

УДК 025.4
ББК 73

ЯЗЫКИ ТЕМАТИЧЕСКОГО ПОИСКА В ЭЛЕКТРОННОМ КАТАЛОГЕ И ИНДИВИДУАЛЬНОЕ ПОИСКОВОЕ ВЗАИМОДЕЙСТВИЕ

© Г.А. Скарук, 2005

Государственная публичная научно-техническая библиотека
Сибирского отделения Российской академии наук
630200, г. Новосибирск, ул. Восход, 15

Требование «дружественности к пользователю» должно предъявляться и к лингвистическому обеспечению электронных библиотечных каталогов. Поиск в электронном каталоге с использованием различных типов ИПЯ рассматривается как когнитивно-коммуникативный процесс. Свойства ИПЯ оцениваются с точки зрения различных концепций комфортности информационной среды. Многие из этих свойств изменяются в электронной среде. Обсуждается новый подход к обеспечению простоты поиска, связанный с замещением функций пользователя автоматическими процедурами. Поиск описывается как опосредованное либо непосредственное взаимодействие множества участников. Решению возникающих в связи с этим проблем способствует применение комплекса ИПЯ.

Ключевые слова: электронный библиотечный каталог, информационный поиск, информационно-поисковые языки, коммуникативное взаимодействие, комфортность поиска.

По мере вовлечения библиотеки в глобальную электронную среду, роста числа ее удаленных пользователей заметно изменяется и характер помощи пользователям интерактивных библиотечных каталогов. Такого рода помощь все больше ориентируется на их самообслуживание, опосредуется через систему «подсказок», инструкций и автоматических процедур. Это требует детального учета особенностей поискового поведения пользователей. И здесь наглядно проявились противоречия между необходимостью формализации, унификации поисковых средств и необходимостью учета индивидуальных коммуникативных механизмов поискового взаимодействия.

Эти противоречия в полной мере проецируются на языки тематического поиска в электронном каталоге (ЭК), ведь именно они выполняют самую важную роль в диалоге пользователя с поисковой системой, описывая семантику информационной потребности, и именно при их использовании проявляются индивидуальные особенности мышления и личностного знания. «Дружественность к пользователю» – одно из основных требований к современным электронным каталогам. На наш взгляд, оно должно относиться и к их лингвистическому обеспечению. Поэтому сегодня к информационно-поисковым языкам (ИПЯ) ЭК применяются многие теоретические положения из этой области.

1. Поиск как процесс коммуникации.

При оценке совместимости пользователя с ЭК информационный поиск чаще всего рассматрива-

ется как процесс «коммуникативного взаимодействия» /1, 2/. Содержание поиска определяется как «диалог», «общение», в результате которого происходит обмен данными, информацией. Проблемы поиска анализируются на разных уровнях взаимодействия: от уровня «пользователь – система», наиболее распространенного в библиотековедении, до уровня «пользователь – создатель информационного ресурса» с позиций эргономики, психологии, психолингвистики, когнитологии.

2. Поиск как когнитивный процесс.

Основанием для применения когнитивного подхода служит трактовка поисковых языков ЭК как инструментов организации знаний, функционирующих в информационной среде /3, 4/. Электронный каталог – один из ее элементов. Соответственно, требования к информационной среде во многом можно предъявить и к поисковым языкам ЭК.

Термин «организация знаний» был введен в библиотековедение Г.И. Блиссом /5/. Несмотря на дискуссии о правомерности его употребления в таком контексте и аспектах его толкования (с позиций теории экспертных систем, семиотики, когнитологии) /6–8/, это понятие широко используется при анализе поисковых языков.

Представители этого подхода используют базовые положения когнитивной лингвистики. Один из основных тезисов звучит так: возможность хранения в человеческой памяти и использования большого объема многообразных знаний обеспечивается определенной упорядоченностью храни-

мых единиц /9, 10/. По утверждению S. Hannabuss, процессы извлечения знания подобны для лингвистических структур ИПС и семантических структур концептуальных представлений /11/. Это направление представлено работами А.Ц. Масевича /12/, Л. В. Сахарного /13, 14/. За рубежом популярны исследования способов отражения в структуре и лексике поисковых языков менталитета отдельных социальных и профессиональных общностей /15, 16/.

