

СТЕНДОВЫЕ ДОКЛАДЫ

Новые подходы в реализации справочного аппарата для создания электронных каталогов и улучшения поиска в них

Горшкова Галина Анатольевна, зав. отделом, Научная библиотека Петрозаводского государственного университета, Петрозаводск

В докладе рассматривается справочный аппарат, разработанный и реализованный в АИБС “Фолиант”. Справочный аппарат разработан с применением Российского коммуникативного формата представления авторитетных / нормативных записей. Он разрабатывается в помощь каталогизаторам, систематизаторам и предметизаторам для создания библиографических записей, а также для улучшения поиска в электронных каталогах. Справочный аппарат в промышленном режиме работает в НБ Петрозаводского государственного университета. Готовится корпоративная версия для библиотек Карелии и библиотек, работающих в АИБС “Фолиант”.

О разработке подобного справочного аппарата мечталось давно. Однако, уровень используемых АБИС не позволял реализовать его в полном объеме. Справочный аппарат (будем его так называть, поскольку иное название трудно пока придумать) включает в себя предметные рубрики, тезаурусы, систему вышестоящих терминов и понятий (элементы систематизации), классификационные индексы, а при желании полочные индексы.

Основой справочного аппарата является “авторитетная запись”. Возьмем это понятие в кавычки, т. к. оно по своему смыслу и содержанию значительно отличается от авторитетной записи, к которой все сегодня привыкли. Но и справочной эту запись тоже назвать нельзя.

Представляемый справочный аппарат реализован в АИБС “Фолиант” с использованием Российского коммуникативного формата представления авторитетных / нормативных записей. Реализация такого справочного аппарата в “доморощенном” формате, как показывает практика, приводит к тому, что его использование затруднено другими библиотеками и корпоративными библиотечными объединениями. Конвертирование данных из одной АИБС в другую, из одной версии программы в другую приводит к потерям информации и требует значительных временных затрат на последующее редактирование данных.

Для создания справочного аппарата из коммуникативного формата задействованы поля: 200 Имя лица, 210 Наименование организации, 215

Географическое заглавие, 250 Тематическая предметная рубрика. К сожалению, для решения задач, которые мы поставили перед собой, этих полей оказалось недостаточно. Одной из поставленных задач при разработке справочного аппарата предусматривается контроль за внесением сведений практически во все подполя создаваемой библиографической записи. Поэтому пришлось самостоятельно добавить поля 290 Место издания, 291 Издательство, 292 Дата издания, 293 Объем, 294 Сведения об иллюстрациях, 295 Размер, 296 Сопроводительный материал, 297 Форма, вид и т. п., 298 Читательское назначение, 299 Сведения, относящиеся к заглавию. И так как этот перечень не полностью покрывает заполняемые для библиографической записи поля, то в последствии будут добавляться новые поля для дальнейшего развития справочного аппарата. Если возникнет вопрос с обозначениями самостоятельно добавленных полей, то их всегда можно заменить. В АИБС “Фолиант” замена обозначений в перечне полей влечет за собой автоматическую замену их в “авторитетных записях”.

Каждая “авторитетная запись”, создаваемая с использованием коммуникативного формата имеет определенную структуру и содержит определенный перечень полей:

- поле 2-го блока (например, 250, которое никогда не повторяется);
- поле 4-го блока (например, 450, которое повторяется необходимое количество раз);
- поле 5-го блока (например, 550, которое повторяется необходимое количество раз);
- поле 675 Индекс УДК;
- поле 686 Индекс ББК.

В “авторитетную запись” могут быть добавлены и классификационные индексы других систем. Добавлено поле 9-го локального блока - Полочный индекс, отсутствующее в коммуникативном формате.

В полях 2-го блока находятся термины, выбранные в качестве основополагающих и внесенные по правилам составления предметной рубрики или по энциклопедическим изданиям.

В полях 4-го блока отражаются синонимы и варианты терминов, внесенных в поля 2-го блока.

В поля 5-го блока включаются вышестоящие термины по отношению к терминам, внесенным в поля 2-го блока. Цепочки вышестоящих терминов формируются на основе классификации определенной науки.

