

ИНФОРМАЦИОННО-ПОИСКОВЫЕ ЯЗЫКИ И СИСТЕМЫ

УДК 025.4.036

doi: 10.33186/1027-3689-2021-11-131-146

Е. М. Зайцева

ГПНТБ России, Москва, Российская Федерация

Организация поискового интерфейса библиотечно-информационных систем: простота *versus* эффективность

Аннотация. Основная цель статьи – обобщить проблемные вопросы, связанные с организацией поиска, в особенности тематического, в библиотечно-информационных системах, выявить возможности их решения. Рассмотрены основные требования к поисковому интерфейсу электронных каталогов, отражённые в зарубежной и отечественной литературе. Выявлены и проанализированы проблемы, связанные со стремлением разработчиков библиотечно-информационных систем обеспечить простоту поискового интерфейса. Описаны варианты первых поисковых страниц электронных каталогов и электронных библиотек российских и зарубежных библиотек. Предложен вариант поисковой страницы, включающий единую поисковую строку и возможность расширенного поиска с наглядным отражением опций библиографического, тематического, полнотекстового и комбинированного поиска. Даны определения всех предлагаемых видов поиска на основе действующих стандартов СИБИД или собственные дефиниции – при отсутствии рассматриваемых терминов в стандартах. Мотивировано обязательное присутствие тематического поиска в библиотечно-информационной системе, при этом желательно различных видов и разных уровней сложности. Сформулированы с детальными пояснениями основные требования к поисковым средствам библиотечно-информационных систем.

Статья подготовлена в рамках государственного задания ГПНТБ России 730000Ф.99.1.БВ09АА00006.

Ключевые слова: поисковый интерфейс, электронные каталоги, электронные библиотеки, электронные архивы, библиографический поиск, тематический поиск, полнотекстовый поиск, комбинированный поиск, классификационные системы, предметные рубрики, ключевые слова

INFORMATION RETRIEVAL LANGUAGES AND SYSTEMS

UDC 025.4.036

doi: 10.33186/1027-3689-2021-11-131-146

Ekaterina M. Zaitseva

*Russian National Public Library for Science and Technology,
Moscow, Russian Federation*

Structuring information retrieval interface in LIS: Simplicity vs. efficiency

Abstract. The author attempts to generalize on the problems of providing retrieval, in particular, subject search, in library information systems and to find solutions. The key requirements to e-catalog retrieval interface as considered in foreign and domestic publications are presented. The problems of providing simplicity of retrieval interface are analyzed. The diverse first search pages of OPACs and e-catalog of Russian and foreign libraries are described. The author suggests a search page version comprising unit search box with the possibility of advanced search and user friendly option of bibliographic, subject, full-text and complex search. Every proposed search types are defined based on SIBID standards, or in case the terms are not provided by existing standards, the author provides her own definitions for them. The requirement of mandatory subject search in library information system, desirably for various types and degrees of complexity, is substantiated. Key requirements to retrieval tools of library information systems are formulated and detailed.

The article is prepared within the framework of the State Order to RNPLS&T No. 730000F.99.1.BV09AA00006.

Keywords: search interface, OPAC, electronic catalog, electronic library, digital archives, bibliographic search, subject search, full-text search, complex search, classification systems, subject headings, keywords

Поисковые интерфейсы библиотечно-информационных систем в России и за рубежом продолжают совершенствоваться в рамках действующих систем или в связи с внедрением нового программного обеспечения. Однако поисковые возможности систем по-прежнему вызывают нарекания со стороны пользователей. Попробуем разобраться,

что же не устраивает пользователей, и наметить выходы из сложившейся ситуации.

С начала 2000-х гг. в зарубежной и отечественной литературе [1–3] в качестве основных требований к поисковому интерфейсу электронных каталогов стали выдвигаться следующие критерии: простота, понятность, наглядность, комфортность. В результате это привело к появлению единой точки входа для доступа к ресурсам – единой поисковой строке. Этот вход был перенесён в электронные библиотеки и открытые архивы, что сблизило все указанные системы с поисковыми машинами интернета. Единое поисковое окно с вводом в него текста запроса на естественном языке представлено сегодня на сайтах многих российских и зарубежных библиотек.

Анализ поисковых интерфейсов электронных каталогов и электронных библиотек крупнейших отечественных и зарубежных библиотек, представленных на их веб-сайтах, показывает, что существуют разные варианты первых поисковых страниц.