3. Комфортность поиска.

Как показал анализ литературы, в это понятие вкладывается очень разное содержание. В контексте когнитивного подхода формируются две точки зрения, одна из которых ориентирована на «сопровождение» интеллектуальной деятельности, другая – на ее активизацию. Чтобы это проиллюстрировать, сопоставим две, на первый взгляд, полярные позиции.

Развивая мысль И. Дальберга /3/, Ю.А. Шрейдер применял понятие комфортности именно к организации знаний. По его мнению, комфортную информационную среду должны характеризовать возможность корректировки информационной потребности и легкость использования, способность «не предъявлять существенных и неожиданных трудностей», не требовать от пользователя «смены привычных установок» /4/.

Статья Т.Б. Соколовской была написана много раньше и представляет точку зрения, немного забытую сегодня. В ней комфортность описывается как «такой способ организации информационной среды, который создает условия для наиболее эффективной переработки информации», «наличие информационных предпосылок, стимулирующих творческое мышление» /17/. Отсюда следует, по нашему мнению, что информационная среда должна быть организована таким образом, чтобы не только и не столько фиксировать привычные, но и способствовать выявлению новых связей между объектами, принципов их размежевания и подобия, тенденций развития. Информация должна быть представлена в разных контекстах, включена в различные структуры. С этой точки зрения требование легкости использования, конечно, не самое значимое.

Справедливость такого взгляда подтвердили результаты анкетирования пользователей систематического каталога, проведенного в ГПНТБ СО РАН /18/. Это свойство информационной среды оказалось наиболее востребовано научными работниками, которые использовали самые разнообразные поисковые стратегии, обращались и к поиску по узким темам с помощью рубрик алфавитно-предметного указателя к каталогу, и к просмотру семантического окружения темы путем движения по уровням иерархической классифика-

ции. В то же время студенты, ориентирующиеся на простоту поиска, затем чаще других категорий отказывались от литературы из-за ее несоответствия теме (44% отказов).

Попробуем применить эти положения к языкам тематического поиска. В первом случае во главе угла окажется принцип «естественности» ИПЯ /19, 20/, близости к естественному языку, соответствия принципам организации знания в процессах речемыслительной деятельности. Во втором – рассматривается принцип «интегральности» /17, 21/, подразумевающий многоуровневое, многоаспектное представление информации. Аргументы такого плана часто звучат в дискуссиях о преимуществах разных типов поисковых языков.

Решающими при оценке «естественности» ИПЯ, как правило, оказываются соображения о совпадении плана выражения (а отчасти и плана содержания) и естественного, и поискового языков /22, 23/. С этих позиций чаще всего предпочтение отдается вербальным ИПЯ. Но, учитывая, что в современных ЭК язык индексирования запроса может отличаться от языка индексирования документа, этот аргумент уже нельзя относить только к вербальным языкам.

Кроме того, и здесь есть определенные проблемы, которые не всегда осознают пользователи. Одна из них – распространенное представление о совпадении плана выражения и плана содержания естественного языка и контролируемых вербальных ИПЯ, с которым не соглашаются авторитетные специалисты /14, 24/. Об этом писал Л.В. Сахарный, указывая на затруднения, возникающие из-за «обманчивой “естественности”», приводящей к возможным несовпадениям с привычным планом выражения (формой) специфического для ИПС плана содержания» /14, с. 19/.

Другая проблема – несовпадение индивидуального словаря пользователя и лексики контролируемого ИПЯ. По данным нашего анкетирования /20/, сразу нашли соответствующие теме рубрики алфавитно-предметного указателя только 40% студентов, 46% ученых, 56% специалистов.

Оба подхода (с точки зрения «естественности» и интегральности) применяются при анализе принципов структурной организации нормативных словарей контролируемых ИПЯ, структуры сложных индексов, способов задания синтагматических отношений между лексическими единицами.