В поле 675 находятся классификационные индексы/индекс УДК, относящиеся к терминам, внесенным в поля 2-го блока.

В поле 686 находятся классификационные индексы/индекс ББК, относящиеся к терминам, внесенным в поля 2-го блока.

В поле 9-го блока может быть включен полочный индекс по рабочим таблицам библиотеки, что значительно облегчает работу систематизаторов.

“Авторитетная запись”, являющаяся основой справочного аппарата, содержит:

- Предметную рубрику (тематическую предметную рубрику, имя лица, имя организации, географическое наименование, издательство и т. д.), составленную с использованием методики предметизации;
- Варианты написания и произношения предметной рубрики, а также синонимы термина, представленного в виде предметной рубрики, в том числе на иностранных языках;
- Все вышестоящие термины, определенные с помощью таблиц УДК, ББК и на основе классификации определенной науки или отрасли знаний (систематизация терминов);
- Индексы УДК и ББК;
- Полочные индексы НБ ПетрГУ.

Продемонстрирую на примерах тематической предметной рубрики, что представляет собой “авторитетная запись”:

250a Амарант, род	250a Гербы	250x Символика
450a Род Амарант	450a Символика гербов	
450a Щирица	450a Символика герба	
450a Щирица, род	450a Герб	450x Символика
450a Род Щирица	450a Гербовая символика	
450a Amaranthus	450a Символика на гербах	
550a Амарантовые, семейство	450a Символика на гербе	
550a Гвоздичные, порядок	550a Геральдика	
550a Двудольные, класс	550a	Вспомогательные исторические дисциплины
550a Цветковые растения	550a Исторические науки	
550a Высшие растения	686a 63.215	686\$2 rubbk
550a Растения	550x Систематика	940a 63.2
550a Растения		
550a Ботаника		
550a Биологические науки		
550a Естественные науки		
675a 582.663		
686a 28.592.72	686\$2	rubbk
940a 58		
250a Русская литература	250x Романы	250z 20 в.
450a Русская литература	450x Романы	450z 20в.
450a Русская литература 20 в.	450x Романы	
450a Русская литература 20в.	450x Романы	
450a Русский роман 20 в.		
450a Русский роман 20в.		
450a Русский роман XX в.		
450a Русский роман XX века		
450a Русские романы 20 в.		
450a Русские романы 20в.		
450a Русские романы XX в.		
450a Русские романы XX века		
250a Уголовное право	250x История	250y Белоруссия 250z 20 в.
450a Уголовное право	450x История	450y Белоруссия 450z 20в.
450a Уголовное право	450x История	450y Беларусь 450z 20 в.

450a Уголовное право 450x История 450у Беларусь 450z 20в.
450a История уголовного права Белоруссии 20 в.
450a История уголовного права Белоруссии 20в.
450a История уголовного права Белоруссии XX в.
450a История уголовного права Белоруссии XX века
450a История уголовного права Беларуси 20 в.
450a История уголовного права Беларуси 20в.
450a История уголовного права Беларуси XX в.
450a История уголовного права Беларуси XX века
450a История белорусского уголовного права 20 в.
450a История белорусского уголовного права 20в.
450a История белорусского уголовного права XX в.
450a История белорусского уголовного права XX века
450a Белорусское уголовное право 450x История 450z 20 в.
450a Белорусское уголовное право 450x История 450z 20в.

Из приведенных примеров видно, что “авторитетные записи” бывают двух видов. До определенного уровня детализации они содержат основную предметную рубрику, варианты и вышестоящие предметные рубрики. С определенного уровня детализации в записях присутствуют только основная предметная рубрика и ее варианты.

Все “авторитетные записи” находятся в соответствующих базах данных. Держать их в едином массиве не лучший вариант. С отдельными базами данных работать удобнее. Однако, следует иметь в виду, что записи физически в базах данных не присутствуют. MARCовская запись не хранится в АИБС “Фолиант” в первоизданном виде. Она хранится в виде системных словарей на уровне сведений подполей, а затем при необходимости происходит визуализация записей в базах данных.

