

ЭЛЕКТРОННАЯ БИБЛИОТЕКА УЧЕБНОГО ЗАВЕДЕНИЯ С ВИЗУАЛЬНЫМИ ИНТЕРФЕЙСАМИ ОТОБРАЖЕНИЯ И УТОЧНЕНИЯ ИНФОРМАЦИИ ПРИ ПОИСКЕ

В.М. Рувинская, В.В. Лотарь

ОНПУ, 65044, Одеса, проспект Шевченка, 1, тел. (0482)68 8543,

iohnl@te.net.ua

Intersog Ukraine Ltd., Одеса, вул. Новосельського, 81, тел. (048)722-45-65

vadim.lot@gmail.com

Предложены способы отображения и уточнения информации при поиске в электронной библиотеке на основе категорий, временных и статистических данных, контекста для выделенных пользователем ключевых слов, ключевых слов, определенных авторами материалов. Выделены различные типы пользователей в зависимости от их осведомленности об искомом ресурсе, построены сценарии поиска для каждого типа. Описана разработанная электронная библиотека методических материалов с использованием усовершенствованного поиска информации.

Methods were proposed for showing and refining the information during searching in the electronic library: based on categories, temporal and statistical data, based on context for user-selected keywords, based on the keywords identified by the authors of materials. Different types of users were defined depending on their awareness of the resources they search. Searching scenarios were constructed for each user's type. The developed electronic library of methodic materials using advanced information search is described.

Постановка задачи

Электронные библиотеки обеспечивают автоматизацию библиотечных процессов: комплектования, каталогизации, книговыдачи и т. п. [1]. Использование электронных библиотек с доступом через Интернет расширяет возможности доступа к книгам, журналам и другим публикациями, хранящимся как в виде твердых публикаций, так и в электронном виде, для различных категорий пользователей, в том числе и для неопытных [2]. Поэтому актуальной становится задача обеспечения поиска нужных публикаций в электронной библиотеке, причем, такого, чтобы в зависимости от категории пользователя система могла приспосабливаться и давать возможность уточнять поисковые запросы по ходу процесса поиска.

Обычно применяют расширенный поиск, используя такие критерии как автор, ключевые слова, год издания или диапазон дат, тип материала. Также объединяют эти критерии с помощью логических связей «и» или «или». Критерии могут быть различными для различных типов материалов. Недостатком такого рода поиска являются большие объемы получаемой информации [3]. Предлагается новый подход для электронных библиотек, позволяющий пользователю уточнять полученные результаты на различных этапах. Этого можно добиться путем внедрения визуальных интерфейсов для уточнения результатов запросов. Внедрение в систему различных визуальных интерфейсов уточнения результатов поиска позволит пользователю удобно и понятно найти необходимые ресурсы и позволит сократить время поиска.

Объект исследования – поиск информации в библиотечной системе. Предметом исследования стали визуальные интерфейсы уточнения запросов при поиске информации в библиотечной системе. Цель исследования – усовершенствование поиска информации в библиотеке за счет уменьшения времени поиска, удобства и простоты работы с системой. Для достижения поставленной цели решаются следующие задачи: изучения работы электронных библиотек вообще и электронных библиотек учебных заведений, а также библиотеки Одесского национального политехнического университета (ОНПУ); исследование способов поиска информации для электронных библиотечных систем, а также разработка электронной библиотеки методических материалов ОНПУ и оценка ее эффективности.

Технологии поиска информации в электронной библиотеке

Простой вариант поиска информации. Материалы, хранящиеся в электронной библиотеке можно отнести к некоторым категориям. Например, электронные методические материалы библиотеки университета относятся к следующим категориям: институты, специальности, направления, учебные дисциплины и тип методического материала (учебник, учебное пособие, конспект лекций, методические указания к лабораторным работам, практическим занятиям, курсовому и дипломному проектированию и т. п.); авторефераты диссертаций – научная специальность, иерархия научных направлений по УДК, тип материала (статья, монография, диссертация, автореферат) и т. д. Причем, множество категорий, как правило, имеет иерархическую структуру и может быть представлено в виде дерева.