1. «Жёсткая» единая поисковая строка, в которой обязательно нужно что-то ввести, чтобы перейти к другим вариантам поиска (например, сайт РНБ; рис. 1).

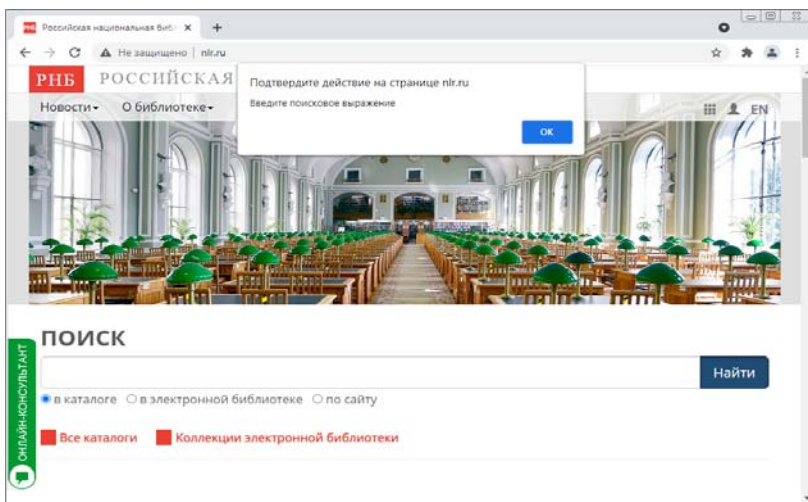


Рис. 1. Первая поисковая страница электронного каталога Российской национальной библиотеки

2. «Мягкая» единая поисковая строка, из которой даже по пустому запросу возможен переход к другим видам поиска (например, сайт РГБ; рис. 2).

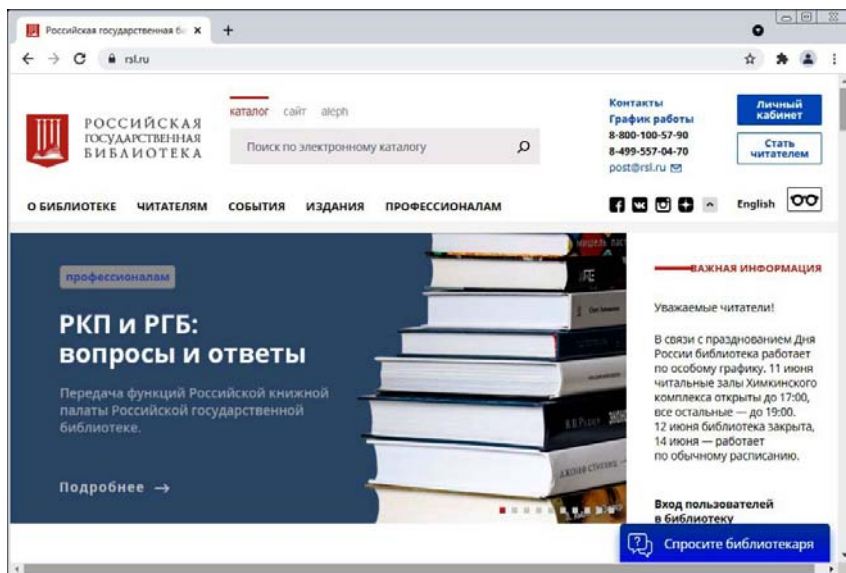


Рис. 2. Первая поисковая страница электронного каталога Российской государственной библиотеки

3. Единая поисковая строка с параллельным отражением обобщённой опции расширенного поиска, позволяющей переходить к разным поисковым возможностям (например, сайты Британской библиотеки и ГПНТБ России; рис. 3 и 4).

