

# №13-14 2000 год

## РАСШИРЕНИЕ ВОЗМОЖНОСТЕЙ ЧИТАТЕЛЕЙ ЦНСХБ В СВЯЗИ С РАЗВИТИЕМ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Центральная научная сельскохозяйственная библиотека, созданная в 1930 г. при Всесоюзной академии сельскохозяйственных наук (ныне Россельхозакадемия), – национальная библиотека России в области сельского хозяйства и продовольствия, в фонде которой насчитывается около 1,8 млн томов отечественной и 1,2 млн томов иностранной литературы по всем вопросам сельского хозяйства и смежным отраслям знаний (естественные науки, пищевая и перерабатывающая промышленность).

В настоящее время ЦНСХБ так же, как и другие аналогичные бюджетные учреждения, лишилась части государственных ассигнований, но за счет большей хозяйственной самостоятельности обрела новые формы привлечения средств и возможности улучшить обслуживание читателей. Огромную роль в этом сыграло развитие компьютерной техники и использование информационных технологий.

С 1985 г. информация о входящем потоке документов стала собираться в базе данных, создававшейся совместными усилиями ЦНСХБ и ВНИИТЭИагропром. Наряду с традиционным каталогом (на бумажных карточках) появился электронный каталог. Сегодня читатели, работающие в зале, проводят поиск библиографии так, как им удобней. Отказ от бумажного каталога не планируется.

В начале 1990-х годов библиотека приобрела несколько мировых баз данных по сельскому хозяйству на компакт-дисках и предоставила читателям доступ к этой информации.

С появлением сети Интернет ЦНСХБ стала частью мирового информационного сообщества. Любой желающий за доступную плату может обучиться основам поиска информации в Интернете или получить машинное время для самостоятельной работы. Библиотека участвует в ряде федеральных и международных проектов ("Libnet", "Libweb", "Российские библиотеки в Интернет" и др.), благодаря чему с компьютеров нашего Интернет-класса можно получить более полную (иногда полнотекстовую, а не только библиографическую) информацию, чем общедоступным образом.

Заметно возросли возможности тех, кто использует информационные ресурсы ЦНСХБ через Интернет. Посетив наш сервер ([www.cnsnb.ru](http://www.cnsnb.ru)), Вы можете получить доступ к нескольким мировым базам данных по сельскому хозяйству, сведения о новых поступлениях (каталог с постатейной росписью обновляется еженедельно), познакомиться с нашими информационными партнерами и, разумеется, поработать с базами данных собственной генерации ЦНСХБ.

Остановимся на возможностях используемой в ЦНСХБ базы данных (БД) "Агрос" и особенностях поиска в ней. БД содержит информацию о статьях из отечественных и зарубежных журналов, периодических и продолжающихся изданиях, сборниках, материалах конференций, книгах и других источниках, имеющихся в фонде ЦНСХБ. Она разбита на 3 подбазы: 1985–1989 гг., 1990–1995 гг. и с 1996 г. по настоящее время.

Для того, чтобы документ впоследствии мог быть найден в БД, при поступлении в библиотеку он подвергается так называемой аналитико-синтетической обработке. Ее проводят сотрудники отдела систематизации – специалисты в области сельскохозяйственной биологии или сельского хозяйства. В соответствии с рядом нормативных документов они индексируют каждый документ по рубрикам ГАСНТИ (АСНТИсельхоз), комплексно-системного каталога ЦНСХБ (КСК), шифрам централизованной библиографической информации, индексам УДК; указывают ключевые слова. Для иностранных документов (при необходимости и для отечественных) составляется аннотация на русском языке, кратко выражающая цель проведенного исследования и суть статьи.

В качестве ключевых слов выступают отобранные из текста документа слова или словосочетания, раскрывающие важнейшие смысловые аспекты документа. Обязательно отражается объект исследования. Обычно индексируются действия с объектом, условия опыта, область применения результатов, география исследования.

