

КАЗАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
им. В.И.УЛЬЯНОВА-ЛЕНИНА
ИЖЕВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
ИНСТИТУТ ПРОБЛЕМ ИНФОРМАТИКИ
АКАДЕМИИ НАУК РЕСПУБЛИКИ ТАТАРСТАН

**Современные информационные технологии
и письменное наследие:
от древних текстов к электронным библиотекам**

El'Manuscript-08

**Материалы Международной научной конференции
(Казань, 26–30 августа 2008 г.)**

Международная научная конференция проводится при финансовой поддержке



Российского фонда фундаментальных исследований (проект № 08-07-06042Г),



компаний ICL-КПО ВС (корпорация Fujitsu Limited),



ООО «Логические системы» (Logic Systems)



Издательство
Казанского государственного университета
2008

К вопросу о возможностях поисковых машин в электронных библиотеках цифровых образовательных ресурсов

Н. М. Ежова

Мурманская академия экономики и управления, Мурманск, Россия

Создание электронных библиотек цифровых образовательных ресурсов (цифровые хранилища) – это одно из наиболее актуальных современных направлений государственной политики в построении виртуального образовательного пространства России, включающего в себя информационные ресурсы всех регионов и обеспечивающего всеобщий доступ к ним.

За 2001–2005 годы в рамках Федеральной целевой программы «Развитие единой образовательной информационной среды» образовательные учреждения оснащались компьютерной техникой и программным обеспечением. Параллельно с этим было начато построение образовательных телекоммуникационных сетей, проведено обучение эффективному использованию информационных технологий нескольких десятков тысяч педагогов, разработаны новые электронные образовательные ресурсы. По этой же программе начал создаваться ряд федеральных порталов для организации свободного бесплатного доступа [Лобачёв 2004] к образовательным ресурсам.

С 2006 года действует приоритетный национальный проект «Образование» (URL: http://www.rost.ru/projects/education/education_main.shtml), в рамках которого в 2006–2007 гг. решалась задача подключения к сети Интернет более 50 тысяч учреждений общего образования¹.

Все порталы и сайты, представляющие информацию об образовательных ресурсах российского Интернет-пространства, являют собой огромные хранилища. В них располагаются десятки тысяч ссылок на эти ресурсы и их электронные копии, хранящиеся непосредственно в библиотеках таких порталов и сайтов.

Приёмы классификации и структурирования находящейся в хранилищах материальных носителей информации (архивы, библиотеки, фильмотеки, медиатеки) разработаны давно и широко известны. Компьютерная среда хранения позволяет реализовать все традиционные приёмы (поиск по алфавиту, по названию книги или по фамилии автора, по категории и т. д.) и предложить новые формы – поиск по фрагменту контекста, по дате занесения в хранилище и проч.

Существенное отличие компьютерных хранилищ информации от материальных состоит в том, что фактическое расположение ресурсов не играет никакой роли.

¹ См., например, http://catalog.iot.ru/index.php?option=com_content&task=view&id=3&Itemid=11.

К примеру, в реальной библиотеке вынуждены повторять информацию о книгах (и других источниках), если хотят предоставить посетителю возможность поиска и по фамилии автора (по алфавиту), и по содержанию (тематический каталог). В виртуальной библиотеке дублировать информацию для её группировки по различным признакам необходимости нет.

Другое преимущество компьютерных (информационных) технологий заключается в возможности ускорить поиск необходимых «читателю» данных в массивах информации.

Однако на данный момент количество организованных образовательных ресурсов стало исчисляться десятками тысяч. К примеру, оцифрованных ресурсов в Библиотеке Единого окна² на 24.02.2008 имелось по самым скромным подсчётам более 17 тысяч, в Библиотеке Единой Коллекции³ – более 50 тысяч.

Найти необходимую информацию сейчас столь сложно, что учителя, в большинстве своём неуверенно владеющие технологией поиска, предпочитают обойтись своими наработками, чем разбираться в той «свалке», что представляет собой «ответ» на запрос пользователя. Ведь нелегко, чтобы человек мог пересмотреть и десятую долю содержимого библиотеки, даже при ограничении области поиска с помощью тематических каталогов или рубрикаторов.