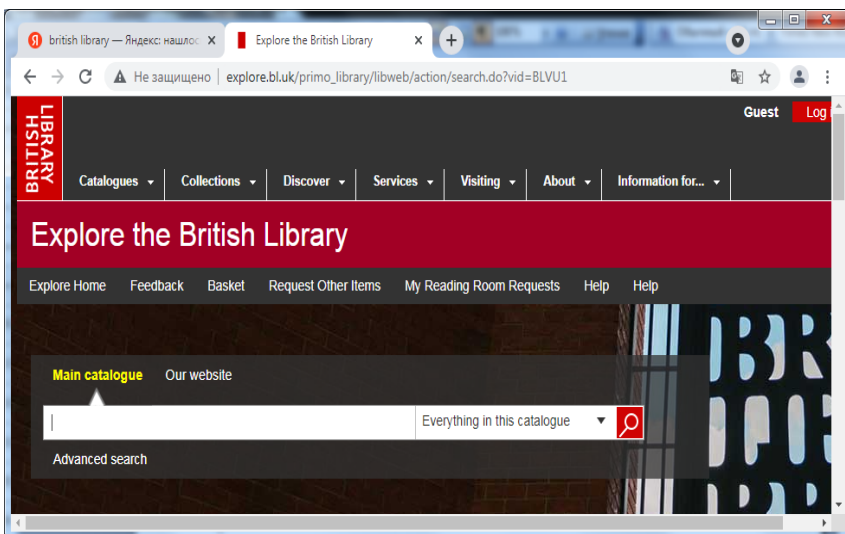


Рис. 3. Первая поисковая страница электронного каталога Британской библиотеки

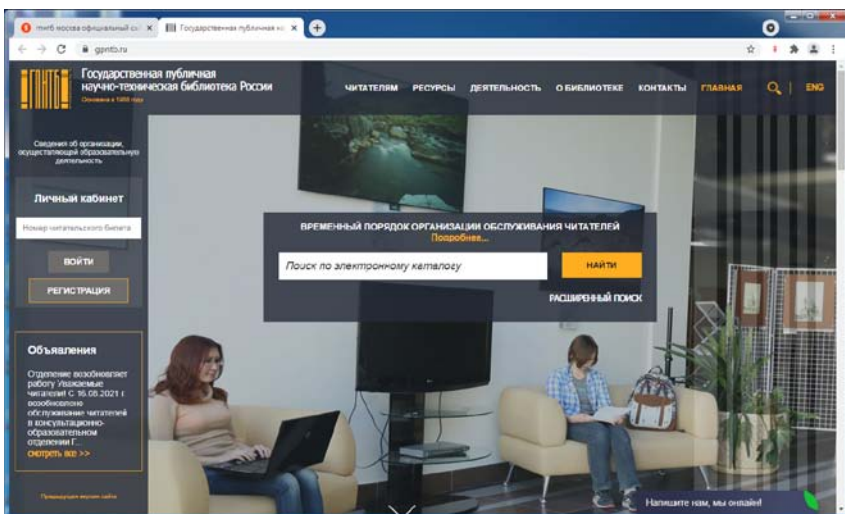


Рис. 4. Первая поисковая страница электронного каталога ГПНТБ России

4. Единая поисковая строка с параллельным отражением основных поисковых опций с их краткой характеристикой, например, сайт Библиотеки Конгресса США (рис. 5): *Browse* – поиск с использованием различных словарей; *Advanced Search* – простой и комбинированный поиск с использованием разных элементов; *Keyword Search* – поиск по ключевым словам, содержащимся в поле заглавия, автора, предметных рубрик или во всех полях.

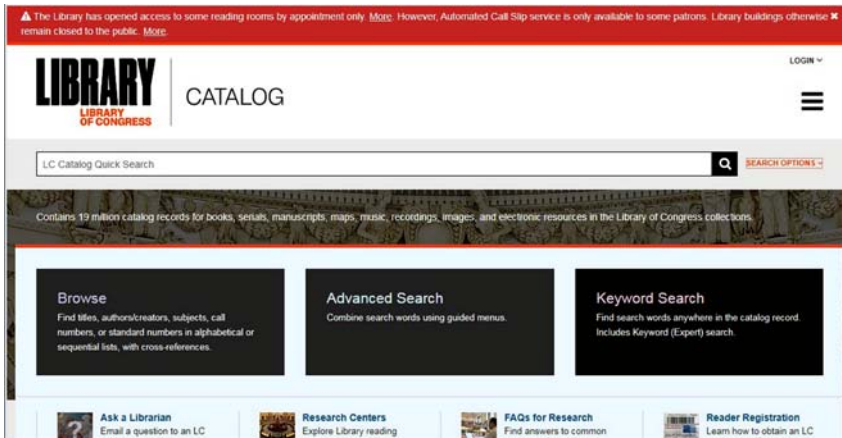


Рис. 5. Первая поисковая страница электронного каталога Библиотеки Конгресса США

Первые два варианта первых поисковых страниц, на мой взгляд, менее удачны, поскольку изначально накладывают некоторые ограничения на реализацию поисковых потребностей пользователей и не раскрывают перед ними в явном виде имеющиеся поисковые средства. Пользователь, ставящий задачу не просто быстро найти хотя бы что-то по интересующей его теме, а желающий получить достаточно полную или максимально полную информацию по конкретной теме, вправе ожидать, что ему будут предоставлены понятные и удобные средства тематического поиска, а на первой странице желательно видеть указание на их наличие.

