

Сергей Константинович Канн

кандидат исторических наук, старший научный сотрудник отделения Государственной публичной научно-технической библиотеки Сибирского отделения Российской академии наук (Новосибирск)

Модели развития библиотечных сайтов в свете вебометрики

В данной статье рассматриваются четыре коммуникативные модели развития библиотечных сайтов, описывающие способы их продвижения и взаимодействия с внешней средой. Проведенная типологизация моделей позволяет разработать индивидуальные планы действий и необходимые рекомендации по наращиванию и продвижению веб-ресурсов, а также по повышению эффективности библиотечного обслуживания в интернете.

Ключевые слова: библиотечный сайт, эффективные модели развития, продвижение, веб-ресурсы, индикаторы развития, вебометрика, веб-аналитика, библиотеки Сибири и Дальнего Востока.

This article discusses four communicative models of library sites development, describing the ways of their promotion and interaction with the external environment. The typology of these models allows us to develop individual action plans and necessary recommendations for building up and promoting web resources, as well as for improving the efficiency of library services in the Internet.

Keywords: library website, effective models of development, promotion, web resources, development indicators, webometrics, web analytics, libraries of Siberia and the Russian Far East.

Среди разных атрибутов, описывающих специфику сайтов, особый интерес представляют способы их взаимодействия с внешней средой – своего рода коммуникативные модели поведения, объективно складывающиеся в процессе эволюции. Очевидно, что в настоящее время, в силу объективно-исторических причин, еще преждевременно говорить о каких-то осмысленных «стратегиях» развития библиотечного интернета и о достижении каких-либо заранее намеченных целей.

Прежде всего стоит отметить, что только единицы сайтов декларируют понятную и ясно сформулированную «политику» развития. Очевидно, что ее реализация – это вопрос особый, и в рамках данной статьи мы его разбирать не будем. Но помимо этого есть еще одно обстоятельство, препятствующее выработке долговременной стратегии, – быстрая эволюция самой сети. Вместе с ней меняются инструменты, технологии, условия привычного бытования сайтов. В результате любые попытки постулировать некую устойчивую «стратегию» развития библиотечного веба устаревают так же быстро, как и появляются. Рациональный взгляд на проблему подсказывает, что в данный момент можно стремиться лишь к тому, чтобы выработать более эффективные подходы к развитию интернета в библиотеках.

В данном случае большую практическую пользу могла бы принести типологизация существующих моделей библиотечных сайтов. Рассматривая такое моделирование как основу для создания планов в условиях неопределенности, можно корректировать цели сетевого развития, повышать эффективность его отдачи. При этом для изучения количественных и качественных аспектов управления веб-ресурсами можно привлекать весь арсенал имеющихся и новых средств вебометрики, ориентированной на технологии будущего.

С учетом названных обстоятельств, начиная с осени 2015 г. мы проводили разностороннее исследование библиотечных сайтов с целью собрать различные метрики их развития (открытые данные о размере и динамике обновления, источниках и структуре трафика, аудитории, посещаемости, времени и глубине просмотра страниц и пр.). Затем, после выявления некоторых закономерностей развития, можно было создать примерную типологию библиотечных сайтов, основываясь на определенных критериях и схожести характеристик. Для сбора данных использовались бесплатные инструменты SEO-анализа и вебометрики, позволяющие по единой схеме провести массовую оценку индикаторов сразу у многих сайтов.

В частности, привлекались онлайн-инструменты *XTool* (*xtool.ru*), позволяющие проводить регулярный мониторинг и SEO-аудит сайтов. Средства XTool взаимодействуют с инструментами «Яндекс» и получают от него информацию об индексации ресурсов (визуализируется на графиках), о скорости обновления страниц, показателях тИЦ (тематического индекса цитирования) и других метриках (табл. 1). XTool может извлекать и воспроизводить сведения о ключевых словах, приводящих пользователей на сайты.

Таблица 1

**Веб-метрики сайтов центральных научных библиотек
по данным XTool и «Яндекса» на 25.01.2018 г.**

Библиотека	URL	Возраст сайта (дней / лет)	Страниц на сайте («Яндекс»)	Скорость обновления (страниц в сутки)	ТИЦ
РГБ	rsl.ru	7414 / 20,3	53 228	0,8	0
ГПНТБ России	gpntb.ru	7479 / 20,5	35 688	0,04	4100
РНБ	nlr.ru	7422 / 20,3	7 814 669	0,01	5300
БЕН РАН	benran.ru	5686 / 15,6	24 299	0,59	2600
ГПНТБ СО РАН	spsl.nsc.ru	7287 / 20,0	27 066	0,02	2300
ГПИБ	shpl.ru	7141 / 19,6	2 831	1,23	1500
ЦНМБ	scsml.rssi.ru	7593 / 20,8	142	0	2200
НПБ	gnpbu.ru	6873 / 18,8	294	0,19	2000
Вологодская ОУНБ	booksite.ru	6611 / 18,1	399 185	0,17	3000
Отделение ГПНТБ СО РАН	prometeus.nsc.ru	7443 / 20,4	32 631	0,01	2100