Созданы и дополняются базы данных: “Предметные рубрики”, “Имя лица”, “Организации”, “Географические наименования”, “Место издания”, “Издательства” и др. Сегодня они содержат ок. 1 млн. 200 тыс. “авторитетных записей”.

Справочный аппарат разрабатывается с конца 2006 г. Естественно за такой короткий срок невозможно было и программу разработать и обеспечить наполнение баз данных таким количеством “авторитетных записей”. Поэтому изначально во всех базах данных записи были созданы автоматически на основе сведений системных словарей, формируемых в АИБС “Фолиант”. Сейчас работа продолжается: неправильные записи (а их достаточно много) удаляются, правильные редактируются и дополняются, создаются новые записи, отсутствующие в базах данных.

Порядок работы по созданию “авторитетных записей”, правила и особенности их создания вопрос отдельный, поэтому здесь он не рассматривается.

Все базы данных закрепляются за соответствующими полями библиографических записей. Например, база данных “Имя лица” закреплена за полями 700, 701, 702, 600, база данных “Предметные рубрики” – за полем 606, база данных “Издательство” – за полем 210 и т. д.

Рассмотрим теперь, как же такие “авторитетные записи” работают, позволяя облегчать работу каталогизаторов, систематизаторов и предметизаторов в процессе создания и редактирования библиографических записей и значительно улучшая и облегчая поиск информации в электронных каталогах.

Отлажена технология взаимодействия между библиографическими и “авторитетными записями”.

Постепенно меняется порядок работы с библиографическими записями. Если к полю или подполю имеется хоть какая-то база данных “авторитетных записей”, то сотрудник, создающий библиографическую запись, просто обязан при работе использовать ее. При отсутствии нужного термина он вносится сотрудником. Сейчас, когда объем “авторитетных записей” еще не очень велик, нет диктата в данном вопросе. В дальнейшем сотрудник не сможет попасть в поле или подполе, минуя справочный аппарат.

После вывода базы данных с “авторитетными записями” на уровень поля сотрудник производит поиск необходимого ему термина. При этом ему совсем необязательно знать, как, например, правильно он должен внести предметную рубрику. В запросе он может, как дилетант, набрать любой из вариантов, внесенных в поля 2-го или 4-го блока. В библиографическую же запись будет автоматически перенесен вариант их 2-х полей, т. е. выбранный в качестве основного.

Следом за основным вариантом, расположенным в полях 200, 210, 250 и т. п., в библиографическую запись автоматически будут перенесены вышестоящие термины, расположенные в полях 500, 510, 550 и т. п. В библиографической записи при этом будет сформировано автоматически столько полей, сколько 5-х полей имеется в “авторитетной записи”.

Автоматически в библиографическую запись будут перенесены индексы УДК, ББК и полочный индекс. Конечно же, технология перенесения сведений имеет много нюансов.

Поиск информации в электронном каталоге и базах данных для пользователей строится с учетом всех терминов, внесенных в поля 2-го и 4-го блоков, т. е. учитываются все варианты и синонимы, внесенные в “авторитетные записи”. Поэтому такие запросы пользователей, как “Влияние света на рост растений”, “Русские романы XX века”; “Зимняя война”, “История Франции XVII века”, “Царствование Александра I”, “Отношение А. С. Пушкина к литературе”, “Мат. анализ”, “Сопромат” “Землетрясения” (ошибка сделана специально) будут выполнены по электронному каталогу, хотя такие варианты в библиографических записях отсутствуют. Естественно, чтобы такой поиск действовал, эти варианты должны реально присутствовать в “авторитетных записях”.

Теперь хочется остановиться на том объеме работы, который предстоит проделать. Справочный аппарат будет включать со временем несколько миллионов “авторитетных записей”, и довольно объемным по находящимся в них сведениям. На первый взгляд объем просто фантастический, и кажется,

что выполнить его невозможно. Но это только на первый взгляд. Если подойти к работе творчески и с умом, то она обозрима. Только не надо хвататься за все сразу.