Третий и четвёртый варианты первых поисковых страниц в разной степени обеспечивают отражение возможных поисковых опций, при этом в четвёртом варианте это отражено более детально.

В качестве ещё одного, пятого, варианта предлагается первая поисковая страница, обеспечивающая, на мой взгляд, необходимую структуризацию поисковых средств, поскольку включает единую поисковую строку и возможность расширенного поиска с одновременным указанием основных опций, принципиально различающихся по используемым поисковым элементам:

1. Библиографический поиск, определяемый как поиск по элементам библиографического описания. Предлагаемое определение отличается от дефиниций, представленных в стандартах – ГОСТе 7.73-96 [4] и ГОСТе 7.0-99 [5], оно более узкое. Для сравнения приведём определение библиографического поиска, данные в стандартах:

библиографический поиск – «информационный поиск в библиографической базе данных» (ГОСТ 7.73-96);

библиографический поиск – «информационный поиск, осуществляемый на основании библиографических данных» (ГОСТ 7.0-99).

В обоих указанных ГОСТах термин *библиографический поиск* имеет широкое толкование, учитывая тот факт, что в системе СИБИД библиографические данные (сведения) трактуются как «сведения о документе, используемые при составлении библиографической записи» [6].

2. Тематический поиск, определяемый как поиск по поисковым образам документов (ПОД), сформированным на классификационном или вербальном языке, который предлагается подразделять на следующие виды:

классификационный поиск, определяемый как поиск с использованием различных классификаций;

предметный поиск, определяемый как поиск по ПОДах, сформированным на языке предметных рубрик или ключевых слов.

В библиотечной теории и практике термины *тематический поиск* и *предметный поиск* трактуются по-разному. В первой трактовке они представляются как эквивалентные понятия. Вторая трактовка рассматривает тематический поиск как более широкое понятие по сравнению с предметным поиском, под которым может пониматься только поиск по предметным рубрикам или поиск по предметным рубрикам и ключевым словам. В этой статье используется последняя трактовка, поскольку она позволяет выделить разные виды тематического поиска. Следует отметить, что в стандартах СИБИД определения тематического

поиска и его видов, а также рассматриваемого ниже комбинированного поиска, отсутствуют.

3. Полнотекстовый поиск – «поиск, при котором в качестве поискового образа документа используется его полный текст или существенные части текста» [4].

4. Комбинированный поиск, определяемый как поиск, в котором реализуются совместно разные возможности библиографического или тематического поиска либо используются средства нескольких видов поиска.

Такое отражение поисковых опций представляется оптимальным, предоставляющим пользователям наиболее полные поисковые возможности в дифференцированном виде. К предлагаемому варианту приближен поисковый интерфейс электронного каталога ГПНТБ России (рис. 6), который раскрывается при входе в расширенный поиск и предоставляет опции поиска по базовым библиографическим элементам (автор, заглавие, издательство, год издания), а также поиска с использованием классификаций УДК и ГРНТИ. В профессиональном поиске (профи) реализуются следующие опции поиска: по предметным рубрикам, ключевым словам и комбинированного. Поиск по полным текстам проводится в единой поисковой строке.

Представляется желательным реализовывать полнотекстовый поиск отдельно ввиду того, что в единой строке объединены разные виды поиска. Также кажется целесообразным представлять поиск по предметным рубрикам и ключевым словам в расширенном поиске, поскольку до профессионального поиска пользователь может просто не добраться.

Тематический поиск в том или ином виде должен быть обязательно представлен в библиотечно-информационной системе, так как он по-прежнему является одним из самых востребованных видов поиска. Анализ годовой статистики поискового поведения пользователей, проведённый в ГПНТБ СО РАН [7], показал, что в расширенном поиске по-прежнему чаще всего используется тематический поиск, на втором месте – обращение к ресурсам по автору.