В первом, самом простом варианте поиска, предполагается, что пользователи сначала поочередно проходят по спискам каждой из категорий и отмечают в каждом списке необходимый пункт. После этого подсистема поиска выбирает все материалы, относящиеся к выбранным категориям, и показывает на экране их названия, а также требуемую дополнительную информацию, например, автора (авторов), дату издания и т. п. Таким образом, технология поиска материала в электронной библиотеке выглядит так, как показано на рис. 1.

© В.М. Рувинская, В.В. Лотарь, 2010

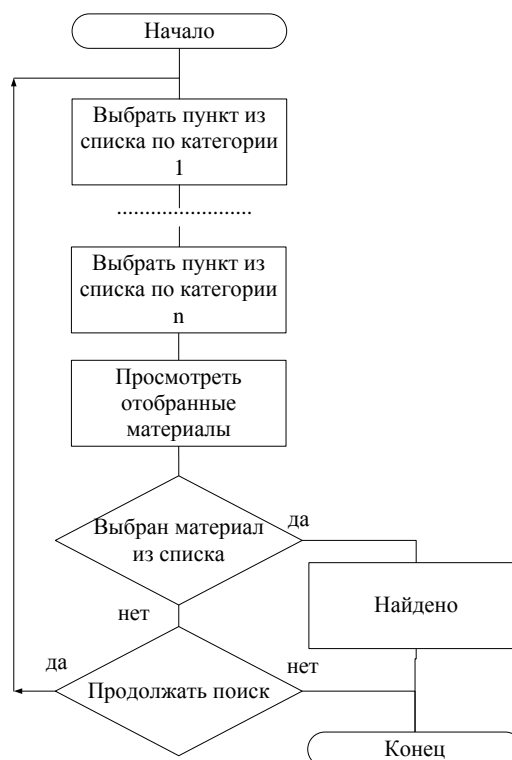


Рис. 1. Технология поиска (вариант 1)

Ниже представлена технология поиска в электронной библиотеке методических материалов.

1. Выбрать институт из справочника институтов.
2. Выбрать специальность из справочника специальностей.
3. Выбрать направление из справочника направлений.
4. Выбрать дисциплину из справочника дисциплин.
5. Выбрать тип методического материала.
6. Просмотреть отобранные методические материалы.
7. Если из списка выбран методический материал по автору, названию, году издания, то перейти к п. 9, иначе – к п. 8.
8. Если необходимо продолжить поиск, то – к п. 1, иначе – к п. 10.
9. Запрос пользователя удовлетворен: найден и / или загружен методический материал.
10. Конец.

Такой вариант достаточно прост для реализации, однако он имеет много недостатков. Главный состоит в том, что пользователь может ничего не знать обо всех или некоторых категориях. И тогда он либо вообще не получит никакого результата, либо получит для просмотра очень большой список материалов, в котором будет трудно разобраться. Кроме того, при этой технологии строго фиксирована последовательность выбора категорий, и таким образом жестко ограничена свобода пользователя.

Усовершенствованный поиск в электронных библиотеках. Предлагается ввести дополнительные возможности поиска информации и соответственно новые визуальные интерфейсы для отображения и уточнения полученной информации.

Рассмотрены визуальные интерфейсы, используемые на сайтах различных поисковых систем, и впервые такого рода интерфейсы были адаптированы для поиска в автоматизированных библиотечных информационных системах.

Особое внимание следует уделить особенностям визуализации полученной после поиска информации, а также гибкому уточнению запросов на поиск как на этапе выбора категорий и критериев поиска, так и на последующих этапах, когда пользователю уже предъявлен для рассмотрения список найденных материалов. Обобщенная технология поиска в электронной библиотеке показана на рис. 2.