Кроме того, XTool в режиме онлайн проверяет внешние ссылки (*backlinks*) с любого сайта и генерирует полный список с указанием позиций ссылок, URL доноров и акцепторов, ТИЦ доноров, уровня вложенности страниц и анкоров. Одновременно система оценивает степень доверия к сайту и «авторитетность» ресурсов (*траст ХТ*). Хорошие, «трастовые», ресурсы получают ХТ = 8–10, «средняки» – 4–7, а некачественные и откровенно плохие сайты – ниже 3. Правда, надо подчеркнуть, что эта оценка – чисто технологическая, то есть связанная исключительно с процедурами SEO-оптимизации. Она имеет мало общего с критериями библиотечной оценки контента.

Самая существенная часть показателей внешнего трафика, посещаемости и веб-аудитории была получена с помощью вебметрической

платформы *SimilarWeb*, созданной еще в 2009 г. Этот инструмент позволяет структурировать трафик сайтов по его источникам – напрямую, ссылочному (реферальному) и поисковому, а также по заходам из социальных сетей и почтовых рассылок. Во всем мире SimilarWeb широко используется для выполнения разнообразных бизнес-исследований и глубокого вебметрического анализа данных. О возможностях системы говорит хотя бы тот факт, что она обрабатывает до 5 миллиардов событий в сутки без существенного ущерба для точности данных.

В процессе работы по сбору вебметрических данных о сайтах библиотек мы сформировали две выборки. В первую вошли сайты центральных научных библиотек, извлеченные из top-10 «Яндекс-Каталога» (представлены в табл. 1). Во второй выборке были собраны 26 сайтов региональных библиотек Сибири и Дальнего Востока, выявленные из самых разных источников, включая и профессиональный справочник ГПНТБ СО РАН «Научные библиотеки Сибири и Дальнего Востока (областные, краевые, национальные, окружные универсальные)»¹.

Отдельно по каждой выборке определялись *максимальные* и *минимальные* значения показателей, а затем в границах найденных интервалов каждая выборка делилась на четыре группы, сходные по параметрам развития. Таким образом стало возможно, хотя бы в первом приближении, получить *типологию* библиотечных сайтов, сгруппированных по некоторому способу существования этих сайтов в веб-среде. Разумеется, включение каждого сайта конкретно в одну из «моделей развития» очень условно и, как будет показано ниже, не лишено изъянов. Тем не менее проведенное нами деление позволяет судить как о некоторых важных тенденциях, так и о свойствах сайтов, полезных для их самоидентификации.

В качестве главного критерия группировки был выбран *способ взаимодействия сайтов с внешней средой* – то есть преобладающие для каждого сайта структура и источники веб-трафика, включая трафик из социальных сетей. Понятно, что метрики, полученные от сайтов

¹ Западная Сибирь. URL: http://www.spsl.nsc.ru/win/zap_sbr.htm;

Восточная Сибирь. URL: http://www.spsl.nsc.ru/win/vost_sbr.htm;

Дальний Восток. URL: http://www.spsl.nsc.ru/win/daln_vostk.htm (дата обращения: 16.02.2018 г.).

центральных библиотек, заметно отличаются от таковых же у небольших и сравнительно молодых региональных сайтов Сибири и Дальнего Востока. Так, например, сильно разнятся имеющиеся данные о размерах сайтов. Даже у центральных библиотек число страниц на сайтах может различаться на порядки (от полутора сотен у ЦНМБ до 8 млн у РНБ (см. табл. 1). Кроме того, бросается в глаза большой разброс цифр в *индексах поисковых систем* по каждой библиотеке в отдельности. Причины такой нестабильности подробно разобраны в аналитической статье по вебметрике академических институтов РАН [1, с. 1018–1021].

В таблице 2 представлены критерии группировки и отнесения сайтов к некоторым *моделям развития*, выработанные в процессе исследования. Точного совпадения проектированной схемы с реальностью, естественно, не получилось, особенно у региональных библиотек. Да и трудно было ожидать чего-то другого в таком изменчивом деле, как развитие интернета. Однако данный шаблон послужил некоторым ориентиром в проведенной работе.