Начинает повторяться знакомая ситуация: учитель не может получить доступ к учебно-методическим материалам, представленным в оцифрованном виде. И если ранее это было связано с ограниченными тиражами соответствующей литературы, с недостаточным освещением методического опыта в периодической печати, то теперь аналогичный эффект вызывает обилие информации подобного рода, к тому же не всегда достоверной. В результате огромные объёмы электронных цифровых ресурсов, созданные коллективами профессиональных разработчиков, могут остаться невостребованными.

Сложившееся положение вызывает сильнейшее беспокойство многих исследователей и, более того, руководителей образовательных Интернет-проектов различного уровня. В связи с этим, на наш взгляд, одной из первоочередных задач для решения данной проблемы должна быть разработка требований к возможностям поиска, предлагаемого в электронных хранилищах.

Первоначально необходимо подробно исследовать, каким именно образом работают уже существующие поисковые машины⁴ на тех или иных порталах или сайтах.

Здесь мы представим некоторые результаты обзора возможностей поиска на отдельных сайтах и порталах российского уровня. Это порталы:

² Информационный портал по русскоязычным образовательным ресурсам «Единое окно доступа к образовательным ресурсам» (ИС «Единое окно») (URL: <http://window.edu.ru>).

³ Проект «Единая коллекция цифровых образовательных ресурсов» (school-collection.edu.ru).

⁴ Программный инструментальный портал (сайта), осуществляющий функции поиска.

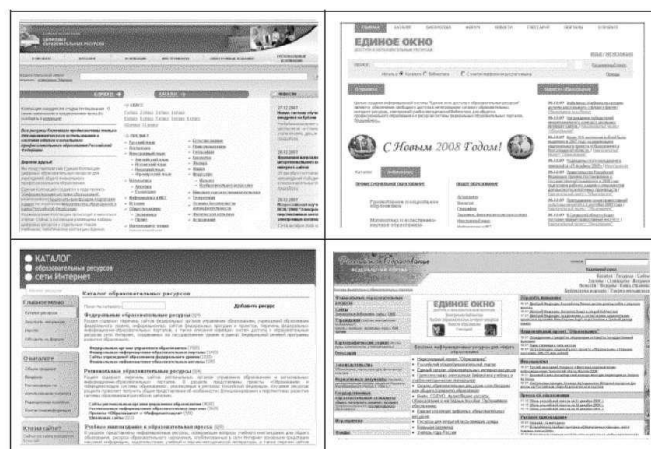


Рис.1

- Единая коллекция цифровых образовательных ресурсов (URL: <http://school-collection.edu.ru>) (рис. 1, вверху слева),
- Единое окно доступа к образовательным ресурсам (URL: <http://window.edu.ru>) (рис.1, вверху справа),
- Каталог образовательных ресурсов сети Интернет (URL: <http://catalog.iot.ru>) (рис.1, внизу слева),
- Российское образование (URL: <http://www.edu.ru>) (рис. 1, внизу справа).

Почти везде предлагается два известных вида поиска: поиск и расширенный поиск, причём, традиционно на «верхних» страницах пользователю доступен только поиск.

Для поиска посетитель портала для нахождения нужной информации вводит соответствующее слово. Параметры такого поиска заданы «по умолчанию», т.е. определяются разработчиком («хозяйном») портала или сайта.

Рассмотрим, какие параметры именно этого вида поиска предлагают разработчики вышеперечисленных порталов.

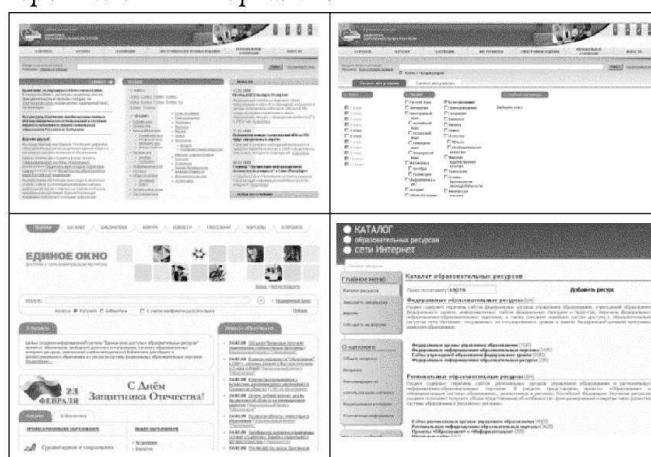


Рис. 2

На сайте «Единая коллекция цифровых образовательных ресурсов» можно задать ограничения области поиска по предмету или по классу, из-за чего приходится «делать лишний шаг»: после выбора необходимого раздела тематического рубрикатора происходит переход на новую «страницу» (рис. 2, вверху).