При оценке гибкости структуры поисковых образов документа и запроса речь ранее велась прежде всего о преимуществах посткоординатных ИПЯ. Проблемой же предкоординатных ИПЯ всегда считались трудности психологического характера при формулировке поискового предписания, связанные с потребностью варьирования формы высказывания в речевой коммуникации /13/.

Но сегодня благодаря появлению возможности поиска в ЭК в режиме посткоординации (поиска по каждому отдельному элементу сложного индекса) эта проблема снимается.

Нормативные словари (авторитетные файлы) ИПЯ оцениваются по двум позициям: с точки зрения их обозримости и с точки зрения жесткости структуры. Большинство авторов указывают на преимущества по обозримости иерархических классификаций /25, 26/. Требование обозримости предъявляется и к выданным в ответ на запрос массивам библиографических записей. Один из возможных вариантов – структурирование выдачи большого объема. Такое сегодня возможно в карточных систематических каталогах. Его могли бы обеспечить семантически мощные иерархические классификации. Но, по утверждению Э.Р. Сукиасяна, этот механизм пока нигде не реализован /27/. Взгляды же на жесткость структуры словарей также меняются, так как авторитетные файлы обеспечивают поэлементный доступ к сложным предметным рубрикам и классификационным индексам.

4. Простота поиска.

Самый очевидный способ обеспечить легкость использования ИПЯ – сокращение количества составных элементов (лексических единиц, грамматических средств) и связей между ними в самом языке. Именно в таком понимании принцип доступности был сначала реализован в зарубежных ЭК, и именно этот аспект проблемы был предметом дискуссий в середине 1980-х гг. Например, J.A. Adams /28/, V. Raval /29/ высказывали мнение, что ориентация на дружелюбность к пользователю ведет к простым и неглубоким стратегиям поиска и в конечном итоге ограничивает доступ к информации.

Такое понимание простоты поиска справедливо лишь в определенных пределах. Это наглядно подтверждает история развития поисковых машин Интернета, где постепенно усложняются механизмы поиска и вводятся элементы структурирования информации. Значительное большинство участников анкетирования /30/ сообщили, что ориентируются на поисковые машины, обеспечивающие большую полноту и релевантность поиска, и лишь 22,6% – на простоту использования этих машин.

Новый подход к обеспечению простоты использования поисковых языков основывается на рассмотрении ЭК как экспертной посреднической системы. Появляются возможности представления структуры и лексики ИПЯ с разной полнотой для разных категорий пользователей; автоматического управления построением и корректировкой поискового предписания (в том числе автоматического перевода запроса с естественного языка на ИПЯ); автоматического выбора ИПЯ. Но нельзя забывать, что такое замещение функций пользователя может

иметь и положительные, и отрицательные последствия. По мнению Э.Р. Сукиасяна, важно, чтобы при этом не было утрачено когнитивное значение каталога и за читателем сохранилась возможность выбора путей поиска /31/.

В связи с этим изменяются и традиционные оценки сложности использования поисковых языков. Так, менее значимыми становятся аргументы по поводу сложности ориентации в структуре иерархических классификаций. Необходимо учитывать, что каждый пользователь карточного систематического каталога в любом случае должен был выполнять один и тот же набор шагов. В ЭК же для неопытных пользователей часть этих процедур может выполняться «по умолчанию». В то же время возможности раскрытия структуры и содержания ИПЯ в авторитетных файлах гораздо шире, чем в печатных таблицах классификации. К ним могут обратиться пользователи с высокой квалификацией.

5. Вариативность поиска.

Другое следствие рассмотрения проблем информационного поиска на уровне индивидуальных межсубъектных взаимодействий – усиление внимания к проблеме вариативности его процессов и результатов. В противовес подходу к оценке ЭК на уровне «пользователь – система» сегодня исследователи ориентируются на оценку деятельности значительно большего числа участников. А значит, оценивается взаимодействие в процессе поиска «когнитивных структур», «дискурсов», «тезауросов» всех задействованных в поиске сообществ.

