

**Подсистема электронного каталога АБИС библиотеки  
МГТУ им. Н. Э. Баумана «Яуза»**

**Electronic Catalog Subsystem  
of MSTU «Yauza» Library ALIS**

**Підсистема електронного каталогу АБІС бібліотеки  
МДТУ ім. Н. Е. Баумана «Яуза»**

*А. Е. Шиваров*

*Московский Государственный Технический Университет им. Н. Э. Баумана, Москва, Россия*

*Alexander Shivarov*

*N. E. Bauman Moscow State Technical University, Moscow, Russia*

*О. С. Шиваров*

*Московський державний технічний університет ім. Н. Е. Баумана, Москва, Росія*

В докладе рассматриваются вопросы эффективности и удобства электронных каталогов современных АБИС, в частности, проблема сложности формулирования поискового запроса для неподготовленных пользователей. В качестве решения данной проблемы предлагается наряду с поисковыми возможностями электронного каталога использовать навигационные. Рассматриваются принципы построения подсистемы электронного каталога АБИС библиотеки МГТУ им. Н. Э. Баумана «Яуза», позволяющие организовать навигацию в электронном каталоге при помощи именованных и неименованных подмножеств записей.

The paper addresses the issues of efficiency and friendliness of modern ALIS electronic catalogs, in particular, the complexity of retrieval query formulation by untrained users. To solve the problem, the author offers to use navigation functions along with the retrieval ones. The structure of the electronic catalog subsystem of «Yauza» Library ALIS (N. E. Bauman MSTU) enabling the navigation through the electronic catalog driven by named and unnamed record subsets.

В доповіді розглядаються питання ефективності та зручності електронних каталогів сучасних АБІС, а саме, проблема складності формулювання пошукового запиту для не підготовлених користувачів. В якості вирішення даної проблеми пропонується поряд з пошуковими можливостями електронного каталогу використовувати навігаційні. Розглядаються принципи побудови підсистеми електронного каталогу АБІС бібліотеки МДТУ ім. Н. Е. Баумана «Яуза», що дозволяють організувати навігацію в електронному каталозі за допомогою іменованих та неіменованих підмножин записів.

На ранних этапах развития технологий автоматизации библиотек электронный каталог представлял собой автоматизированный вариант традиционного карточного каталога [1, 2]. Со временем эта ситуация изменялась, функциональные возможности электронных каталогов расширялись, но даже на современном этапе по мнению многих исследователей электронные каталоги не обладают требуемой функциональностью и удобством для пользователя [1, 3, 4].

В процессе использования электронного каталога с целью удовлетворения информационной потребности пользователь может осуществлять взаимодействие двух видов:

- поисковое;
- навигационное.

Под поисковым будем понимать взаимодействие, происходящее по следующему сценарию: пользователь формулирует и отправляет в систему поисковый запрос, в ответ на который она формирует выборку релевантных записей.

Под навигационным понимается взаимодействие, при котором система предлагает пользователю несколько вариантов дальнейших действий, среди которых пользователь выбирает наиболее подходящее, уточняя тем самым свою информационную потребность для системы.

Исследования в области пользовательского поведения показали, что большинство начинающих пользователей информационно-поисковых систем находят наиболее сложным процесс формулирования поискового запроса [5]. Основной причиной этих сложностей является недостаточное знакомство с предметной областью, в которой производится поиск. Действительно, для того, чтобы

сформулировать хороший поисковый запрос, требуется знание предметной области. Однако во многих случаях причиной информационного поиска является именно необходимость расширить знания о данной предметной области. Наличие подобного «замкнутого круга» демонстрирует неприменимость традиционного механизма поиска в каталоге по запросам для пользователей с неглубоким знанием предметной области. В первую очередь это относится к студентам – основным пользователям библиотеки университета.

Решить данную проблему позволяет использование навигации вместо поиска. Исследователи сходятся во мнении, что навигация является более простым и удобным для пользователей способом для нахождения нужных записей в каталоге[1, 4]. При этом организация данных конкретного каталога для предоставления возможности навигации представляет собой отдельную задачу в каждом конкретном случае.

Подсистема электронного каталога автоматизированной библиотечно-информационной системы библиотеки МГТУ им. Н. Э. Баумана «Яуза» предоставляет пользователю комбинированную поисково-навигационную модель электронного каталога. Наряду с традиционными поисковыми возможностями в пользовательский интерфейс добавлены средства навигации. Основой навигационной составляющей являются именованные и неименованные подмножества записей электронного каталога, позволяющие пользователю переходить из одного состояния взаимодействия с электронным каталогом в другое, выбирая одно из рекомендованных релевантных подмножеств на данном этапе работы с системой.

Именованные подмножества формируются на основе справочных полей библиографического описания и получают имя из соответствующих справочников, например, классификаторов, справочников авторов, издательств и т. п. Примерами являются подмножества записей из рубрики «языки программирования», автора «Савельев», издательства МГТУ им. Н. Э. Баумана.

Неименованные подмножества не имеют четкого образующего признака вроде предметной рубрики, автора и т. п. Подмножества данного вида формируются системой автоматически с использованием косвенных признаков, поэтому они не имеют естественного имени. Примерами таких подмножеств являются записи на часто заказываемые совместно книги, записи похожие на заданную и т. д.

