

Опыт организации работ по конвертированию БД в НТБ СибГТУ*

Научно-техническая библиотека Сибирского государственного технологического университета (НТБ СибГТУ) в течение 10 лет автоматизировала библиотечные процессы на основе АИБС «Библиотека 4.02». Но библиотечные информационные технологии развиваются, появляются новые программы, предоставляющие качественно иные возможности. В 2002 г. принято решение перейти на ИРБИС. Планируя эту работу, мы не хотели повторить опыт некоторых библиотек, которые при переходе с «Библиотеки 4.02» на ИРБИС не смогли сконвертировать свои записи и начали создавать электронный каталог заново.

Как известно, в ИРБИС существует стандартный конвертор для «Библиотеки 4.02» и в профессиональной печати не было публикаций о трудностях перехода. Поэтому на процесс конвертирования мы запланировали всего один месяц. Однако попытавшись «перебросить различные типы записей из нескольких каталогов, мы получили крайне неудовлетворительный результат. Большая часть записей не конвертировалась совсем, остальные содержали грубые ошибки в описании.

Попытавшись выяснить причины, мы сделали вывод, что один человек (главный библиотекарь ОВКТ) с этой работой не справится и создали временный творческий коллектив (ВТК) «Конвертация». Его целью было конвертирование баз данных НТБ СибГТУ из АИБС «Библиотека 4.02» в ИРБИС, главной задачей – создание технического задания на конвертирование и файла конвертора.

В состав ВТК вошли шесть человек: пять библиотекарей (главные специалисты научно-библиографического и отдела обработки документов, заведующая методическим отделом и заместитель директора по научно-методической работе) и один программист. Они обеспечивают руководителя первичной информацией для создания ТЗ (заполнение формы ТЗ без пункта «Условия» – об этом ниже), выявляют ошибки, исправляют их лично или дают задания сотрудникам отделов, а также отчитываются перед руководителем ВТК.

Руководитель ВТК организует работу, проводит анализ первичной информации, сводит ее в единую таблицу, проверяет; заполняя пункт «Условия», готовит ТЗ для передачи программисту; будучи экспертом (для данной библиотеки) в форматах машиночитаемой каталогизации, дает консультации членам ВТК; отчитывается перед руководством библиотеки, предлагает конкретные решения. С программистом общается по электронной почте только руководитель ВТК (программист работает на 0,25 ставки).

После составления ТЗ всегда возникают дополнительные вопросы, они решаются только руководителем. При невозможности самостоятельно решить проблему руководитель связывается с внешними специалистами, консультантами.

На основе информации, полученной от ВТК, руководство библиотеки принимает управленческие решения, влияющие на все уровни иерархии.

Как показывает опыт, работа по созданию ТЗ сложная, аналитическая, требующая больших временных и интеллектуальных затрат, что необходимо учитывать при формировании группы, предназначенной для этой работы.

Таблица 1

Этапы работы ВТК «Конвертация»

Месяц	Этап работы
январь	Попытка конвертировать различные типы записей с помощью стандартного конвертора (ОВКТ). Отрицательные результаты
февраль	Создание ВТК. Определение целей, задач, методов работы; подбор литературы
март	Изучение литературы. Изучение АРМ «Каталогизатор» ИРБИС. Создание таблицы соответствия полей. Попытка создать на ее основе конвертор. Отрицательный результат

апрель	Семинар в ГПНТБ СО РАН по изучению формата RUSMARC. Новая технология работы над ТЗ. Новая форма для создания ТЗ. Самостоятельная работа членов ВТК по переброске информации в ИРБИС.
май	Удаление из ЭК «Статьи» записей до 1996 г. Просмотр наполнения каждого поля ЭК «Статьи», выявление ошибок. Начало создания ТЗ.
июнь–август	Создание ТЗ на конвертацию ЭК «Статьи». Исправление грубых ошибок в ЭК «Книги». Создание четких инструкций по заполнению полей в ЭК «Книги», «Статьи», «Труды» для исключения повторения выявленных ошибок («Б-ка»). Редактирование ЭК «Статьи», «Труды». Редактирование ЭК «Библиотекведение» для переброски. Создание конвертора на ЭК «Статьи».
сентябрь–октябрь	Создание ТЗ на конвертацию ЭК «Труды сотрудников СибГТУ». Накопление первичной информации по ЭК «Книги». Редактирование ЭК «Труды». Создание конвертора на ЭК «Труды».