Роль отдельных действующих лиц в цепочке взаимодействий, обеспечивающих эффективный поиск, легко проследить, суммировав данные публикаций, описывающих отдельные его стороны. Когнитивные структуры конструкторов и производителей автоматизированной библиотечно-информационной системы представлены в специфической структуре конкретных баз данных, функциях пользовательских интерфейсов, критериях релевантности, алгоритмах поиска, реализуемых системой /32/. Автор излагает собственную концепцию, вводит новые термины и понятия, произвольно структурирует текст. Субъективизм индекса проявляется не только при выделении основного содержания (предмета, ведущего признака) авторского текста, но и в произвольном отборе терминов из нормативных словарей (либо текста), определении их роли в поисковом образе документа /12, 14/. Доказательство тому – результаты многочисленных экспериментов по устойчивости индексирования /33, 34/. Е.Н. Пименов, характеризуя влияние прагматических факторов на процесс и результат индексирования, выделяет 12 таких условий, в том числе общее назначение базы данных, объем обрабатываемой информации,

состав документов в системе, штаты и квалификацию персонала /35/. Кроме того, как справедливо отмечает Э.Р. Сукиасян, индексы имеют дело не с идеальной моделью ИПЯ (ИПЯ как системой), а с его овеществленным носителем, например с классификационными таблицами конкретного ИПЯ определенного года издания /36/. Выбирая поисковые термины, информационный посредник рассматривает тему (предмет) поискового запроса в перспективе многих занятых в дискурсах сообществ и при этом оценивает и синтезирует элементы отличия и сходства /37/. Эти же операции должен выполнять и пользователь каталога. Данный процесс усложняется, когда искомые знания относятся к «разным, но исследующим одну сущность мира системам знания» /38/.

Отсюда вывод о потенциальной ценности всех участвующих в поиске когнитивных структур. П. Ингверсен говорит в этом контексте о необходимости многоуровневого представления как семантических объектов (документов, текстов), так и характеристик потребностей пользователя /32/. А значит, мы вновь возвращаемся к принципу интегральности информационной среды и необходимости применения при поиске нескольких ИПЯ. С другой стороны, на примере поиска по полным текстам обосновывается неприемлемость механического наращивания числа «точек доступа» и, как следствие, вывод, что в любой поисковой среде не удастся избежать применения специальных языков индексирования /39/.

Кроме того, для решения этого комплекса взаимосвязанных проблем в ЭК предлагается расширение семантического поля запроса. Механизмы такого расширения в наибольшей степени предусмотрены в цифровых иерархических классификациях.

Очевидно, что возможности различных типов ИПЯ по обеспечению рассмотренных в статье принципов дополняют друг друга и будут необходимы в разных поисковых ситуациях. В то же время эти возможности настолько меняются в электронной среде, что свойства многих ИПЯ сближаются, многие традиционные аргументы «за» или «против» применения тех или иных лингвистических средств в ЭК теряют силу. Но в любом случае итогом дискуссий об их преимуществах, как нам кажется, должен стать вывод о необходимости комплексного применения языков тематического поиска.