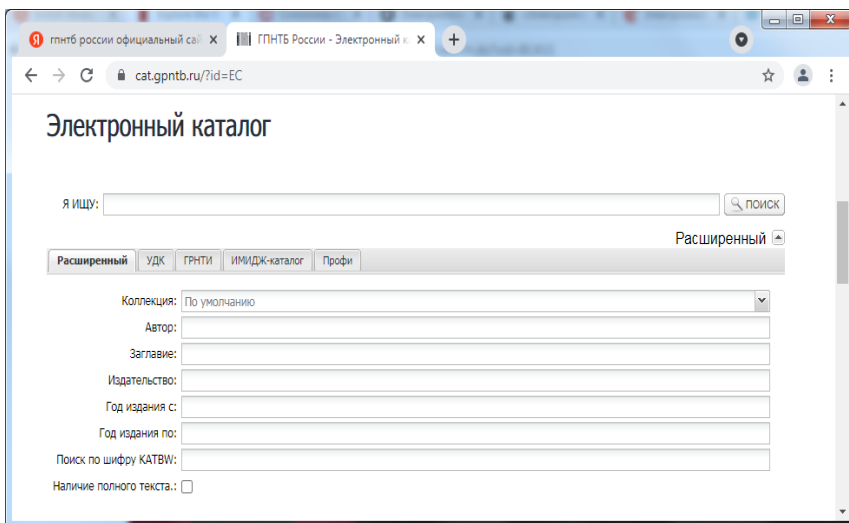


Рис. 6. Расширенный поиск в электронном каталоге ГПНТБ России

Стремление разработчиков библиотечно-информационных систем максимально упростить поиск и поисковые средства и как результат применение пользователем при поиске только единой строки облегчает формулировку запроса, делает поиск более быстрым, но приводит к значительному информационному шуму, особенно если поиск ведётся по всем полям библиографических записей и полным текстам документов, и самое главное – резко снижает pertinентность поиска. Для получения адекватных результатов пользователь должен иметь разные, достаточно богатые и понятные инструменты для создания запроса.

В качестве основной задачи разработчикам следует выдвигать не обеспечение простоты поиска, а создание удобного, понятного поискового интерфейса и методов повышения pertinентности поиска. Пользователю трудно, используя несколько слов, чётко выразить свою информационную потребность в единой поисковой строке. Для облегчения работы его нужно обеспечить необходимыми средствами. Эти средства должны быть разными по своему типу и уровню сложности, с учётом различных потребностей и уровня подготовки пользователя.

Анализ поисковых интерфейсов библиотечно-информационных систем позволил обобщить и конкретизировать известные требования к поисковым средствам (многоуровневость, наглядность, понятность), а также сформулировать дополнительные существенные, на наш взгляд, требования.

1. Требование многоуровневости, многоступенчатости поискового интерфейса. Это требование отмечено, в частности, в Руководстве по обеспечению тематического доступа в национальной библиографии, разработанном ИФЛА: «Обеспечьте интерфейс с возможностью работы с различной степенью детализации. Рекомендуется предложить в дополнение к интерфейсу с простым окном поиска один или несколько других уровней, предлагающих опции расширенного поиска» [8].

В выполнении этого требования разработчики всех библиотечно-информационных систем единодушны. Во всех поисковых интерфейсах мы видим как минимум две возможности – простого и расширенного поиска, а в некоторых системах присутствует ещё и профессиональный поиск. Но на этом однообразие систем заканчивается. А хотелось бы, чтобы однообразие распространялось и на набор поисковых средств, и на форму их представления, что значительно облегчило бы работу пользователя, который может обращаться за информацией в разные библиотеки.

2. Требование наглядности. Все виды поиска должны быть наглядно представлены, обеспечивая возможность быстрого выбора. Такую наглядность может обеспечить пятый вариант основной поисковой страницы, предложенный выше.

3. Требование понятности. Все поисковые средства должны быть понятны и иметь однозначное толкование.

4. Требование дифференцированности. Все поисковые средства должны быть чётко дифференцированы.

Последние два требования представляются весьма важными, поскольку в противном случае сам поиск будет осложнён и пользователю будет трудно разобраться в результатах поиска.

В качестве примера обратимся к поиску в электронном каталоге на сайте РГБ. Здесь в расширенном поиске представлен поиск по элементу «Содержание». При обращении к «Справке о расширенном поиске» выясняется, что поиск по содержанию – это поиск «по ключевым

словам, персоналиям, организациям, географическим названиям и мероприятиям, упоминающимся в документе, словам из аннотации к документу, наименованиям делений (рубрик) Библиотечно-библиографической классификации».

Таким образом, этот элемент включает в себя и библиографические, и тематические данные и, строго говоря, не может лежать в основе собственно тематического поиска. На тематический поиск (в строгом понимании) пользователь может выйти только в профессиональном поиске, который сложен для обычного пользователя.

