущества и значение серой литературы:

- 1) ее можно издать и распространить в короткие сроки с небольшими затратами, поэтому она поступает к читателям быстрее, чем журнальные статьи и монографии;
- 2) серая литература дает детальную информацию, которая интересует многих людей. Часто она подробнее, чем информация, которая появляется позже в журнальных статьях и монографиях;
- 3) актуальность темы – часто информация, изданная в виде серой литературы, повторно как официальный документ не издается; серая литература важна в различных научных исследованиях, особенно в межотраслевых, где необходима самая оперативная информация.

Государства Европейского Сообщества (ЕС) в 1980 г. создали систему доставки документов SIGLE – System for Information on Grey Literature in Europe (Система информации по серой литературе в Европе). Эта система охватывала серую литературу, созданную в странах ЕС и была доступна всему миру в режиме онлайн. Основная цель системы – обеспечить доступ к документам серой литературы. Вначале она объединяла страны ЕС, но теперь она открыта и другим странам Европы, становясь Всеевропейской системой. Рабочий язык – английский.

Европейские центры информации и документации объединились в ассоциацию EAGLE – European Association for Grey Literature Exploration (Европейская ассоциация пользователей серой литературы). EAGLE руководит системой работы SIGLE с помощью Генеральной Ассамблеи, Исполнительного и Технического комитетов. Центр обработки данных ответственен за их получение и ввод, а также за все технические процессы.

Он находится в Карлсруе (Германия).

В данное время систему SIGLE образуют 16 стран. В каждой стране есть национальный центр, который отвечает за:

- 1) комплектование серой литературы своей страны;
- 2) ввод библиографических описаний в базу данных SIGLE;
- 3) обеспечение доступности документов читателям;
- 4) доставку документов заказчику.

В конце 2002 г. база данных SIGLE состояла из 793 000 записей, в течение года вводится примерно 55 000 записей, база данных пополняется каждый месяц. Наибольший вклад делают страны-участницы: Великобритания – приблизительно 24 000 записей в год, это примерно 49 % от общего количества записей,

Германия	–	8 000 записей	или	24 %
Франция	–	4 000	х	10 %
Нидерланды	–	4 000	х	10 %
Латвия	–	400-500	х	1 %

В базе данных SIGLE представлены практически все отрасли науки, но более всего:

техника	31 %
естественные науки	26 %
общественные науки	20 %
биология и медицина	15 %

База данных SIGLE доступна как в формате CD-ROM (выпускает SilverPlatter), так и в режиме онлайн из STN International (The Scientific and Technical Information Network). Латвийская академическая библиотека использует только STN International, т. к. финансово это более выгодно, чем CD-ROM.

Поиск в базе данных STN International можно проводить по различным параметрам: предметному указателю, ключевым словам в названии, другим ключевым словам, автору, коллективному автору, государству, месту и году издания документа, языку, номеру научного отчета. Можно также комбинировать поиск по нескольким параметрам.

ЛАБ вступила в Ассоциацию EAGLE в 1996 г., но первые пробные записи были отосланы в EAGLE в 1995 г. Начиная с 1996 г. записи отправляются регулярно, приблизительно 400-500 в год. До конца 2002 г. было отослано 2780 записей.

ЛАБ как ассоциированный член имеет свои обязанности, которые необходимо выполнять:

- 1) регулярно платить членские взносы,
- 2) регулярно пересылать свои записи в Технический центр,
- 3) гарантировать доступ к серой литературе и пересылку ее по требованию.

Как и другие члены Ассоциации, ЛАБ является владельцем базы данных серой литературы. Владельцы – теоретически – могут и индивидуально продать эту информацию, но соблюдая те же условия, что и центр. Практически для ЛАБ это невыгодно. Будучи членом ассоциации EAGLE, ЛАБ тем не менее не имеет никаких скидок как пользователь STN International. Популяризация латвийской науки и достижений латвийских ученых – вот главное приобретение от участия в базе данных серой литературы. К сожалению, мы не можем сказать насколько велик интерес зарубежных исследователей к записям, которые посылает ЛАБ, так как в базе данных SIGLE все слито воедино. Отдел МБА нашей библиотеки получал запросы на некоторые серые издания, но там не были указаны источники информации о запрашиваемой литературе.

В ЛАБ существует давняя традиция комплектования серой литературы. Библиотека Мисиньша (или отдел латышской литературы) уже с первых дней существования комплектовала все, что издано в Латвии, независимо от содержания работы, следовательно, и серую литературу. С 1946 г. наша библиотека, став библиотекой Академии наук, комплекзует диссертации, отчеты о различных научных работах и ведомственные издания.

Сегодня комплектование ЛАБ серой литературой происходит из двух основных источников:

- 1) обязательный бесплатный экземпляр;
- 2) ЛАБ получает отчеты о проектных исследованиях, заказанных учреждениями государственного управления; кроме того, мы получаем научные отчеты исследовательских институтов, проекты которых финансируются Советом по науке Латвии. Это дает нам возможность предположить, что в нашей библиотеке наиболее полно представлена серая литература и мы можем работать как центр серой литературы в Латвии.