На сайте «Единое окно доступа к образовательным ресурсам» пользователь может указать, интересуют ли его тексты, содержащие только указанную форму слова, или он ищет материалы «с учётом морфологии русского языка» (рис. 2, внизу слева).

На сайте «Каталог образовательных ресурсов сети Интернет» параметры поиска вообще не предусмотрены (рис. 2, внизу справа).

Из-за того что на каждом из трёх порталов предусмотрены свои «правила игры», учитель (преподаватель), «путешествуя» по сайтам (порталам) сходного назначения, будет вынужден каждый раз перестраиваться при поиске материала.

Отсюда вывод: необходимо предусмотреть единообразие подобных ограничений, к важнейшим из которых мы относим возможности:

(а) уточнений: ищутся или только строгие соответствия запросу, или также различные словоформы указанного слова (словосочетания), к примеру, не только «прогрессия», но и «прогрессии», «прогрессию» и т. п.;

(б) указаний: если задан поиск словосочетания, то ищутся или только документы, содержащие данное словосочетание, или также документы, в которых эти слова встречаются по отдельности (например, не только «безопасность жизнедеятельности», но и «безопасность» и «жизнедеятельность»).

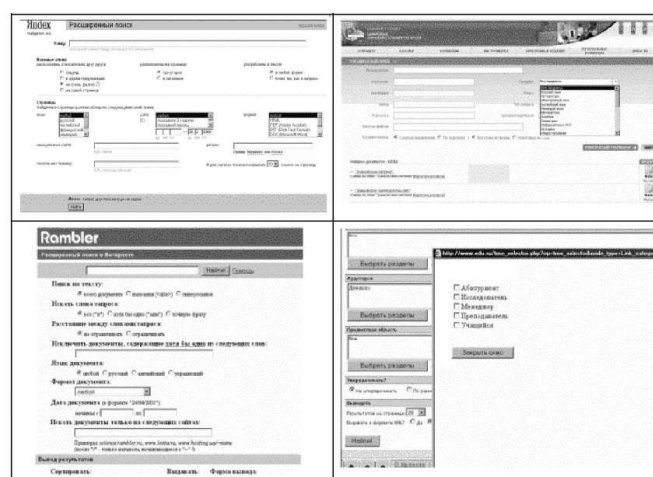


Рис. 3

Кроме того, если учесть, что «посетитель» хранилищ может искать конкретный учебник (ресурс), необходимо дать ему возможность вводить, как особо важные реквизиты, фамилию автора и название ресурса.

Если же говорить о расширенном поиске (опция, позволяющая задать большой спектр дополнительных реквизитов поиска), то различия между порталами в наборах предлагаемых реквизитов ещё значительнее.

На рис. 3 приведено сравнение содержимого расширенного поиска на серверах, предназначенных для свободного поиска в Интернете (на рисунке – слева), и на образовательных сайтах и порталах (на рисунке – справа).

С первого же взгляда видно, что «страницы» расширенного поиска на сайтах сходного образовательного содержания отличаются между собой существенно, чем такие же «страницы» на серверах, позволяющих искать произвольную информацию во всём мировом пространстве.

В результате от пользователя при работе с расширенным поиском на этих порталах потребуются дополнительные усилия для понимания специфики работы конкретной поисковой системы.

Отсюда вывод: необходимо предусмотреть единообразие в составе наборов реквизитов для расширенного поиска.

Подведём краткие итоги.

Первый шаг уже сделан: для работы учителя (преподавателя) создано и активно развивается пространство, наполненное только образовательным содержанием.

