

Повесть о двух библиотеках,
которые обкрадывают своих пользователей
или
Суровая необходимость настройки
сервисов доступа к электронным ресурсам

Наталья Николаевна Литвинова
Российская государственная библиотека

1-ая международная конференция НЭИКОН о российских и
иностраных электронных ресурсах, публикациях и журналах
Будва, 29 сентября – 4 октября 2013 г.



Необходимость сервисов доступа к множеству разнородных ресурсов библиотеки из одной точки

До конца 2000-х годов в библиотеках, как правило, поиск в библиотечных ресурсах можно было вести отдельно в:

- электронном каталоге, отражающем, в основном, печатный фонд;
- электронной библиотеке (репозитарии), включающей электронные документы, размещенные на серверах библиотеки;
- сетевых удаленных ресурсах (СУР), в основном лицензионных.

Для поиска во всех ресурсах из одной точки использовалась технология распределенного (федеративного) поиска.

Сервисы типа discovery (web-scale discovery)

Принцип работы:

- создание единого индекса, формируемого индексированием всех типов ресурсов, представляющих интерес для пользователей библиотеки;
- поиск через этот индекс (из одной точки)
- ссылки от результатов поиска, представленных библиографическими записями, к полным текстам документов, доступных данной библиотеке.

Сервисы Discovery:

Summon® - Serials Solutions, Proquest

EBSCO Discovery Service™ (EDS) - EBSCO

Primo® - ExLibris

WorldCat® Local - OCLC

Marshall Breeding:

“by 2016 we can expect discovery products to reach a state of full maturity and near universal deployment among academic libraries”

Сервисы Discovery в России

В настоящее время российские библиотеки внедрили в основном сервис EDS.

EDS внедрен в 16 библиотеках, 14 из которых – университетские и две – публичные.

Часть библиотек внедрила EDS для поиска во всех своих ресурсах, часть – для поиска только в сетевых удаленных ресурсах.

Последний вариант внедрения функционально близок к использованию обширных реферативно-библиографических баз данных с внедренными ссылками на доступные полные тексты (SciFinder, Inspec, Medline ...)

Как обеспечиваются ссылки именно на доступные документы (appropriate copies)?

Основной механизм внедрения ссылок – сервис Link Resolver

(определитель ссылок) – сервис, который индивидуально создает ссылки на полные тексты документов. Работа определителя ссылок заключается в сопоставлении ссылок к ресурсу-источнику с полными текстами этих документов, в которых они опубликованы.

Ресурс-источник – это документ, в котором опубликованы полные тексты этих документов. Библиографически сопоставлены ссылки к ресурсу-источнику с полными текстами этих документов, в которых они опубликованы. OpenURL содержит:

http://linksource.ebsco.com/linking.aspx?sid=ukoln:ariadne&genre=article &title=Information%20gateways:%20collaboration%20on%20content &title=Online%20Information%20Review&issn=1468-4527&volume=24 &spage=40&epage=45&artnum=1&aulast=Heery&aufirst=Rachel

- base URL - адрес определителя ссылок, применяемого библиотекой;
- идентификатор ресурса-источника;
- цепочку метаданных библиографической записи.

Как формируются ссылки?

Определитель ссылок извлекает цепочку метаданных и информацию о том, какой ресурс-источник ее сформировал. Его задача – через индекс найти описанный документ по *обогащенным* метаданным и определить, есть ли у библиотеки доступ к нему.

Если доступ есть, определитель ссылок формирует ссылки, одну или несколько, на полные тексты документов на разных платформах издателей и/или агрегаторов.

Если доступа к документу у библиотеки нет, то в зависимости от настроек, примененных библиотекой, определитель ссылок либо вообще не формирует ссылок, либо формирует их на определенные библиотекой сервисы, которые могут помочь в получении документа: МБА или поисковые сервисы Интернета.

Как определитель ссылок понимает, есть ли у библиотеки доступ к документу?

Для этого библиотека должна настроить **общую базу знаний**, содержащую сведения о:

- всех доступных коллекциях (пакетов) документов на разных платформах;
- всех документах (журналах, книгах, ...) входящих в эти коллекции;
- периодах доступности этих документов (особенно актуально для периодики).

В результате настройки формируется **локальная база знаний**.

Настройка базы знаний – важнейшее звено в обеспечении адекватной работы определителя ссылок!!!