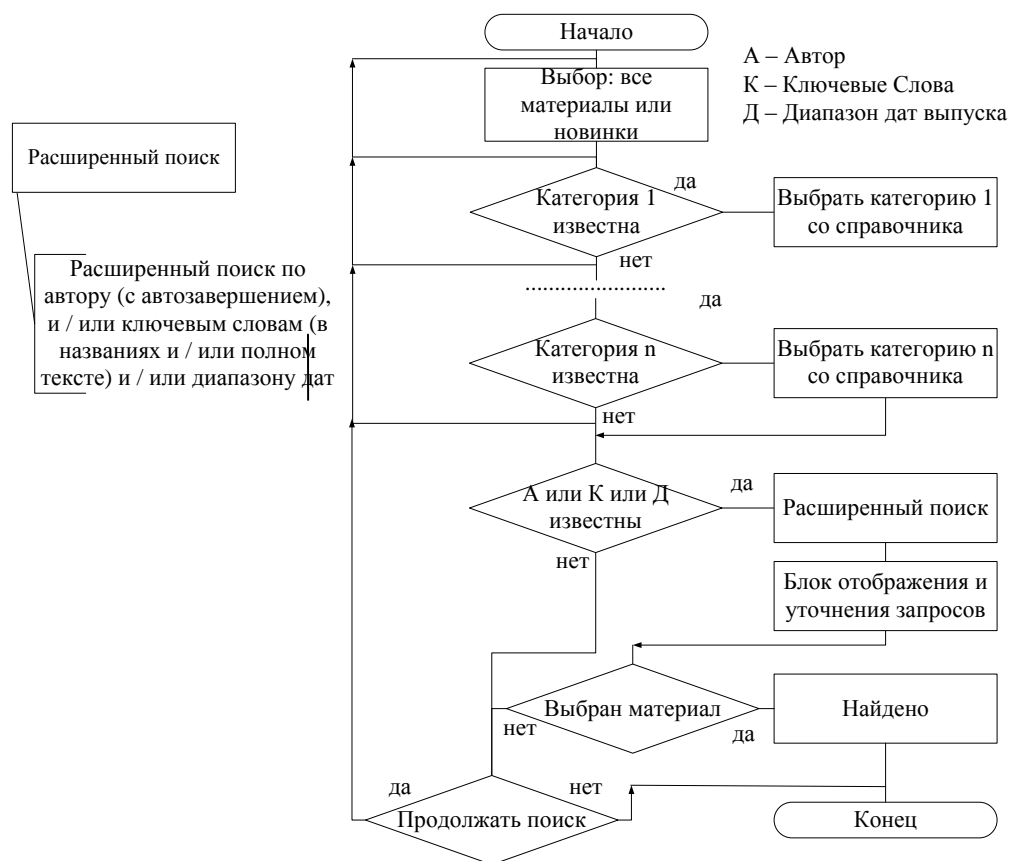


Рис. 2. Технология поиска в электронной библиотеке (вариант II)

Рассмотрим далее предложенные нами способы уточнения запросов на поиск информации в электронных библиотеках.

Уточнение запросов на основе категорий

1. Пользователь может выбирать категории в любой последовательности, либо вообще их не выбирать, а только использовать расширенный поиск; либо наоборот – использовать поиск только по категориям; либо использовать и то, и другое вместе.

2. На этапе, когда уже найдены материалы, он может вернуться и изменить выбор категорий, тем самым уточняя пространство поиска. Возможны два способа отображения:

- пользователь видит на экране найденные материалы и для каждого из них дерево категорий, к которым материал относится, в виде ссылок, с помощью которых он может перейти на справочники конкретных категорий. Таким образом пользователь может уяснить для себя смысл категорий, о которых он ранее не знал (через найденные материалы), посмотреть, что еще полезное для него есть в этой категории;

- пользователь в результате поиска видит дерево, но не полное, а только из тех категорий, в которых присутствуют найденные материалы. Для подсвеченной категории в отдельном фрейме окна высвечиваются названия найденных материалов, авторы, год издания. Аналогично первому способу, пользователь может перейти на список по какой-то категории.

Такой подход для уточнения результатов поиска на основе категорий эффективен в случае, когда пользователь мало осведомлен о категориях, а увидев не только их названия, но и какие материалы к ним относятся, и отобрав нужные категории (множественный выбор), может уменьшить само результирующее дерево. И дальше уже проходить по уменьшенному дереву.

Уточнение запросов на основе временных и статистических данных

1. Учитывать при сортировке результатов ранг популярности, а также дату издания.

2. Как и с твердыми публикациями (есть каталог новых поступлений), можно отдельно показывать только новинки, например за последний год.