В ЦНСХБ разработан и используется в практике индексирования тезаурус по сельскому хозяйству и продовольствию, содержащий наиболее употребительные термины (так называемые дескрипторы), используемые в данной области науки, и отражающий иерархические, синонимические и ассоциативные связи между ними. Вышестоящий дескриптор (В) означает родовые понятия либо целое по отношению к данному дескриптору, являющемуся частью этого целого. Нижестоящий дескриптор (Н) означает либо видовое понятие, либо часть по отношению к понятию, представленному вышестоящим дескриптором. Ассоциативный дескриптор (А) связан с данным дескриптором семантической связью, характер которой не указывается. В тезаурусе отражены также слова, запрещенные к употреблению в качестве терминов индексирования и представляющие собой синонимы или омонимы дескрипторов (например, белки–животные и белки–молекулы). От них дается отсылка ИСП (используй) или СМ (смотри). Если термин используется в разных областях знания, но в каждой имеет специфическое значение, дается уточнение (например, вместо понятия *плод* вводятся понятия *плод ботанический* и *плод утробный*).

Использование тезауруса позволяет преодолеть неоднозначность естественного языка (разные формулировки одного и того же понятия), единообразно и точно переводить слова естественного языка на доступный компьютерной обработке информационно-поисковый язык.

Разные элементы описания документа разнесены в БД в отдельные поля, обозначенные 3–4-буквенным кодом. Поиск можно вести по всему тексту описания документа или по отдельным полям (их сочетанию). Правила построения запросов с примерами содержатся в отдельном файле, доступном пользователю.

При поиске информации в БД “Агрос” используется информационно-поисковая система (ИПС) “Артефакт”, учитывающая многие нюансы русского и английского языков. Система “узнает” знакомое слово в любом роде, числе или падеже. Можно указать в запросе только начальные буквы, а после них поставить “звездочку” или, напротив, указать в кавычках только точную форму слова. Можно задать порядок следования слов и расстояние между ними. Двоеточие между терминами означает, что первое слово должно стоять раньше второго. Цифра после двоеточия (в том числе 0) определяет максимальное количество слов, разделяющих понятия. Можно указать, что искомые слова могут появиться в любом порядке в пределах одного предложения или группы нескольких смежных предложений. Для этого к формулировке запроса добавим /*п*п, где *п* – желаемое число предложений.

Для построения сложных поисковых предписаний в ИПС имеются логические операции, задаваемые операторами И, ИЛИ, НЕ.

В ИПС “Артефакт” реализована система иерархических связей между терминами тезауруса. В поле ТЕР\ (термины тезауруса) вводятся дескрипторы тезауруса, а в поле РТЗ\ (расширение по тезаурусу) автоматически попадают все дескрипторы, стоящие выше по отношению к употребленному. Если написанные систематизатором ключевые слова не вошли в тезаурус, они отражаются в поле КЛС\ (ключевые слова). Разработчиками программы дорабатывается автоматическое приведение синонимичных понятий к единому варианту. К сожалению, при составлении поискового предписания действующий в ИПС тезаурус не доступен пользователю. Поэтому понять, какие из используемых формулировок являются дескрипторами тезауруса, а какие – ключевыми словами, можно лишь, просмотрев найденные документы.

В ответ на запрос пользователей ИПС “Артефакт” создает подборку документов в виде их библиографического описания, реферата или аннотации и других элементов поискового образа документа (ключевые слова, индексы УДК, рубрики КСК и т.д.). Первоисточники (журналы, книги, авторефераты) могут быть при необходимости получены пользователем в печатном виде. Шифр хранения указан в соответствующем поле описания (ШХР). Группа найденных документов может быть сохранена на компьютере пользователя в виде текстового файла.