Иллюстрация: настройка базы знаний в сервисе A to Z

Выбор провайдеров

The screenshot displays the EBSCO A to Z Administrator interface. The main window is titled 'Collection Editor - Windows Internet Explorer' and shows the configuration for a collection named 'Energy economics'. The interface includes a navigation pane on the left with 'Title List Wizard' and 'Administrator' sections. The main content area shows details for the 'Energy economics' collection, including ISSN numbers, managed coverage (1979 to present), and publisher (Elsevier). There are sections for selecting vendors, packages, and custom links. A 'Notes' section at the bottom contains specific access instructions.

Выбор коллекций (пакетов, баз данных)

Задание периодов доступа (для коллекции в целом или отдельных документов)

Energy economics

Elsevier Science : ScienceDirect Subject Collections - Economics, Econometrics, and Finance
Print ISSN: 0140-9883 Managed Coverage: 1979 to present Subject: Social Sciences -- Industries. Land use. Labor -- Special industries and trades
Online ISSN: 1873-6181 Publisher: Elsevier

Proxy Server: Save Cancel

Custom Coverage Begin: End:

Custom Embargo Days:

Do not display this resource on the End User Site.

LinkSource CustomLinks:

	Label	Category
<input type="checkbox"/>	American Association for the Advancement of Science (AAAS)	Full Text
<input type="checkbox"/>	American Chemical Society Web Edition (Issue)	Full Text
<input type="checkbox"/>	American Institute of Physics (AIP) (Article)	Full Text
<input type="checkbox"/>	American Mathematical Society - MathSciNet (Article)	Full Text
<input type="checkbox"/>	American Physical Society - Meetings	Full Text

Notes:

	Label	Note
<input type="checkbox"/>	East View	(Журналы на русском языке)
<input type="checkbox"/>	AMS	AMS journal
<input type="checkbox"/>	Science	Your Access: 1997 -
<input type="checkbox"/>	Log-Pass	(Доступ по логину и паролю - вводятся сотрудниками РГБ)

Каковы следствия неудовлетворительной настройки базы знаний?

Недостатки в настройке базы знаний ведут к появлению ошибочных ссылок:

- Ложные положительные (false positive): документ реально недоступен, но ссылка на него сформирована

Это порождает **разочарование пользователей в сервисе**

- Ложные отрицательные (false negative): документ реально доступен, но ссылка на него не формируется.

Это и есть **КРАЖА у ПОЛЬЗОВАТЕЛЕЙ**

Библиотека X: поиск статей из журнала *Current Sociology* (издается с 1952 г.) издательства Sage

The screenshot shows a web browser window displaying search results on the Ebscohost platform. The URL in the address bar is <http://ehis.ebscohost.com/eds/resultsadvanced?sid=325c2c4b-0ffb>. The page features a search results summary on the left and a list of search results on the right.

Search Results Summary (Left Panel):

- Уточнить критерии поиска
- 1,617 Результаты для
- Найти все мои термины: SO "Current Sociology"
- Расширители: Искать в полных текстах
- Уточнить результаты:
 - Полный текст (онлайн)
 - Научные (рецензируемые) журналы
- Дата издания: 1952 - 2013
- Виды документов

Search Results List (Right Panel):

- Страница: 1 2 3 4 5 Далее >
- Первая дата ▾ Параметры страницы ▾
- 721 Результаты для...
- Найти все мои термины: SO "Current Sociology"
- Расширители: Искать в полных текстах
- Ограничители:
 - Полный текст (онлайн)
 - Научные (рецензируемые) журналы
- Дата издания: 19520101-20131231
- Уточнить результаты:
 - Полный текст (онлайн)
 - Научные (рецензируемые) журналы
- Дата издания: 1999 - 2013

Search Results (Main Content Area):

- журнал
- Ссылка на полный текст
- 2. [Which way forward for **sociology** of work? An introduction.](#)
By: Castillo, Juan Jose. *Current Sociology*, Apr99, Vol. 47 Issue 2, p1, 4p
База данных: Publisher Provided Full Text Searching File
- Ссылка на полный текст
- 3. [The emergence, centrality and end of work.](#)
By: Freyssenet, Michel. *Current Sociology*, Apr99, Vol. 47 Issue 2, p5, 16p
База данных: Publisher Provided Full Text Searching File
- Ссылка на полный текст

Annotation:

Ограничение по наличию полного текста отсеивает все записи из архивных выпусков журнала, поскольку библиотека не прописала в БЗ архив Sage