Уточнение запросов на основе контекста для введенных пользователем ключевых слов

Пользователь ввел ключевые слова для поиска, а ему представляются фразы и словосочетания из названий и/или из всего текста материалов, содержащие эти ключевые слова. Пользователь выбирает требуемые фразы и словосочетания, что значительно сужает поиск.

Уточнение запросов на основе ключевых слов, определенных авторами материалов

Если речь идет о научных материалах: статьях, авторефератах диссертаций и т. п., то, как правило, требуется, чтобы автор привел ключевые слова, используемые в тексте. Их и предлагается использовать для уточнения запросов при поиске в электронной библиотеке.

1. Расширенный поиск по ключевым полям, введенным пользователем, вводится не только в названиях материалов и/или полном тексте материала, а изначально – в ключевых словах, определенных авторами материалов.

2. Выводятся «облака» тегов, содержащие ключевые слова из материалов. Разные тематики (категории) выделяются различными цветами и шрифтами. Пользователь выбирает часть ключевых слов и получает только те материалы, где эти слова есть. Это можно делать на различных этапах поиска: и для тех категорий, которые пользователь уже выбрал, и после расширенного поиска. Такой поиск может быть реализован как отдельный режим, который пользователь выбирает в нужный момент.

3. «Облака» тегов имеют существенный недостаток, связанный с тем, что эти «облака» расположены по экрану хаотично. Предлагается использовать т.н. «карусельный» интерфейс. При этом близкие по смыслу ключевые слова (в нашем случае они относятся к одной категории), соотносятся с кругом (он носит название диска). То есть на экране мы видим несколько дисков разных цветов, относящихся к тем категориям верхнего уровня, для которых уже найдены материалы. Далее возможны несколько вариантов. На дисках могут быть ключевые слова (если их немного) либо при выборе диска в отдельном фрейме появляются ключевые слова, относящиеся к материалам диска. Можно также перейти от текущего диска к диску категории нижнего уровня. Аналогично «облакам» выбираем интересующие ключевые слова и получаем названия и авторов материалов.

Технология поиска в электронной библиотеке методических материалов.

1. Выбрать режим: просматривать все материалы или новинки.
2. Если известен институт, то переход к п. 3, иначе – к п. 4.
3. Выбрать институт из справочника институтов.
4. Если известна специальность, то перейти к п. 5, иначе – к п. 6.
5. Выбрать специальность из справочника специальностей.
6. Если известно направление, то перейти к п. 7, иначе – к п. 8.
7. Выбрать направление из справочника направлений.
8. Если известна дисциплина, то перейти к п. 9, иначе – к п. 10.
9. Выбрать дисциплину из справочника дисциплин.
10. Если известен тип, то перейти к п. 11, иначе – к п. 12.
11. Выбрать тип методического материала.
12. Если известны автор или ключевые слова или диапазон дат выпуска, то переход к п. 13, иначе – к п. 16.
13. Расширенный поиск по автору (с автозавершением), и / или ключевым словам (в названиях и / или полном тексте) и / или диапазону дат.
14. Просмотр и / или уточнение результатов
 - используется /не используется ранг для сортировки;
 - обычный, древовидный (по категориям или по методикам) способ вывода;
 - используются/не используются для поиска фразы и словосочетания с введенными пользователем ключевыми словами;
 - используются /не используются для поиска ключевые слова, определенные авторами методических материалов.
15. Если выбран из списка методический материал, то перейти к п. 17, иначе – к п. 16.
16. Если необходимо продолжить поиск, то – к п. 1, либо к п. 2, п. 4, п. 6, п. 8, п. 10, п. 12, п. 14, иначе – к п. 18.
17. Запрос пользователя удовлетворен: найден и / или загружен методический материал.
18. Конец.

Типы пользователей. Далее рассмотрим действия для различных типов пользователей. Каждый пользователь обладает некоторыми знаниями о том, что он хочет найти. Отметим также, что пользователь в ходе поиска нужного ему материала может найти для себя полезным и другой ресурс. Было выявлено пять типов пользователей в соответствии с их осведомленностью об искомом ресурсе.

1. Пользователь знает практически все об искомом ресурсе.
2. Пользователь знает (или вспомнит, если увидит) фамилию автора.
3. Пользователь знает приблизительное название материала.
4. Пользователь знает приблизительное название дисциплины, по которой хочет найти материалы.
5. Пользователь практически ничего не знает об искомом ресурсе.