Для работы программиста изначально планировалась простая форма для ТЗ (табл. 2); но она оказалась непригодной.

Таблица 2

Библиотека 4.02		ИРБИС	
метка	пример	метка	пример
091a (инд. ББК)	Ч481.24	621	Ч481.24
040a (назв.орг.)	НТБ СибГТУ	902	НТБ СибГТУ

В табл. 3 представлен окончательный вариант формы ТЗ и некоторые примеры ее заполнения. Появился пункт «Условие», где необходимо подробно расписывать программисту правила деления информации по полям и подробно описывать варианты заполнения.

Таблица 3

«Библиотека 4.02»			ИРБИС (автоматически)			ИРБИС (должно быть)	
метка	пример	условие	метка	пример	условие	метка	пример
100a 1 авт.	Абрамова Е.А. и др. 2001	убрать после 2 точки и др., и др.	700a 463j	Абрамова Е.А. и др. 2001	– –	700a 700b 463j	Абрамова Е.А. 2001
773d Вых. данные 9 вар.	2001.–№3.– С.23–33	Номер м.б. и №, и лат. №. То же со страницами.	463j	2001.–№3.–С.23–33	Если после 4-х цифр идет – , то после 1-й.- и до 2-й.- инф. в 463v, после .- С. – в 463s	463j 463s 463v	2001 23– 33 №3
773g Смеш. инф. 8 вар.	Vol. 10, № 4.- С.243 Vol.13,№12.- Р.97 Т.88,№3,- С.56 Т.74, Вып.4.-С.684	Часть м.б.:Т.;Vol.;Vol. Номер м.б. лат. №, и просто №; Вып.4.	463v	Vol.10,№4.-С.243 Vol.13,№12.-Р.97 Т.88,№3.-С.56 Т.74, Вып.4.-С.684	Инф. до, в 463v, после, до .- в 463, после .- С.(или.- Р.) в 463^1 и 463s	463v 463h 463^1 463s	Vol.10 №4 С. 243
100a 100b 111a	МИФИ ИТ СибГТУ ИТ	М.б. такие варианты: только 100, 100 или	710a 710b	МИФИ ИТ не перебрасывается	1) при переброске записей с б/ф	711/1a 711/1b 711/2a	МИФИ ИТ СибГТУ ИТ

111e	только 111a, 111e, или 100.	уровнем а, установить раб.лист ASP (по умолчанию PVK, PVKASP).	711/2b
------	-----------------------------	--	--------

Из представленных примеров видно, что существуют различные по уровню сложности поля. Чтобы составить техническое задание для некоторых из них (как в последнем примере), потребовалось много времени и интеллектуальных усилий (так как часть информации не перебрасывалась и устанавливались другие рабочие листы).

Для удобства работы членов ВТК создана схема подготовки ТЗ на конвертацию, где описаны все этапы.

Работа велась по каждому заполненному полю в программе «Библиотека 4.02». Сначала проверялось наличие поискового элемента в списке (их может быть не более 18, поэтому приходилось удалять предыдущий, вставляя новый). Если поисковый элемент существовал, сразу начинался сплошной просмотр на варианты заполнения. Параллельно выявлялись ошибки, грубые или единичные исправлялись сразу.

Если поискового элемента не было, то руководитель ВТК создавал его, реактуализировал записи. Эта работа для ЭК «Труды», «Статьи» занимала до 30 минут, а для ЭК «Книги» – 2–3 часа. Причем во время реактуализации каталогом нельзя пользоваться, поэтому работа проводилась после закрытия библиотеки.