В рассматриваемом примере в профессиональном поиске, помимо элемента «Содержание», представлены элементы «ББК», «Тема» и «Ключевые слова» без каких-либо пояснений. В поле «ББК» нужно вводить индексы ББК, которые рядовой пользователь вряд ли знает, тем более – их буквенные версии, которые применяет РГБ. Поиск по «Теме» ведётся в среде рубрик ББК, что можно понять только при открытии описаний найденных документов. Из каких полей библиографической записи слова трактуются как «Ключевые слова», нигде не поясняется.

Из приведённого примера напрашивается вывод: желательно использовать более понятные, дифференцированные, конкретные названия для поисковых элементов, такие как «Индексы ББК», «Рубрики ББК», «Предметные рубрики», «Ключевые слова (с конкретизацией)», и избегать таких элементов, как «Содержание», «Тема», имеющих комплексное или неясное толкование.

Особенно коварен и неоднозначен для понимания (без конкретизации) поисковый элемент «Ключевые слова»: в различных библиотечно-информационных системах он трактуется по-разному – ключевые слова: в поле ключевых слов, в предметных рубриках, в заглавиях, в любых полях и др.

5. Требование поискового многообразия. В библиотечно-информационной системе необходимо обеспечить три основных вида поиска: библиографический (по элементам библиографического описания), тематический (желательно и классификационный, и предметный) и полнотекстовый. Комбинированный поиск может фигурировать в расширенном и/или профессиональном поиске. В профессиональном поиске представляется целесообразным расширение набора поисковых элементов и введение более полных возможностей их комбинирования.

6. Требование многоуровневости (наличия нескольких уровней сложности) тематического поиска:

1) простой – в единой поисковой строке. При этом если поиск ведётся по всем полям библиографических записей, желательно обеспечить ранжирование результатов: в первую очередь нужно выводить документы, содержащие искомые термины в заглавиях, предметных рубриках, ключевых словах, аннотациях;

2) средний – по упрощённой классификационной системе. Так в БЕН РАН в тематическом поиске в электронном каталоге применяется специальный рубрикатор, построенный на основе рубрик сокращённой схемы УДК [9]. В системе открытого архива ГПНТБ России предполагается использовать упрощённый тематический рубрикатор, специально разработанный на основе УДК, с указанием соответствий по ГРНТИ и ББК. Его применение призвано уменьшить информационный шум, возникающий при поиске в единой поисковой строке;

3) сложный – по полной классификационной схеме, словарю предметных рубрик, тезаурусу.

7. Требование наличия навигации по схеме/словарю в тематическом поиске. Для успешного функционирования сложного тематического поиска необходимо предоставлять в поисковом интерфейсе навигацию по полной классификационной схеме, словарю предметных рубрик, тезаурусу с возможностью отбора индексов, рубрик, терминов для поиска, а также обеспечивать вербальный вход в классификацию или словарь. Желательно также дать хорошую визуализацию имеющихся в наличии ресурсов с указанием количества документов, представленных по конкретному индексу, рубрике, термину. В качестве примера приведём фрагмент навигации по полной классификации УДК в электронном каталоге ГПНТБ России, организованной по указанным принципам (рис. 7).

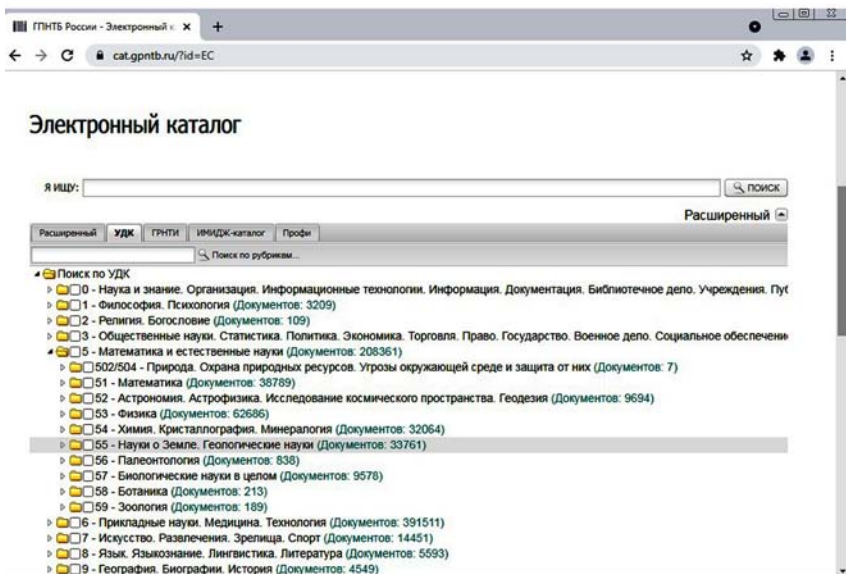


Рис. 7. Навигация по полной классификации УДК в электронном каталоге ГПНТБ России