Таблица 2

Критерии группировки библиотечных сайтов по структуре трафика

Модель	Источники трафика (%)					
	поиск		ссылки		соцсети	
	Ц*	Р	Ц	Р	Ц	Р
Поисковая	выше 75	выше 75	5–10	0–5	2–3	0–2
Реферальная	25–40	30–50	выше 20	выше 15	0–2	0–2
Социальная	40–50	50–70	15–20	10–15	выше 5	выше 2
Диверсифицированная	50–75	70–80	10–15	5–10	3–5	2

* Ц – центральные библиотеки; Р – библиотеки Сибири и Дальнего Востока.

Весьма условная группировка и дальнейший анализ позволили разработать четыре ведущих модели библиотечных сайтов:

- 1) *поисковую*;
- 2) *реферальную* (ссылочную);
- 3) *социальную*;
- 4) *переходную* (диверсифицированную).

У первых двух отчетливо видны противоположные подходы к генерации трафика с точки зрения *открытости* и *закрытости* представленной информации. У остальных наблюдается сбалансированное распределение визитов по всем источникам трафика. Из них развитая *социальная* модель эволюционирует, главным образом за счет продвижения ресурсов в *социальных сетях*, уделяя большое внимание созданию новых продуктов и услуг. А *диверсифицированная* модель, по существу, является *переходной* ступенью от открыто-закрытых форм сайтов (№ 1–2) к развитой социальной модели (№ 3).

В рамках *первой модели* доминирующим источником трафика является *органический поиск*, то есть приход пользователей из поисковых выдач «Яндекса», «Гугла» и других поисковиков. Ресурсы таких сайтов свободно индексируются поисковыми машинами и без излишне затянутых процедур (аутентификации, настройки сессий) предоставляются всем пользователям. Это и позволяет назвать эту модель *открытой*, так как она базируется на открытой индексации страниц, статическом коде и сессиях вида *open access*.

Вторая модель опирается на пользователей, приходящих на сайт преимущественно *по ссылкам*. Главным источником трафика для подобного «закрытого» сайта², как и главным фактором роста его посещаемости, являются известные, тематически близкие ему «доноры» ссылок. Как правило, сайты такого типа сокращают до минимума ту часть своего трафика, которая связана с поисковыми и социальными источниками. В то же время постоянно растущую часть визитов (до 20–50 % и более) начинают занимать рефералы³ и прямые заходы на сайт с «домашних» страниц, закладок и вводов URL. Сайт как будто «изолируется» от внешнего мира, переходя в автономный режим «только для своих», хорошо знакомых гостей.

Третья модель носит самый продвинутый характер. Ее можно охарактеризовать как открытую общественную модель с развитой

² В случае некоторых библиотечных сайтов можно говорить даже о своего рода «изоляции» и закрытости от внешнего мира.

³ Мы используем расширительную трактовку этого термина (от англ. *to refer* – посылать, направлять) – как всякого посетителя, приходящего на сайт по внешним ссылкам. Вместе с тем главное, более узкое значение термина отсылает к зарегистрированным пользователям разного рода коммерческих проектов, участвующим в реферальных программах, приносящих денежный доход.

ссылочной частью трафика и очень высокой долей визитов из социальных сетей. Библиотечные сайты этого типа являются результатом длительного и устойчивого развития. Они отличаются большой клиентской базой, стабильным и надежным трафиком, широкими технологическими возможностями, разнообразными веб-ресурсами с большим количеством специализированных баз данных и коллекций полнотекстовых документов.

Наконец, **четвертая модель** характеризуется распределенной и не всегда устойчивой структурой всех видов трафика – прямого, ссылочного, поискового и социального⁴. При этом крупные, давно существующие сайты центральных библиотек способны сохранять относительную стабильность в течение долгого времени, тогда как региональные сайты часто мигрируют в соседние группы – к более продвинутой социальной (№ 3) или к абсолютно «закрытой» реферальной (№ 2) модели. По многим материально-техническим и организационным причинам им трудно поддерживать длительную устойчивость, поэтому их показатели наиболее изменчивы.

Разбиение десятки сайтов центральных научных библиотек по источникам трафика выделило по три поисковых и переходных модели и по две реферальных и социальных (табл. 3).

Таблица 3

Группировка сайтов центральных научных библиотек *

URL	Поиск (%)	Ссылки (%)	Соцсети (%)	Прямой (%)
1) поисковая модель				
spsl.nsc.ru	81,03	4,74	0,66	13,34
gnpbu.ru	72,72	5,37	2,30	19,10
booksite.ru	77,93	5,86	2,31	13,51
2) реферальная модель				
benran.ru	35,64	33,65	0,65	30,06
scsml.rssi.ru	28,99	21,90	2,10	47,01

⁴ В случае прямого (*direct*) трафика, как правило, речь идет о внутренних заходах сотрудников и читателей с компьютеров библиотеки или по ранее сделанным закладкам в браузерах.