Для первого случая может быть использован как первый вариант поиска (рис. 1), так и второй (рис. 2), так как пользователь, знающий практически все об искомом методическом материале, может легко найти нужный ресурс, используя любой вид поиска. Для остальных типов пользователей приемлем лишь второй вариант поиска.

Для второго случая, когда пользователь приблизительно знает фамилию автора, предложена следующая технология поиска ресурсов.

Шаг 1) пользователь переходит на страницу расширенного поиска.

Шаг 2) если пользователь знает фамилию автора, он вводит фамилию в поле «Автор», если же он не может вспомнить ее, то пользователь набирает любую букву в поле «Автор» и система выдает подсказку (отсортированный список авторов), после чего пользователь выбирает из списка фамилию нужного автора.

Шаг 3) пользователь определяется с видом отображаемых результатов, выбирая один из визуальных интерфейсов.

Шаг 4) если выбран обычный интерфейс отображения результатов, то пользователь просматривает таблицу с полученными результатами и, если материал найден, переходит к шагу 9, иначе – шаг 5.

Шаг 5) если выбран древовидный интерфейс отображения результатов, то пользователь в полученном дереве, кроме названия методического материала, видит к какой специальности и дисциплине относится тот или иной методический материал, что дает ему более полную информацию и в случае успеха переходит к шагу 9, иначе – шаг 6.

Шаг 6) если выбран детализированный интерфейс отображения результатов, то пользователь получает наиболее полную информацию о ресурсах и может выполнить просмотр дерева и прочитает аннотацию к материалу, а в случае успеха перейти к шагу 9, иначе – шаг 7.

Шаг 7) ресурс не был найден либо из-за неверного выбора автора, либо из-за невнимательности; пользователь переходит к шагу 2, если он попал в шаг 7 первый раз, иначе – шаг 8.

Шаг 8) скорее всего материал не был найден из-за невнимательности, и пользователю необходимо выполнить поиск еще раз, выбрав обычный вид отображения результатов и просмотреть каждый материал из списка, а в случае успеха перейти к шагу 9.

Шаг 9) пользователь переходит на страницу с материалами.

Шаг 10) пользователь скачивает выбранный ресурс.

Для третьего случая, когда пользователь приблизительно знает название, предложена следующая технология поиска.

Шаг 1) пользователь переходит на страницу, содержащую справочник материалов, и вводит критерий поиска в поле быстрого поиска, получает список найденных методичек и, если необходимая методичка найдена, переходит к шагу 11, иначе – шаг 2.

Шаг 2) пользователь переходит на страницу расширенного поиска.

Шаг 3) исходя из своих знаний, пользователь формирует критерий поиска.

Шаг 4) пользователь определяется с видом отображаемых результатов, выбирая один из визуальных интерфейсов.

Шаг 5) если выбран обычный интерфейс отображения результатов, то пользователь просматривает таблицу с полученными результатами и, если материал найден, переходит к шагу 10, иначе – шаг 6.

Шаг 6) если выбран древовидный интерфейс отображения результатов, то пользователь в полученном дереве, кроме названия методического материала, видит к какой специальности и дисциплине относится тот или иной методический материал, что дает ему более полную информацию и в случае успеха переходит к шагу 10, иначе – шаг 7.

Шаг 7) если выбран детализированный интерфейс отображения результатов, то пользователь получает наиболее полную информацию о ресурсах и может выполнить просмотр дерева и прочитает аннотацию к материалу, а в случае успеха перейти на шаг 10, иначе – шаг 8.

Шаг 8) ресурс не был найден из-за: либо введены неверные критерии поиска, либо из-за невнимательности; пользователь переходит к шагу 3, если он попал в шаг 8 первый раз, иначе – шаг 9.

Шаг 9) скорее всего материал не был найден из-за невнимательности и пользователю необходимо выполнить поиск еще раз, выбрав обычный вид отображения результатов и просмотреть каждый материал из списка, а в случае успеха перейти к шагу 10.

Шаг 10) пользователь переходит на страницу с материалами.