Библиотека У - поиск статей из журнала Current Sociology издательства Sage

The screenshot shows a search results page for 'Current Sociology' on the EBSCOhost platform. The search criteria are 'SO current sociology'. The results are filtered to show only 'Full text (online)' documents. The search results list several articles, including 'PREFACE', 'FIREWORD', and 'HISTORY AND ORGANIZATION OF SOCIOLOGICAL STUDIES HISTOIRE ET ORGANISATION DES ETUDES SOCIOLOGIQUES'. A text box highlights the 'Ограничители' (Restrictions) section, which is set to 'Full text (online)'. The text box contains the following text:

Ограничение по наличию полного текста отсеивает записи текущих выпусков журнала из-за того, что библиотека не прописала в БЗ текущие выпуски Sage

В коллекции Sage более 400 журналов, значит, ложные ссылки порождены для десятков и сотен тысяч статей из них.

И это отнюдь не единичные примеры!

Особенно «не везет»:

- архивам на платформах издателей (архивы на российской платформе не прописаны вообще ни у одной библиотеки)
- коллекциям журналов открытого доступа

Упрощенное описание цепочки взаимодействия, поддерживающей работу Link Resolver (LR)

Кто генерирует	Что генерирует	Кто использует
Издатели	Поддержка: отраслевые Поддержка: проект KBART (UK	Метаданные
Издатели	SG-NISO) – Knowledge Bases and Related Tools Рынок	Информация о коллекциях и их в них документах
Общие средства поиска контента: A&I базы данных Resolvers Поисковые сервисы	Поддержка: проект IOTA (NISO) - Improving OpenURLs through Analytics Рынок	Обобщенные базы знаний владельцев Link Resolvers
Владельцы Link Resolvers	Поддержка: проект KBART (UKSG-NISO) – Knowledge Bases and Related Tools Рынок	Обобщенные базы знаний
Библиотеки	Рынок	Обобщенная база знаний (настроенная под ресурсы библиотеки)
		Общие средства поиска контента: A&I базы данных Поисковые сервисы
		Link Resolvers

Потери в цепочке генерации-поставки данных для формирования правильных ссылок

В библиотеках выполнен ряд работ, оценивающих эффективность работы Link Resolvers.

Данные одной из них (Digging into the Data - LTR, October 2010)

Отобрано SourceURL (OpenURL) – 351

Link Resolver, отработав, дал следующие результаты:

- положительные результаты (есть доступный полный текст) – **169 – 343 ссылки**
- отрицательные результаты (нет доступного полного текста) – **182**

Библиотека проверила реальную доступность:

- из 343 положительных ссылок вывели на доступные документы - **292 (85%)**. Из них **59 не предоставляли полный текст на открывшейся странице, требуя дальнейших поисков (тоже ошибка!)**
- из 182 отрицательных ссылок **44** оказались ложными: документы были доступны библиотеке. Из остальных **138 – 21 (15%)** были найдены в свободном доступе через Google (Scholar).

Итог: из 525 (343+182) результатов всего правильных – 430 (82%)

Причины ошибочных результатов (распределились примерно поровну):

- недостаточность метаданных и/или ошибки в метаданных источника;
- ошибки «перевода» LR из SourceURL в TargetURL
- неспособность контент-провайдера (цели) выполнить правильное разрешение TargetURL
- ошибки в базе знаний

Выводы об эффективности работы LR

1. При отличной настройке библиотекой базы знаний процент ошибочных результатов работы Link Resolver – не менее 15%
2. При плохой настройке базы знаний процент ошибочных результатов может превысить (и превышает!) 50%

Вопрос (риторический)

Чего стоит сервис, дающий ошибочные результаты более чем в половине случаев?

Библиотеки должны концентрировать усилия на настройке БЗ и анализе сбоев в работе LR

Две цитаты из статьи: Xiaotian Chen. Broken-Link Reports from SFX Users: How Publishers, Vendors and Libraries Can Do Better. - Serials Review 38 (2012) 222–227

It is always a sign of bad library service when users are the first to find out that links are not working.

Всякий раз когда пользователь первым обнаруживает неверную ссылку, это свидетельствует о плохой работе библиотеки .

Dead links are frustrating, but a quick email reply from a librarian with a PDF full-text attachment will certainly ease frustration, especially when the users are working under the pressure of approaching deadlines.

Неверные ссылки разочаровывают, но быстрая реакция библиотекаря по электронной почте с вложенным pdf-файлом смягчит ситуацию, особенно если пользователь стремится срочно завершить работу к сроку.