Если вариантов заполнения не было (например поле «Год издания» – там только год и никаких ошибок), то запись сразу сохранялась и перебрасывалась в ИРБИС. Но таких полей было 2–3, в остальных же – вариантов заполнения оказывалось очень много. Их все необходимо было выявить и записать. Записи с каждым вариантом сохранялись и перебрасывались в ИРБИС.

После этого проверялась правильность переброски данного поля. Если поле перебрасывалось правильно, оно сразу описывалось в ТЗ. Но если неправильно, то сначала проверялась правильность его внесения в программу «Библиотека». Если выявлялась ошибка, проводились работы по ее исправлению; если же нет, то руководитель искал варианты решения проблемы, найденное решение описывалось в ТЗ.

На практике из-за небольшого количества членов ВТК и загруженности руководителей (зав. методическим отделом и зам. директора) созданием первичной информации занимался и руководитель ВТК (количество полей в ЭК было разделено на части между членами ВТК).

Следует отметить, что в процессе анализа записей выявлено большое количество ошибок при вводе данных в ЭК, что вызвано, на наш взгляд, несколькими факторами:

- ранним началом внедрения автоматизированных технологий – до 1995 г. отсутствовало описание формата машиночитаемой каталогизации (не было опыта коллег);
- незнанием заведующими отделами и главными специалистами, занимающимися созданием ЭК, формата машиночитаемой каталогизации (USMARC, на котором работала библиотека (с 1995 г. описание формата хранилось в фонде НТБ СибГТУ, но не использовалось);
- отсутствием четких инструкций по вводу (единственная инструкция создана в 1995 г. и не редактировалась);
- недостаточным количеством словарей, подключенных к полям (как показывает опыт, наименьшее число ошибок выявлялось в полях, к которым подключены словари);
- низкой степенью ответственности каталогизаторов (пример: весь отдел заполняет поле по одним правилам, а один сотрудник может заполнять то же поле по совершенно другим правилам, известным только ему);
- отсутствием редактора каталогов.

Для продолжения работы все ошибки необходимо было разбить на группы по значимости и классифицировать. Прежде всего ошибки делятся на недочеты программы и ошибки, допущенные

каталогизаторами НТБ. Как правило, недочеты программы выявить несложно и они исправимы.

Ошибки, допущенные каталогизаторами, мы разделили на случайные и системные.

К случайным относятся опечатки (их редактирование не представляется возможным) и занесение информации в другое поле (как правило, рядом стоящее). Последняя ошибка серьезнее, далее мы ее рассмотрим подробнее.

Системные ошибки – допускаемые систематически, возможно, документированные (Советы по автоматизации). Они делятся на те, которые можно исправить программным путем (это зависит от квалификации программиста), и те, которые исправить программным путем нельзя. Это самые серьезные ошибки, вызывающие принципиально неправильное конвертирование. Пример: для безынвентарной литературы в отделе комплектования фондов (где создается первичное описание), проставлялось количество экземпляров. Когда запись попадала в отдел обработки, данные о количестве экземпляров удалялись. Это привело к тому, что безынвентарная литература не перебрасывается в ИРБИС (записи как некачественные логически удаляются). Необходимо нестандартное решение проблемы, но его пока нет.

Работа ВТК привела к следующим результатам:

- отработана технология создания ТЗ на конвертирование, что позволит быстрее создавать технические задания на другие каталоги;
- отредактированы ЭК «Статьи», «Труды», «Библиотечковедение» и подготовлены к конвертации;
- классифицированы ошибки, что позволит принимать взвешенные управленческие решения по планированию и организации дальнейшей работы по конвертированию;
- члены ВТК освоили работу с АРМ «Каталогизатор» в ИРБИС.