Шаг 11) пользователь скачивает выбранный ресурс.

Для четвертого случая, когда пользователь знает приблизительное название дисциплины, была предложена следующая технология.

Шаг 1) пользователь переходит на страницу со справочником дисциплин и вводит в поле быстрого поиска нужную дисциплину.

Шаг 2) просматривает список полученных дисциплин и если дисциплина найдена, то шаг 11, иначе – шаг 3.

Шаг 3) пользователь переходит на страницу расширенного поиска.

Шаг 4) исходя из своих знаний, пользователь формирует критерий поиска.

Шаг 5) пользователь определяется с видом отображаемых результатов, выбирая один из визуальных интерфейсов.

Шаг 6) если выбран обычный интерфейс отображения результатов, то пользователь просматривает таблицу с полученными результатами и, если материал найден, переходит к шагу 11, иначе – шаг 7.

Шаг 7) если выбран древовидный интерфейс отображения результатов, то пользователь в полученном дереве, кроме названия методического материала, видит к какой специальности и дисциплине относится тот или иной методический материал, что дает ему более полную информацию и в случае успеха переходит к шагу 11, иначе – шаг 8.

Шаг 8) если выбран детализированный интерфейс отображения результатов, то пользователь получает наиболее полную информацию о ресурсах и может выполнить просмотр дерева, прочитает аннотацию к материалу, а в случае успеха перейти к шагу 11, иначе – шаг 9.

Шаг 9) ресурс не был найден из-за: либо введены неверные критерии поиска, либо из-за невнимательности; пользователь переходит к шагу 4 если он попал в шаг 9 первый раз, иначе – шаг 10.

Шаг 10) скорее всего материал не был найден из-за невнимательности, и пользователю необходимо выполнить поиск еще раз, выбрав обычный вид отображения результатов и просмотреть каждый материал из списка, а в случае успеха перейти к шагу 11.

Шаг 11) пользователь переходит на страницу с материалами.

Шаг 12) пользователь скачивает выбранный ресурс.

В последнем случае, когда пользователь не знает практически ничего о нужном ему ресурсе, необходимо следовать следующим инструкциям.

Шаг 1) пользователь пытается найти свой факультет в справочнике факультетов, если успех, то шаг 2, иначе шаг 6.

Шаг 2) пользователь пытается найти свое направление в справочнике направлений, если успех, то шаг 3, иначе шаг 6.

Шаг 3) пользователь пытается найти свою специальность в справочнике специальностей, если успех, то шаг 4, иначе шаг 6.

Шаг 4) пользователь пытается найти нужную дисциплину в справочнике дисциплин, если успех, то шаг 5, иначе шаг 6.

Шаг 5) пользователь просматривает список методических материалов, доступных для интересующей его дисциплины, если материал найден, то шаг 15, иначе шаг 6.

Шаг 6) пользователь переходит на страницу расширенного поиска.

Шаг 7) из своих и вновь полученных знаний пользователь формирует критерии поиска.

Шаг 8) пользователь определяется с видом отображаемых результатов, выбирая один из визуальных интерфейсов.

Шаг 9) если выбран обычный интерфейс отображения результатов, то пользователь просматривает таблицу с полученными результатами и если материал найден, переходит к шагу 14, иначе – шаг 10.

Шаг 10) если выбран древовидный интерфейс отображения результатов, то пользователь в полученном дереве кроме названия методического материала, видит к какой специальности и дисциплине относится тот или иной методический материал, что дает ему более полную информацию, и в случае успеха переходит к шагу 14, иначе шаг 11.

Шаг 11) если выбран детализированный интерфейс отображения результатов, то пользователь получает наиболее полную информацию о ресурсах и может выполнить просмотр дерева и прочитать аннотацию к материалу, а в случае успеха перейти к шагу 14, иначе – шаг 12.

Шаг 12) ресурс не был найден из-за: либо введены неверные критерии поиска, либо из-за невнимательности; пользователь переходит к шагу 1 если он попал в шаг 12 первый раз, иначе – шаг 13;

Шаг 13) скорее всего материал не был найден из-за невнимательности и пользователю необходимо выполнить поиск еще раз, выбрав обычный вид отображения результатов и просмотреть каждый материал из списка, а в случае успеха перейти к шагу 14.