В каждой науке или отрасли знания существует ее логическое построение. Именно от этого и стоит оттолкнуться. Вначале следует создать “авторитетные записи” на термины первого, второго, третьего уровней детализации, а затем постепенно углубляться в термины следующих уровней. Как показала практика нашей работы, такой порядок работы значительно сокращает время на создание одной “авторитетной записи”, т. к. создавать записи можно копированием. Правда, эту работу может выполнять далеко не каждый сотрудник. Она сложна и требует определенных знаний и опыта.

Возможно также создание блоков “авторитетных записей”. Например, детализируем термин “уголовное право”. Главное максимально полно учесть все варианты и синонимы в первоначальном блоке. Термины “трудовое право”, “земельное право” и т. п. можно создавать копированием, заменяя слово “уголовное” на слово “трудовое”. И эту работу можно поручить даже студентам.

Возможности нашей библиотеки ограничены, но, тем не менее, мы создали за год около 30 тыс. “авторитетных записей”. На первый взгляд вроде бы и немного, но даже это небольшое количество записей облегчило работу систематизаторам и предметизаторам с разделами “Социальные науки”, “Исторические науки”, “Естественные науки”.

Кроме того, почувствовав реальную значимость справочного аппарата для поиска, сотрудники библиотеки вносят в имеющиеся “авторитетные записи” только варианты, не дорабатывая ее до конца. Очень хорошо, доработать до конца можно и позже.

Решением сотрудников наиболее крупных библиотек Карелии принято решение, что разрабатываемый справочный аппарат будет корпоративным. По технологии предполагается, что он будет находиться на отдельном сервере, а библиотеки будут настраивать свои АРМы “Каталогизатор” для работы с этим сервером (внутри АИБС “Фолиант” взаимодействие библиотек происходит без протокола Z3950). Тогда каждая библиотека сможет внести свой конкретный вклад в разрабатываемый справочный аппарат.

Сегодня все библиотеки, работающие в АИБС “Фолиант”, могут воспользоваться созданным массивом “авторитетных записей”, правда, только для того, чтобы посмотреть. Но и это уже подспорье в работе каталогизаторов и систематизаторов (чаще всего в одном лице).

Создаваемые “авторитетные записи”, как видно из вышесказанного, не привязаны к терминам, внесенным в библиографические записи. В справочном аппарате “авторитетных записей” намного больше, чем реально присутствует в библиографических записях электронного каталога. По мере развития справочного аппарата планируется вывести на поиск системные словари для пользователей, привыкших использовать их при поиске. При этом проводить какую-то дополнительную работу для этого не придется. Будет существовать автоматическая привязка “авторитетных записей” к

терминам, внесенным в библиографические записи. Строка, выведенная в словаре в качестве основной, будет автоматически тянуть за собой все синонимы и варианты.

В заключение хотелось бы перечислить, что реально мы имеем, разрабатывая справочный аппарат такого уровня:

1. Сокращение времени на создание библиографических записей. Каталогизаторы, систематизаторы, предметизаторы будут реально затрачивать меньше времени на работу с документом.
2. Возможность отказа от редактирования созданных библиографических записей, так как сведения в записи переносятся, а не вводятся. В этом случае исключаются опечатки.
3. Полный отказ от редактирования ретроспективных библиографических записей и загруженных из других библиотек, за исключением исправления с них опечаток и орфографических ошибок.
4. Создаваемый справочный аппарат будет корпоративным. Им смогут пользоваться даже небольшие библиотеки, которым недоступны большие по объему авторитетные файлы. Качественно улучшится поиск в их электронных каталогах без какой-либо дополнительной работы со стороны этих библиотек.
5. Качественно улучшается поиск по электронным каталогам. Многие пользователи смогут найти любую информацию без помощи библиографов. Уже сегодня многие читатели приходят в библиотеку, произведя поиск по электронным каталогам дома или на работе, а там библиографа нет.
6. Возможна реализация поиска, как по словарям, так и без них в зависимости от подготовленности пользователя к работе с электронными каталогами.
7. Возможно приближение электронных каталогов российских библиотек пользователям других стран. Если в “авторитетные записи” в поля 4-го блока внести, как вариант, термины на других языках, то можно будет выполнить запрос, сформировав его на соответствующем языке.