Вариант поиска с использованием навигации по полной классификации, возможно, требует больше временных затрат, но зато даёт более эффективные результаты. Варианты поиска по классификационным индексам, требующие знания этих индексов (например, в электронном каталоге РНБ), поиска по ключевым словам, входящим в состав классификационных рубрик (например, в электронном каталоге РГБ) и навигации по классификации без указания индексов (например, в электронном каталоге БЕН РАН), явно уступают первому варианту, поскольку не обеспечивают должной наглядности и простоты ни при формировании запроса, ни при просмотре результатов поиска, усложняя тем самым работу пользователя.

При обобщении требований к поисковым средствам специально не указывалось требование простоты. Представляется неправильным выдвигать это требование, особенно в качестве одного из ведущих.

Упрощение поисковых средств неизбежно ведёт к увеличению информационного шума и снижению pertinентности поиска. Поисковые средства должны быть многообразными и удобными, и только тогда они могут гарантировать эффективный поиск.

СПИСОК ИСТОЧНИКОВ

1. **Breeding M.** Next-Generation Library Catalogs. Chapter 1: Introduction // Library Technology Reports. – 2007. – Vol. 43. – № 4. – P. 5–14. – URL: <https://library-technology.org/document/18344> (дата обращения: 10.08.2021).

2. **Yang Sh. Q., Hofmann M. A.** The Next Generation Library Catalog: A Comparative Study of the OPACs of Koha, Evergreen and Voyager // Information Technology and Libraries. – 2010. – September. – P. 141–150. – URL: <https://ejournals.bc.edu/index.php/ital/article/view/3139/2753> (дата обращения: 10.08.2021).

3. **Скарук Г. А.** Комфортность как критерий оценки качества электронного каталога // Тр. ГПНТБ СО РАН. – Вып. 13. – Т. 2. – Новосибирск, 2018. – С. 177–184.

4. **ГОСТ 7.73-96.** Поиск и распространение информации. Термины и определения : межгосударственный стандарт : взамен ГОСТ 7.27-80 : дата введения 1998-01-01 / разработан Всероссийским институтом научной и технической информации РАН [и др.]. – Москва : Издательство стандартов, 1997. – IV, 16 с. – (Система стандартов по информации, библиотечному и издательскому делу).

5. **ГОСТ 7.0-99.** Информационно-библиотечная деятельность, библиография. Термины и определения : межгосударственный стандарт : взамен ГОСТ 7.0-84, ГОСТ 7.26-80 : дата введения 2000-07-01 / разработан Российской государственной библиотекой [и др.]. – Москва : Издательство стандартов, 1999. – IV, 24 с. – (Система стандартов по информации, библиотечному и издательскому делу).

6. **ГОСТ 7.76-96.** Комплектование фонда документов. Библиографирование. Каталогизация. Термины и определения : межгосударственный стандарт : введён впервые : дата введения 1998-01-01 / разработан Российской государственной библиотекой [и др.]. – Москва : Издательство стандартов, 1997. – III, 53 с. – (Система стандартов по информации, библиотечному и издательскому делу).

7. **Скарук Г. А.** Читательский поиск в электронном каталоге: итоги анализа поисковой статистики ГПНТБ СО РАН / Г. А. Скарук // Науч. и техн. 6-ки. – 2017. – № 12. – С. 63–72.

8. **Руководство** по обеспечению тематического доступа в национальной библиографии / Международная федерация библиотечных ассоциаций и учреждений, Российская национальная библиотека ; [пер. с англ.]. – Санкт-Петербург, 2017 // International Federation of Library Associations and Institutions : официальный сайт. – URL: <https://www.ifla.org/wp-content/uploads/2019/05/assets/hq/publications/series/45-ru.pdf> (дата обращения: 10.08.2021).