Шаг 14) пользователь переходит на страницу с материалами.

Шаг 15) пользователь скачивает выбранный ресурс.

«Методические материалы – Online Система»

Предложенные идеи легли в основу электронной библиотеки методических материалов ОНПУ. Система построена на классической трехуровневой архитектуре. В качестве сервера приложений используется Tomcat, в качестве хранилища данных – БД Oracle 10XE. Само приложение построено на архитектуре Model – View – Controller, которую также называют шаблоном проектирования. Первая версия системы была построена на технологии Spring MVC, вторая использует IceFaces, которые по сути являются надстройкой Java Server Faces. Все эти технологии являются Фреймворками, т. е. библиотеками, позволяющими разработчику избежать рутинного и повторяющегося кодирования. В качестве языка программирования выбран Java.

Система состоит из следующих основных компонентов: подсистема взаимодействия с БД посредством Jdbc и Hibernate, подсистем расширенного поиска и загрузки материалов с ftp-ресурса, подсистема аутентификации и другие. Функциональность системы разграничена на несколько пользовательских ролей, среди которых – администратор, студент, преподаватель, библиотекарь и гость.

Есть и небольшие недостатки у разработанных визуальных интерфейсов. К недостаткам древоподобного интерфейса стоит отнести его громоздкость, а к детализированному интерфейсу – скорость работы, хотя в данном случае она не критична.

Пример поиска. Представим, что пользователю системы необходимо найти материал. Все что он о нем знает – это то, что материал является конспектом лекций и относится к менеджменту. Зайдя на страницу расширенного поиска, пользователь вводит эти критерии, выбирает обычный режим отображения результатов, и система производит поиск. Система выдает результат в виде названий большого количества материалов. Пользователю затруднительно разобраться в полученном списке. Он выбирает древовидный способ отображения результатов. После чего ему предлагается дерево с категориями, таким образом система предлагает пользователю уточнить запрос. Пользователь открывает в дереве нужные ему категории и видит список методических материалов, привязанных к открытым категориям. Для уточнения результатов поиска

пользователь выбирает детализированный способ отображения результатов. После чего система открывает для него дерево категорий и полный список методических материалов. Пользователь может просмотреть детальную информацию по материалу, в частности, аннотацию.

Отметим, что каждый вид отображения результатов поиска хорош для разных типов пользователей. Например, пользователь, который знает практически все об искомом ресурсе, воспользуется обычным поиском, так как при таком подходе на экране отображается минимум дополнительной и в этом случае ненужной для него информации. Пользователи, плохо осведомленные о материале, воспользуются детализированным поиском. Оставшиеся категории выберут древовидный поиск.

Выводы

В результате проведенной работы был предложен инновационный подход для отображения и уточнения данных при поиске с использованием визуальных интерфейсов. Разработанные интерфейсы позволяют в разы увеличить скорость поиска, а главное доступ к интересующей информации пользователей электронной библиотеки.

Была спроектирована и разработана автоматизированная библиотечная информационная система «Методические материалы – Online система» с визуальными интерфейсами отображения и уточнения данных при поиске. В настоящее время она работает в библиотеке ОНПУ.

В будущем предполагаются следующие усовершенствования. Во второй версии системы пользователь может сам двигаться по различным окнам системы на основе различных способов уточнения. Предлагается в дальнейшем, чтобы пользователь мог выбирать свой тип (типы приведены выше), и дальше система по сценариям поиска для каждого типа пользователя вела пользователя наиболее оптимальным для него путем. Сами сценарии могут быть не «защиты» в программе, а вводиться и корректироваться по мере работы системы не программистом, а работниками библиотеки.

1. *Коголовский М.Р.* Стандарты XML и электронные библиотеки [Электронный ресурс]: http://www.ict.edu.ru/ft/002337/stand_xml_lib.pdf
2. *Зеленина Г.Н., Гальцева Т.Е.* Электронный каталог в библиотеке // НТБ. – 1995. – № 10–11. – С. 68–72.
3. *Воройский Ф.С.* Разработка средств организационно-технологического обеспечения АБИС // НТБ. – 2001. – № 9. – С. 71–86.