Вторым важным шагом должна стать унификация решений однотипных задач для работы на разных сайтах (порталах) образовательного содержания, хотя бы федерального уровня (находящихся под единым руководством), необходимая для удобства работы пользователя, должны быть унифицированы параметры поиска материала и содержание карточек описания ресурса (на рис. 4 приведены карточки ресурсов в разных хранилищах).

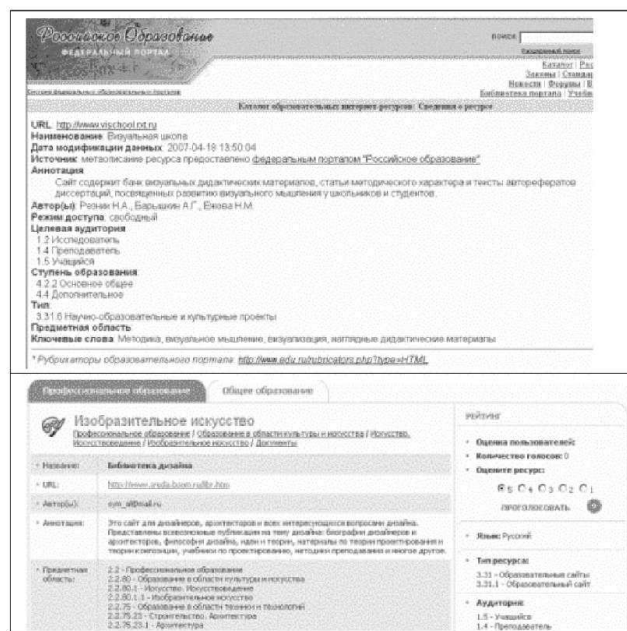


Рис. 4

Сегодня преподаватель (учитель) в поисках нужной информации вынужден последовательно просматривать имеющиеся образовательные сайты и порталы

(пусть и укрупненные до федерального масштаба). Таким образом, на третьем шаге, на наш взгляд, желательно объединение подобных ресурсов и представление результатов поиска в них (одновременно) таким образом, чтобы пользователь мог при первом же взгляде, не заходя на каждый из них, понять, подходят или нет найденные ресурсы для решения его задач.

В поддержку нашего мнения приведём слова Д. Уэльса, основателя Википедии⁵: «Одна из основных задач поисковика - принять решение: этот сайт хорош, а этот - плох. Компьютеры не могут принимать подобные решения. Но мы сами можем прекрасно это делать: нам достаточно просто взглянуть на сайт» [Губайловский 2006].

В заключение подчеркнём, что проблема поиска важна для каждого работающего с электронными хранилищами. Она должна быть решена настолько просто и удобно, чтобы любой учитель, даже слабо подготовленный как пользователь Интернета, мог бы с лёгкостью находить необходимые ему ресурсы и хотел бы пользоваться электронными хранилищами, как самым удобным источником учебно-методических материалов.

Список литературы

- Губайловский 2006 – Губайловский, В. Основатель Википедии разрабатывает новую поисковую систему [Электронный ресурс] / В. Губайловский. – 2006. – Режим доступа : <http://katolik.ru/modules.php?name=News&file=article&sid=16543>, свободный. – Загл. с титул. страницы.
- Лобачёв 2004 – Лобачёв, С. Л. К вопросу о будущем образовательных порталов [Электронный ресурс] / С. Л. Лобачёв // Единая образовательная информационная среда: проблемы и пути развития : материалы Третьей Всерос. научно-практ. конф.-выставки. – Омск, 2004. – Режим доступа : <http://www.ict.edu.ru/vconf/files/6865.doc>, свободный. – Загл. с титул. страницы.

Towards consideration of the possibilities for a search engine for digital educational resources

Nataliya M. Ezhova

Murmansk Academy of Economics and Government, Murmansk, Russia

While a physical library may need to have information repeated for various access points, a digital library does not need such redundancy. However, it is difficult to search across digital libraries, whose internal structure is fluid. The solution of which should be simple and convenient; otherwise, thousands of professionally created digital resources may be left unfound. This paper argues for documenting the fundamental requirements of searching in digital depositories.

⁵ Интернет-энциклопедия, пополняемая читателями.