В целях оптимизации самостоятельного поиска читателей в БД “Агрос” в отделе систематизации ЦНСХБ ведется научное исследование по выявлению факторов, влияющих на полноту и релевантность поиска. Установлено, что основными из них являются качество аналитико-синтетической обработки литературы и знание пользователем как возможностей и особенностей ИПС, так и основ индексации научных документов.

В ходе исследования мы анализировали, как разные варианты составления одного и того же запроса влияют на результат поиска. На первом этапе поиск рекомендуется вести по ключевым словам (в широком смысле), не разбивая запрос на поля, а затем “отсекать все лишнее”, добавляя (или исключая) другие поисковые признаки. При оценке предварительных результатов пользователь должен проанализировать, почему в выборку попали документы, формально отвечающие запросу, но на самом деле не относящиеся к делу, и уточнить формулировку или логику запроса.

Нами выявлен ряд типичных накладок. Бывает, что заданное слово совпадает с названием журнала или фамилией автора. Тогда необходимо указать, в каком поле следует искать термин.

Например, в ответ на запрос *петухов* система находит документы, где Петухов – фамилия автора, слово “петухи” встречается в заглавии статьи и ПЕТУХИ – дескриптор тезауруса. Следует составить запрос либо /АВТ (*петухов*) либо *петухи* /АВТ (*не петухов*). Соответственно будут найдены или документы, в которых Петухов – автор, или документы, в которых речь идет о петухах. Запрос *генетика микроорганизмов* /ЕДХР(*не генетика*) подразумевает, что из выборки исключены статьи, опубликованные в журнале “Генетика”.

Если, указывая в запросе страну, не уточнить, в каком поле следует искать ее название, могут найтись документы, изданные в этой стране, а не описывающие исследования в ней.

Если формулировка запроса обнаружена в поле РТЗ, желательно выбрать более узкий термин.

Пользователю рекомендуется использовать возможность задавать расстояние между искомыми словами или указывать, что слова должны встретиться в пределах одного или нескольких предложений.

Как показал опыт, оптимальное число используемых для поиска понятий – два (понятие может состоять из 2–3 слов, следующих одно за другим); обычно это объект исследования и связанная с ним проблема.

Сужая или расширяя тему запроса, проводя разбивку на поля, уточняя правильность формулировок, используя возможности ИПС “Артефакт”, можно получить достаточно полную и релевантную подборку библиографии по любой теме, отраженной в БД “Агрос”.

В течение 2000 года планируется разработка методического пособия для читателей по самостоятельному поиску, поэтому мы будем благодарны всем, кто поделится с нами своим опытом работы с нашей БД, и постараемся своевременно учесть все пожелания и замечания.

У ряда документов имеется не только библиографическая, но и фактографическая информация (реферат). К сожалению, в настоящее время из-за организационно-финансовых трудностей рефератами снабжается гораздо меньшая часть входного потока документов, чем несколько лет назад. К написанию рефератов привлекаются внештатные сотрудники, а имеющиеся кадры в первую очередь направляются на подготовку изданий ЦНСХБ – реферативных журналов “Ветеринария”, “Экологическая

безопасность в АПК", "Пищевая и перерабатывающая промышленность". С 1999 года в рамках федеральной программы начато финансирование реферативной БД "Повышение плодородия почв России".

Одно из последних достижений информационных технологий в библиотечном деле – электронная доставка, включающая поиск, сканирование и пересылку пользователю документа-первоисточника по электронной почте. В ЦНСХБ проводится предварительная проработка подобного проекта. Не исключено, что он будет воплощен в жизнь в обозримом будущем.

В заключение отмечу, что сотрудники ЦНСХБ всегда стремятся максимально помочь своим читателям (а теперь и пользователям Интернета) в поиске нужной информации, а руководство библиотеки открыто для сотрудничества со всеми заинтересованными организациями и гражданами.

*Е.В.Климова*, к.б.н., ст. науч. сотр.

Отдел систематизации ЦНСХБ РАСХН, г. Москва;

e-mail: evklimov@valise.com