Объемы работ, выполненных членами ВТК:

- просмотрены на варианты заполнения 60 полей ЭК «Статьи», «Труды», «Библиотечковедение» (созданы, реактуализированы поисковые элементы, записи с каждым вариантом переброшены в ИРБИС, описаны в ТЗ);
- отредактировано более 7 500 ошибок в 34 полях ЭК «Статьи», «Труды», «Библиотечковедение»;
- на редактирование затрачено более 500 часов (около 2,5 месяцев по 8 часов ежедневно);
- написано более 60 писем коллегам и программисту для решения возникших проблем (в дальнейших планах ВТК создание ТЗ и конвертирование ЭК «Книги»);
- конвертирование оставшихся ЭК НТБ СибГТУ;
- создание инструкций по вводу информации в ИРБИС.

Проделанная работа позволила сделать следующие выводы:

- нельзя оценивать ЭК только количеством записей, обязательным условием должно быть соответствие введенных в него записей формату машиночитаемой каталогизации (ФМК) и требованиям качественной семантической обработки;
- несоответствие записей ФМК – брак в работе каталогизатора, требующий больших временных, интеллектуальных, организационных (а следовательно, и финансовых) затрат на его устранение;
- уровень ответственности каталогизатора при работе в сложных системах и при возможности выхода созданного продукта во внешнюю среду повышается многократно;
- в библиотеке необходимо вводить систему управления качеством ЭК.

Философия всеобщего управления качеством предусматривает:

- ориентирование всей библиотеки на удовлетворение потребностей пользователей;

- концентрацию усилий на процессах, которые создают и поддерживают качество;
- осознание сотрудниками каждого подразделения, что они – часть предприятия и поэтому должны понимать его потребности и оказывать содействие любому другому подразделению;
- всеобщую нетерпимость к браку на основе принципа «Делать правильно с первого раза».

Выводы из философии качества (применительно к машиночитаемой каталогизации):

- главное внимание необходимо уделять контролю качества не на выходе (когда библиографическая запись уже создана), а управлению качеством на всех этапах ее создания;
- управление качеством должно носить системный характер;
- эффективность систем управления качеством зависит от овладения философией качества работниками всех подразделений и всем персоналом библиотеки.

На основе всего вышеизложенного ВТК предложил комплекс мер по обеспечению качества электронных каталогов НТБ СибГТУ (в течение года эти меры реализованы):

- ввести должность главного редактора каталогов;
- разработать четкие инструкции по заполнению полей с учетом требований формата UNIMARC (ни один сотрудник не имеет права единолично принимать решения по изменению правил каталогизации);
- организовать постоянное повышение квалификации главных специалистов, занимающихся каталогизацией, следить за изменениями в форматах, вовремя вносить изменения в инструкции по вводу;
- использовать как можно большее количество словарей и других лингвистических средств. Контроль за качеством семантической обработки поручить комиссии по ЛО;
- к сотрудникам, допустившим брак в работе, применять меры административного воздействия.

Таким образом, работа по конвертированию данных из одной программной оболочки в другую показала, что профессия каталогизаторов электронных каталогов требует большой ответственности и специфических знаний. При отсутствии этих качеств огромные затраты по созданию ЭК могут оказаться сизифовым трудом.

В НТБ СибГТУ электронный каталог удалось сконвертировать с минимальными потерями, и необходимые выводы были сделаны, но для улучшения качества электронных каталогов во всех вузовских библиотеках города, по нашему мнению, в рамках Методического объединения необходимо:

- создать секцию по машиночитаемой каталогизации с разделением на подсекции по форматам (UNIMARC, MARC21);
- организовать обучение каталогизаторов в рамках секции, сформировать экспертную группу;
- корпоративно закупать средства лингвистического обеспечения.

С автором можно связаться по адресу: olushakova@krw.ru

* Текст доклада на научно-практической конференции "Библиотеки на пути к открытому информационному пространству", Красноярск, 29 окт. 2002 г. в НТБ КрасГАСА.