Список литературы

1. Лавренова, О.А. Лингвистическое обеспечение электронных каталогов: разработка средств тематического поиска / О.А. Лавренова // Мир библиотек сегодня. – 1997. – Вып. 15. – С. 60–70.
2. Тоом, А.А. Причины дискомфорта в диалоге с информационно-поисковой системой и пути его преодоления / А.А. Тоом // НТИ. Сер. 2. Информ. процессы и системы. – 1987. – № 4. – С. 1–6.
3. Дальберг, И. Организация знаний: ее сфера и возможности / И. Дальберг // Организация знаний: проблемы и тенденции. – М., 1993. – С. 7.
4. Шрейдер, Ю.А. Категория комфортности как основной критерий организации знаний / Ю.А. Шрейдер // Междунар. форум по информ. и документации. – 1993. – Т. 18, № 3–4. – С. 17–18.
5. Bliss, H.E. The organization of knowledge in libraries / H.E. Bliss. – N.Y., 1993. – P. 36–37.
6. Мизинцева, М.Ф. Представление знаний в системах искусственного интеллекта / М.Ф. Мизинцева // НТИ. Сер. 1. Орг. и методика информ. работы. – 1998. – № 12. – С. 1–6.
7. Сукиасян, Э.Р. Организация знаний как элемент информационной культуры / Сукиасян Э.Р. // Формирование информационной культуры читателей: роль новых технологий. – Новоуральск, 2000. – С. 14–21.
8. Черный, А.И. О проблемах организации и представления знаний / А.И. Черный // Междунар. форум по информ. и документации. – 1998. – Т. 23, № 2. – С. 3–8.
9. Петров, В.В. На пути к когнитивной модели языка / В.В. Петров, В.И. Герасимов // Новое в зарубеж. лингвистике. Вып. 23. Когнитивные аспекты языка. – 1988. – С. 5–11.
10. Пиаже, Ж. Психогенез знаний и его эпистемологическое значение / Ж. Пиаже // Семиотика. – М., 1983. – С. 90–101.
11. Hannabuss, S. Dialogue and the search for information / S. Hannabuss // Aslib. Proc. – 1989. – Vol. 41, № 3. – P. 85–98.
12. Масевич, А.Ц. Два подхода к теории ИПС в свете современных лингвистических концепций / А.Ц. Масевич // Предмет. поиск в традиц. и нетрадиц. информ.-поисковых системах. – 1989. – Вып. 9. – С. 25–50.
13. Сахарный, Л.В. Ассоциативные структуры в естественном языке и предметизационном ИПЯ (к постановке проблемы) / Л.В. Сахарный // Там же. – 1988. – Вып. 8. – С. 5–35.
14. Сахарный, Л.В. Предметная рубрика как вид текста / Л.В. Сахарный // Там же. – 1987. – Вып. 7. – С. 16–37.
15. Albrechtsen, H. The dynamics of classification systems as boundary object for cooperation in the electronic library / H. Albrechtsen, E.K. Jacob // Libr. Trends. – 1998. – Vol. 47, N 2. – P. 293–312.
16. Olson, A.H. Mapping beyond Dewey's Boundaries: Construction classificatory space for marginalized knowledge domains / A.H. Olson // Ibid. – P. 233–254.
17. Соколовская, Т.Б. О комфортности информационной среды / Т.Б. Соколовская // НТИ. Сер. 2. Информ. процессы и системы. – 1980. – № 11. – С. 5–10.
18. Скарук, Г.А. Помощь пользователю электронного каталога: ожидания читателей и возможности поисковой системы / Г.А. Скарук // Место и роль библиотеки в инфраструктуре образовательного пространства: Сибирь и Дальний Восток. – Новосибирск, 2003. – С. 84–96.
19. Саффиуллина, З.А. АСОД как научная проблема и предмет преподавания / З.А. Саффиуллина // Библиография. – 1998. – № 3. – С. 3–9.
20. Щербинина, Г.С. Философия координатного индексирования / Г.С. Щербинина // Науч. и техн. б-ки. – 2000. – № 9. – С. 72–76.