Построение именованных подмножеств сводится к определению доступных справочников конкретного каталога. Любая справочная структура данных (каталог авторов, издательств, классификационная схема, предметный рубрикатор) порождает соответствующие именованные подмножества.

Для построения неименованных подмножеств необходимо определить косвенные признаки, по которым могут быть связаны документы, в том смысле, что потребность в одном документе вызывает потребность в другом или один документ с некоторой точки зрения похож на другой. Основное применение неименованных подмножеств «похожих» документов – это обеспечение навигации из состояния просмотра подробного описания одного документа к описаниям других, связанных с просматриваемым. Например, если студент, подбирая литературу для работы над курсовым проектом, остановил выбор на некотором издании, то система автоматически рекомендует ему обратить внимание на перечень аналогичных. Аналогами выбранного издания могут считаться документы, которые рекомендованы по той же дисциплине, часто просматриваются вместе с выбранным, либо часто заказываются или выдаются вместе с ним.

Учитывая специфику библиотеки университета, можно предложить следующие возможные признаки ассоциирования документов для формирования неименованных подмножеств:

1. присутствие документов в одном списке рекомендаций по дисциплине;
2. одновременный заказ документов одним читателем;
3. одновременная выдача документов одному читателю;
4. просмотр документов в рамках одного поискового сеанса.

Применение именованных подмножеств для навигации требует от системы оценки их релевантности информационной потребности пользователя на каждом из этапов поиска. Другими словами, требуется найти подмножества, релевантные текущей выборке записей.

При вычислении релевантных подмножеств данного типа для текущей выборки записей следует рассмотреть два случая:

1. образующий выборку записей запрос является строгим фильтром;
2. образующий выборку записей запрос является запросом на поиск по тематике.

Стратегии формирования релевантных подмножеств в этих двух случаях различаются.

В первом случае, когда текущее множество записей четко определено, смысл релевантных подмножеств состоит в том, чтобы предложить варианты уточнения запроса, если выборка получилась слишком большой. В этом случае релевантные подмножества являются уточняющими и могут быть построены по любому из справочных атрибутов описания: по автору, издательству, классификационным рубрикам, типу документа. Если текущее подмножество образовано фильтром по классификационной рубрике, то релевантные подмножества должны включать ее подрубрики, а также рубрики, связанные с ней отношением «см. также».

Во втором случае более разумно предложить в качестве релевантных подмножеств множество релевантных текущему запросу классификационных рубрик. Во многих случаях это позволило бы пользователю быстрее достичь цели, т. к. с большой долей вероятности в одном из используемых рубрикаторов будет найдена подходящая рубрика, которая существенно сузит область поиска. Однако это требует выработки метода оценки релевантности классификационной рубрики поисковому запросу. Релевантность рубрики запросу должна отражать степень удовлетворения информационной потребности пользователя, выраженной запросом, документами рубрики. Поскольку в данном случае нами рассматривается информационная потребность тематического поиска, требуется оценить в какой степени рубрика отражает тематику поискового запроса. В рамках электронного каталога АБИС о рубрике может быть доступна следующая информация, позволяющая судить о ее тематике:

- наименование;
- положение в иерархии рубрикатора;
- связи с другими рубриками;
- документы данной рубрики;
- сведения о действиях пользователей по отношению к данной рубрике в контексте определенных запросов.

Модель оценки релевантности рубрики может использовать любые из этих источников информации. Объединение информации из всех этих источников с учетом их различной значимости для тематической характеристики рубрики возможно с использованием модели взвешенных ключевых слов. В рамках предлагаемой модели каждая рубрика характеризуется нечетким множеством ключевых слов (или словосочетаний)  $Kw(r)$ . Степень принадлежности ключевого слова нечеткому множеству рубрики характеризует вес или значимость данного слова для отражения тематики рубрики. Таким образом,

$$Kw(r) = Kw(name(r)) \cup \lambda \prod_{d \in L(D,r)} Kw(d)L(d,r)$$

где

$name(r)$  – наименование рубрики;

$Kw(name(r))$  – функция извлечения ключевых слов из имени рубрики в виде нечеткого множества;

$Kw(d)$  – функция извлечения ключевых слов из документа в виде нечеткого множества;

$L(D, r)$  – множество документов из рубрики  $r$ ;

$L(d, r)$  – значение степени принадлежности документа  $d$  рубрике  $r$ ;

$\lambda$  – коэффициент, показывающий во сколько раз ключевые слова, извлеченные из документов рубрики, характеризуют ее слабее, чем слова из наименования.

Предложенная модель может быть расширена для учета информации о связанных рубриках, а также информации о действиях пользователей.

## Литература

1. Antelman K., Lynema E., Pace A. K. Toward a 21st Century Library Catalog // Information Technology and Libraries. – 2006. – №25 (3).
2. Borgman C. L. Why Are Online Catalogs Still Hard to Use? // Journal of the American Society for Information Science. – 1996. – №47 (7).
3. Belkin N. J. Anomalous States of Knowledge as a Basis for Information Retrieval // Canadian Journal of Information Science. – 1980. – № 5.
4. Markey K. The Online Library Catalog. Paradise Lost and Paradise Regained? // D-Lib Magazine. – 2007. – №1–2 (13).
5. Belkin N. J. Helping People Find What They Don't Know // Communications of the ACM. – 2000. – № 12.