Особенности обработки серой литературы

Библиографические описания серой литературы для базы данных SIGLE делаются заново, несмотря на то, что в библиотечной базе данных LIBER они уже есть.

Это обусловлено тем, что эти разные системы нельзя совместить.

Требования к библиографическому описанию в SIGLE отличаются от нашего описания, необходимы дополнительные сведения, которых нет в LIBER.

Несколько характерных требований:

1) серая литература разбита на группы и каждая группа имеет свою букву обозначения, например, диссертации обозначают буквой U, материалы конференций – K, библиографии – Z;

2) обязательно надо при каждом документе указать организацию, которая ответственна за содержание данного документа, присвоение ученой степени или за организацию конференции, это не то же самое, что коллективный автор;

3) название документа на английском языке. Раньше это было обязательным элементом, теперь по желанию, но нельзя самовольно переводить название на английский язык, необходимо перевод согласовать с автором или издателем;

4) ключевые слова, если их не дал автор, можно писать самим и количество их не ограничено. Если название не переведено на английский, то в этом поле дают в переводе наиболее важные слова из него;

5) если у документа есть аннотация, ее вписывают, однако не превышая 5000 знаков, желательно на английском языке, но можно и на языке оригинала;

6) обязательно необходимо указать доступность документа – в каком учреждении он находится;

7) в классификации используется не УДК, а своя система, которая создана на основе US COSATI «Subject Category List» (Federal Council for Science and Technology, Committee on Scientific and Technical Information). Этот предметный указатель состоит из 20 основных групп, которые делятся на 250 предметных категорий. Классификационные коды состоят из трехзначной комбинации цифр и букв, например: 03A – загрязнение воздуха; 05L – литература; 05B – библиотековедение, информатика. Каждая запись должна иметь по крайней мере один классификационный код, но не больше трех, из которых один – основной, а другие – дополнительные.

Полнотекстовая база данных серой литературы ЛАБ

Примерно 3 года назад, при поддержке Фонда Сороса в Латвии ЛАБ приступила к созданию полнотекстовой базы данных латвийской серой литературы для популяризации достижений наших ученых. Эта база данных доступна в Интернете и в ней представлены разные виды серой литературы, но больше всего диссертаций, аннотаций к ним и резюме. Если в базе

данных диссертация дана полностью, то она разбита на отдельные разделы, чтобы облегчить читателям поиск. База данных предлагает и список работ, который по желанию автора можно регулярно пополнять и корректировать. Документы сканируются или записываются с дискет, но в любом случае они перед вводом должны быть обработаны. Оформляются документы так, чтобы их можно было использовать в соответствующей программе, кроме того, каждая работа дополняется различными указаниями и пояснениями о текстовом языке и формате, отвечающими требованиям системы. И несмотря на техническое обеспечение процесс этот очень трудоемкий.

Создание полнотекстовой базы данных не требует таких строгих условий, как база SIGLE. Условия довольно свободные и у каждого документа иные.

Одни форматы индексируются, другие нет. Конкретную тему можно искать в базе индексированных документов. Программа дает список тех работ, где указанный термин присутствует и дальше уже можно выбрать отдельный документ.

Работу в полнотекстовой базе данных можно проводить по двум направлениям:

- 1) поиск по ключевым словам;
- 2) подбор документов по спискам.

Программа сначала предлагает все, что есть в базе, а потом по этому списку возможен запрос по отраслям.

Документами, которые введены в полнотекстовую базу данных, читатель может оперировать по своему усмотрению: переслать на другой компьютер, записать на дискету, распечатать и т. д. В базе дана оговорка, что каждый документ является интеллектуальной собственностью автора и его нельзя использовать в коммерческих целях, но библиотеке трудно проследить, выполняет ли читатель данное требование. О том, что эти условия невозможно проконтролировать, каждого автора предупреждают.

Полнотекстовая база данных серой литературы – одна из наиболее запрашиваемых баз ЛАБ, но она пока не столь объемна по ряду объективных и субъективных причин. Например, есть авторы, которые не хотят столь широкой доступности своих работ.

На 1 января 2003 г. в этой базе данных было 150 документов серой литературы. Существует еще одна форма учета, когда учитываются объекты информации, т. е. не вся работа целиком, а отдельные ее части, например, части диссертации, аннотации, списки литературы и т. д. Все части работы,

которые обрабатываются и индексируются как отдельные единицы, в поиске тоже отображены отдельно.

В ассоциации EAGLE ЛАБ представлена как национальный центр, который ответственен за всю серую литературу, публикуемую в Латвии. Но на сегодня в базу данных SIGLE мы посылаем сведения только о литературе, находящейся в наших фондах. Ассоциация все-таки хотела бы, чтобы серая литература Латвии была более полно представлена в базе SIGLE. Это значит, что в будущем нам предстоит привлечь к этой работе и другие крупные библиотеки Латвии, что будет совсем непросто.