21. Новиков, Ю.А. Психологические критерии отбора информации / Ю.А. Новиков // НТИ. Сер. 1. Орг. и методика информ. работы. – 1970. – № 7. – С. 3–8.
22. Гальцева, Т.Е. Электронный каталог в библиотеке вуза / Т.Е. Гальцева // Науч. и техн. б-ки. – 1996. – № 8. – С. 6–11.
23. Зайцева, Е.М. Лингвистическое обеспечение автоматизированных информационно-библиотечных систем: современные требования и направления развития / Е.М. Зайцева // Науч. и техн. б-ки. – 2000. – № 3. – С. 54–57.
24. Ручимская, Е.М. Явление вариативности естественного языка и способы ее устранения в вербальных ИПЯ / Е.М. Ручимская // НТИ. Сер. 1. Орг. и методика информ. работы. – 2000. – № 2. – С. 24–26.
25. Гендина, Н.И. ББК как средство организации знаний и предмет изучения в высшей библиотечной школе: концептуальные подходы / Н.И. Гендина // ББК: Новые горизонты орг. знаний : материалы Междунар. науч. конф. (14–16 октяб. 1998 г., Москва). – М., 1999. – С. 86–100.
26. Maniez, J.A. classifications still relevant in databases? / J.A. Maniez // Doc. Languages and Databases. – Frankfurt/M, 1991. – P. 120–129.
27. Сукиасян, Э.Р. «Предметный вход» в систематический каталог. Как повысить эффективность поисковых процедур [Электронный ресурс] / Э.Р. Сукиасян // Библиотеки и ассоциации в меняющемся мире : материалы Междунар. конф. «Крым-2003». Судак, Авт. Респ. Крым, Украина, 7–15 июня 2003 г. – М., 2003. – Т. 2. – Режим доступа: <http://www.gpntb.ru/win/inter-events/crimea2003/index2.htm>
28. Adams, J.A. The computer catalog: a democratic or authoritarian technology? / J.A. Adams // Libr. J. – 1988. – Vol. 113, N 2. – P. 31–36.
29. Raval, V. Do user-friendly systems enhance control objectives? / V. Raval // J. Account. And EDP. – 1989. – Vol. 5, N 1. – P. 24–27.
30. Vaughan, J. Considerations in the choice of in Internet tools / J. Vaughan // Libr. Hi Tech. – 1999. – Vol. 17, N 1. – P. 89–106.
31. Сукиасян, Э.Р. Homo Quaeiens (Человек ищущий). К проблеме развития познавательных способностей читателя в процессе информационного поиска [Электронный ресурс] / Э.Р. Сукиасян // Науч. и техн. б-ки. – 2002. – № 4. – Режим доступа: http://www.gpntb.ru/win/ntb/ntb2002/4/f4_13.htm
32. Ингверсен, П. Когнитивная перспектива в информационном поиске / П. Ингверсен // Междунар. форум по информ. и документации. – 1994. – Т. 19, № 2. – С. 25–32.
33. Масевич, А.Ц. Проявления субъективного фактора в традиционных каталогах: теоретический и методический подходы к исследованию / А.Ц. Масевич, И.Н. Комарова, В.Б. Лазуркина // Предмет. поиск в традиц. и нетрадиц. информ.-поисковых системах. – 1990. – Вып. 10. – С. 109–125.
34. Connell, T.H. Use of the LCSH system: Realities / T.H. Connell // Cat. and Classif. Quart. – 1996. – Vol. 23, N 1. – P. 73–98.
35. Пименов, Е.Н. О факторах, влияющих на индексирование: индексирование и предметная область / Е.Н. Пименов // НТИ. Сер. 1. Орг. и методика информ. работы. – 2000. – № 2. – С. 15–23.
36. Сукиасян, Э.Р. Системный анализ проблем управления качеством информационно-поисковой системы / Э.Р. Сукиасян // Науч. и техн. б-ки. – 1995. – № 3. – С. 5–15.
37. Iivonen M. From translation to navigation of different discourses: A model of search term selection during the pre-online stage of the search process / M. Iivonen, D.H. Sonnenwald // J. Amer. Soc. Inf. Sci. – 1998. – Vol. 49, N 4. – P. 312–326.
38. Бернштейн, Э.С. Пробельный анализ как метод информатики / Э.С. Бернштейн // НТИ. Сер. 2. Информ. системы и процессы. – 1983. – № 6. – С. 1–15.
39. Fugmann, R. The empirical approach in the evaluation of information systems / R. Fugmann // Know. Organ. – 1999. – Vol. 26, N 1. – P. 3–9.

Материал поступил в редакцию 26.04.2005 г.

Сведения об авторе: Скарук Галина Анатольевна – кандидат педагогических наук, старший научный сотрудник центра научной обработки документов, тел. (383) 266-92-09