(Имеется в виду, что в меню LR есть опция связи с сотрудником библиотеки, который быстро реагирует на проблему)

Почему российские библиотеки не уделяют должного внимания настройкам БЗ и LR?

1. Недооценка серьезности проблемы (*«Мы понимаем, что в некоторых случаях будут ошибки – ну ничего, переживем»*)
2. Преувеличение оценочных объемов работ, необходимых для более точной настройки (*«Это что ж теперь, из-за этих ошибок нам специальных сотрудников выделять на эту работу?»*)

Для эффективного выполнения этих работ необходимо выделить типовые проблемы и операции.

1. Отсутствие квалифицированных кадров

Типовые проблемы настройки (1)

1. Проблема выбора подходящей коллекции из обобщенной базы знаний.

Пример: прописывание архивной коллекции Sage в A to Z.

Предлагаются следующие архивные коллекции:

Sage Deep Backfile Package 2007 (Titles: 292)

Sage Deep Backfile Package 2008 (Titles: 349)

Sage Deep Backfile Package 2009 (Titles: 375)

Sage Deep Backfile Package 2010 (Titles: 398)

Sage Deep Backfile Package 2011 (Titles: 408)

Sage Deep Backfile Package 2012 (Titles: 425)

Sage Deep Backfile Package 2013 (Titles: 445)

Самая близкая по сроку открытия доступа и по количеству журналов – коллекция 2011-го года. На сайте НЭИКОН опубликован список доступных журналов архива Sage. Пробуем формально сравнить два списка.

Результат формального сравнения списков журналов в A to Z и на сайте НЭИКОН (по названиям)

Sage Title Lists - Commented - Microsoft Excel

Главная Вставка Разметка страницы Формулы Данные Рецензирование Вид Разработчик Настройки

Обычный Разметка страницы Представления Во весь экран

Линейка Сетка Панель сообщений

Строка формул Заголовки

Масштаб 100%

Новое окно Разделить Упорядочить все Закрепить области

Скрыть Отобразить

Сохранить рабочую область

Перейти в другое окно

Макросы

А2 Accounting History

	A	B	C	D	E	F	G
	Названия журналов на сайте НЭИКОН (при описании доступа к архиву)	Названия журналов из списка A to Z - Sage Deep Backfile Package 2011	Номер журнала в списке A to Z, совпадающий с названием из списка на сайте	Номер журнала в списке A to Z, отличающийся от названия из списка на сайте несущественными деталями			
1							
2	Accounting History	Accounting history	1				
3	Acta Sociologica	Acta Sociologica	2				
4	Adaptive Behavior	Adaptive behavior	3				
5	Administration & Society	Administration & society	4				
6	Adult Education Quarterly	Adult Education Quarterly	5				
7	Aesthetic Surgery Journal	Aesthetic Surgery Journal	6				
8	Affilia	Affilia	7				
9	American Behavioral Scientist	American Behavioral Scientist	8				
10	American Educational Research Journal	American educational research journal	9				
11	American Journal of Alzheimer's Disease & Other Dementias	American Journal of Alzheimer's Disease and Other Dementias	#Н/Д	10			
12	American Journal of Evaluation	American Journal of Evaluation	11				
13	American Journal of Hospice and Palliative Medicine	American Journal of Hospice & Palliative Medicine	#Н/Д	12			
14	American Journal of Medical Quality	American journal of medical quality	13				
15	American Journal of Sports Medicine, The	American Journal of Sports Medicine	#Н/Д	14			
16	American Politics Research	American Politics Research	15				
17	American Review of Public Administration, The	American Review of Public Administration	#Н/Д	16			
18	Angiology	Angiology (Sage)	#Н/Д	17			
19	ANNALS of the American Academy of Political and Social Science, The	Annals of the American Academy of Political and Social Science	#Н/Д	18			
20	Applied Psychological Measurement	Applied Psychological Measurement	19				
21	Armed Forces & Society	Armed Forces & Society	20				
22	Asia Pacific Journal of Human Resources	Asian Cardiovascular & Thoracic Annals	#Н/Д				
23	Asian Cardiovascular & Thoracic Annals	Asia-Pacific journal of public health	21				
24	Asia-Pacific Journal of Public Health	Assessment	22				
25	Assessment	Australian and New Zealand Journal of Criminology	23				

Готово

Соответствия Сайт - AtoZ Соответствия Рос_архив - AtoZ Лист3

Количество: 72 100%

Основные причины несовпадения названий

1. «Мелочи»: использование артикля на разных позициях, варианты написания союза (“and” “&”), использование аббревиатур, использование «распространений» к основному названию. *Легко обнаруживаются, если в списках есть ISSN.*
2. ПЕРЕИМЕНОВАНИЯ журналов
Это самая серьезная проблема, требующая значительных трудозатрат для ее решения.
Есть рекомендация КВАРТ: размещать все тома и выпуски журнала под тем названием, под которым они были опубликованы, а переименования учитывать при поиске.
Однако в электронной среде есть тенденция собирать все тома и выпуски под современным названием.
3. Переход журнала от одного издательства к другому.