9. **Власова С. А.** Реализация тематического поиска в электронном каталоге Библиотеки по естественным наукам Российской академии наук // Румянц. чтения – 2018 : Б-ки и музеи как культур. и науч. центры: истор. ретроспектива и взгляд в будущее: К 190-летию со времени основания Румянц. музея : материалы Междунар. науч.-практ. конф. – Москва, 2018. – Ч. 1. – С. 145–149.

REFERENCES

1. **Breeding M.** Next-Generation Library Catalogs. Chapter 1: Introduction // Library Technology Reports. – 2007. – Vol. 43. – № 4. – P. 5–14. – URL: <https://librarytechnology.org/document/18344> (data obrashcheniya: 10.08.2021).

2. **Yang Sh. Q., Hofmann M. A.** The Next Generation Library Catalog: A Comparative Study of the OPACs of Koha, Evergreen and Voyager // Information Technology and Libraries. – 2010. – September. – P. 141–150. – URL: <https://ejournals.bc.edu/index.php/ital/article/view/3139/2753> (data obrashcheniya: 10.08.2021).

3. **Skaruk G. A.** Komfortnost kak kriteriy otsenki kachestva elektronnoogo kataloga // Tr. GPNTB SO RAN. – Вып. 13. – Т. 2. – Novosibirsk, 2018. – S. 177–184.

4. **GOST 7.73-96.** Poisk i rasprostranenie informatsii. Terminy i opredeleniya : mezhgosudarstvennyy standart : vzamen GOST 7.27-80 : data vvedeniya 1998-01-01 / razrabotan Vserossiyskim institutom nauchnoy i tehlicheskoy informatsii RAN [i dr.]. – Moskva : Izdatelstvo standartov, 1997. – IV, 16 s. – (Sistema standartov po informatsii, bibliotechnomu i izdatelskomu delu).

5. **GOST 7.0-99.** Informatsionno-bibliotechnaya deyatel'nost, bibliografiya. Terminy i opredeleniya : mezhgosudarstvennyy standart : vzamen GOST 7.0-84, GOST 7.26-80 : data vvedeniya 2000-07-01 / razrabotan Rossiyskoy gosudarstvennoy bibliotekoy [i dr.]. – Moskva : Izdatelstvo standartov, 1999. – IV, 24 s. – (Sistema standartov po informatsii, bibliotechnomu i izdatelskomu delu).

6. **GOST 7.76-96.** Komplektovanie fonda dokumentov. Bibliografirovanie. Katalogizatsiya. Terminy i opredeleniya : mezhgosudarstvennyy standart : vveden v pervyye : data vvedeniya 1998-01-01 / razrabotan Rossiyskoy gosudarstvennoy bibliotekoy [i dr.]. – Moskva : Izdatelstvo standartov, 1997. – III, 53 s. – (Sistema standartov po informatsii, bibliotechnomu i izdatelskomu delu).

7. **Skaruk G. A.** Chitatelskiy poisk v elektronnom kataloge: itogi analiza poiskovoy statistiki GPNTB SO RAN / G. A. Skaruk // Nauch. i tehn. b-ki. – 2017. – № 12. – S. 63–72.

8. **Rukovodstvo** po obespecheniyu tematicheskogo dostupa v natsionalnoy bibliografii / Mezhdunarodnaya federatsiya biblioteknykh assotsiatsiy i uchrezhdeniy, Rossiyskaya natsionalnaya biblioteka ; [per. s angl.]. – Sankt-Peterburg, 2017 // International Federation of Library Associations and Institutions : ofitsialnyy sayt. – URL: <https://www.ifla.org/wp-content/uploads/2019/05/assets/hq/publications/series/45-ru.pdf> (data obrashcheniya: 10.08.2021).

9. **Vlasova S. A.** Realizatsiya tematicheskogo poiska v elektronnom kataloge Biblioteki po estestvennym naukam Rossiyskoy akademii nauk // Rumyants. chteniya – 2018 : B-ki i muzei kak kultur. i nauch. tsentry: istor. retrospektiva i vzglyad v budushchee: K 190-letiyu so vremeni osnovaniya Rumyants. muzeya : materialy Mezhdunar. nauch.-prakt. konf. – Moskva, 2018. – Ch. 1. – S. 145–149.

Информация об авторе / Information about the author

Зайцева Екатерина Михайловна – канд. филол. наук, ведущий научный сотрудник ГПНТБ России, Москва, Российская Федерация
katja@gpntb.ru

Ekaterina M. Zaitseva – Cand. Sc. (Philology), Leading Researcher, Russian National Public Library for Science and Technology, Moscow, Russian Federation
katja@gpntb.ru