Пример: журнал Sage “*Asia Pacific Journal of Human Resources*”

Данные из Ulrich’s Periodical Directory:

Former titles (**until 1991**): **Asia Pacific Human Resource Management** (Australia) (1032-3627) (**until 1988**): Human Resource Management Australia (Australia) (0156-904X) (**until 1978**): Personnel Management (United States) (0048-346X) (**until 1955**): Journal of the Institute of Personnel Management (Australia)

На сайте Sage:

Не только все выпуски собраны под современным названием, но и в библиографических записях фигурирует ТОЛЬКО СОВРЕМЕННОЕ НАЗВАНИЕ

Пример: Ron Cacioppe, Pat Warren-Langford, and Libby Bell **Trends in Human Resource Development and Training**, *Asia Pacific Journal of Human Resources*, **May 1990** 28: 55-72.

По данным Google Scholar эта статья цитируется 12 раз

Для сравнения: *ссылка на эту статью* в списке литературы статьи Garavan T. N., Costine P., Heraty N. The emergence of strategic human resource development //Journal of European Industrial Training. – 1995. – Т. 19. – №. 10. – С. 4-10.

Cacioppe, R., Warren-Longford, P., Bell L. (1990), "Trends in human resource development and training", *Asia Pacific Human Resource Management*, **May 1990**, pp. 55-67.

Этот журнал перешел на Wiley. Архив на Sage также доступен. Его придется прописать индивидуально.

Типовые проблемы настройки (2)

2. Выбор режима поддержки выбранной коллекции (пополняемый – непополняемый)

Пример:

Архив журналов Oxford University Press.

A to Z предлагает только один вариант, не привязанный к «чужим» консорциумам: Oxford Journals Archive (Titles: 218).

Это *пополняемая коллекция*.

Лицензия НЭИКОН содержит 169 названий. Только эти названия можно прописать в своей БЗ и *запретить A to Z пополнять эту коллекцию*.

Типовые проблемы настройки (3)

3. Задание периодов доступа.

Пример: для коллекций журналов Эльзевир в обобщенной БЗ дается полный период существования журнала.

Для каждой библиотеки реальный период доступа зависит от того, когда она начала подписываться. Этот период и нужно прописать.

Что нужно для облегчения решения этих проблем?

1. Требовать, чтобы в лицензии на ресурс состав ресурса был раскрыт в форме, наиболее удобной для последующего сопоставления с коллекциями обобщенной БЗ.

Модельная лицензия консорциума NESLI требует предоставления данных согласно требованиям KBART, то есть полного перечня в Excel с указанием ISSN, начальной и конечной дат в определенном формате.

2. Требовать, чтобы в лицензии было прописано, будут ли доступны библиотеке пополнения коллекции.

3. При организации коллективного (консорциумного) доступа требовать, чтобы полная информация о доступных изданиях была доведена до сведения библиотек.

4. Обмениваться информацией с коллегами из других библиотек.

Проблема кадров: какие сотрудники нужны?

1. Аккуратный исполнительный сотрудник (АИС), способный скрупулезно работать со списками (*неполный рабочий день*)



2. Творческий эрудированный сотрудник (ТЭС), способный понимать и обобщать проблемы настройки и формулировать задачи для АИС (*неполный рабочий день*)



Как обеспечить участие ТЭС (варианты)?

1. Выделить определенное время **штатного сотрудника библиотеки** для того, чтобы он сформулировал проблемы и пути их решения и обозначил их для себя и для АИС.
2. Использовать *сотрудника другой библиотеки*, более опытного в этой области, для консультаций и ОБУЧЕНИЯ штатных сотрудников.
3. Использовать *сотрудника другой библиотеки* для первичной настройки сервиса и ОБУЧЕНИЯ штатных сотрудников.

Решение проблем откладывать нельзя!



Благодарю за внимание!

Вопросы?

Наталия Николаевна Литвинова

nlit@rsl.ru