Окончание табл. 3

URL	Поиск (%)	Ссылки (%)	Соцсети (%)	Прямой (%)
3) социальная модель				
nlr.ru	49,13	18,57	3,87	26,32
shpl.ru	45,12	18,61	6,59	27,43
4) переходная модель				
rsl.ru	62,41	11,89	2,13	22,05
gpntb.ru	69,46	10,40	0,82	17,52
prometeus.nsc.ru	73,14	10,91	3,52	12,43

* По данным, полученным от SimilarWeb 25.01.2018 г.

Подчеркнем еще раз, что распределение сайтов по названным группам весьма условно и относится только к одной стороне их функционала, а именно, к *структуре трафика*. При этом границы групп крайне размыты и непостоянны. Длительный мониторинг за ряд лет выявил способность сайтов не только наращивать, но и терять накопленный «вес» показателей. Тем не менее результаты выполненной работы позволяют видеть некоторые закономерности и прогнозировать будущие тенденции (переходы сайтов от одного типа развития к другому).

Более детальный анализ показателей посещаемости, несмотря на некоторый разброс цифр в пределах групп, показывает схожесть ключевых характеристик по каждой модели (табл. 4).

Таблица 4

Показатели посещаемости сайтов центральных научных библиотек по оценке SimilarWeb от 25.01.2018 г.*

URL	Визитов (в месяц)	Время на сайте	Глубина (страниц)	Доля отказов (%)
spsl.nsc.ru	219 400	3:25	2,27	62,79
gnpbu.ru	118 870	2:06	3,76	56,33
booksite.ru	659 740	1:05	1,89	75,91
среднее по 1-й модели	332 670	2:12	2,64	65,01
benran.ru *	4 600	1:38	3,54	40,60
scsml.rssi.ru *	18 400	0:44	1,43	32,80

URL	Визитов (в месяц)	Время на сайте	Глубина (страниц)	Доля отказов (%)
среднее по 2-й модели	11 500	1:11	2,49	36,7
nlr.ru	370 280	5:29	6,93	44,21
shpl.ru	212 520	5:03	17,36	48,19
среднее по 3-й модели	291 400	5:16	12,15	46,2
rsl.ru	1 770 000	3:25	5,25	59,36
gpntb.ru	113 910	1:45	2,71	57,45
prometeus.nsc.ru *	31 300	0:37	1,46	50,20
среднее по 4-й модели	638 400	1:56	3,14	55,67

* По трем отмеченным сайтам использованы данные за 30.10.2017 г.

Последняя оценка метрик посещаемости с помощью SimilarWeb от 25.01.2018 г. не позволила получить данные по трем сайтам (БЕН РАН, ЦНМБ и отделения ГПНТБ СО РАН). По этой причине в таблицу 4 включены более ранние замеры сайтов за октябрь 2017 г. В каком-то смысле, они должны быть менее полными, чем январские, так как, например, данные SimilarWeb по источникам трафика у отделения ГПНТБ СО РАН за 25.01.2018 г. заметно улучшились, что позволило сайту переместиться из группы *поисковых* (1-я модель) в группу *диверсифицированных* (4-я модель). Показатели поискового трафика *prometeus.nsc.ru* за три месяца снизились на 2,76 % (с 75,9 до 73,14 %). В то же время реферальный трафик увеличился на 1,71 % (с 9,2 до 10,91 %), а социальный – на 2,7 % (с 0,82 до 3,52 %). Исходя из этой оценки, все показатели, внесенные в таблицу 4 по трем названным библиотекам, также должны были улучшиться.

Здесь бросаются в глаза высокие метрики глубины просмотра и времени присутствия у сайтов 3-й (*социальной*) модели и низкий уровень отказов у 2-й (*реферальной*) модели. Оба результата вполне объяснимы с точки зрения способов, по которым идет развитие. «Социальные» сайты РНБ и ГПИБ располагают большим объемом качественных полнотекстовых ресурсов, отличаются скоростью обновления и эффективным продвижением в социальных сетях (табл. 1, 3). С другой стороны, специфика сайтов «реферального» типа (БЕН РАН, ЦНМБ), выраженная в их ресурсной базе, навигации и

веб-аудитории, усиливает прямой и ссылочный трафик (табл. 3), а он, в свою очередь, подразумевает множественность переходов, не связанных с одной страницей.

Во второй выборке сайтов региональных библиотек Сибири и Дальнего Востока исследование выявило намного более пеструю и малопредсказуемую картину, чем у центральных библиотек. Сбор показателей, как и по первой выборке, производился с интервалом в три месяца – 30.10.2017 и 25.01.2018 г. Но данные за январь обнаружили не только значительные колебания по ряду параметров, но и совершенно необъяснимые «всплески» значений отдельных метрик, прежде всего в сфере социального трафика.

Известно, что взаимодействие с социальными сетями крайне противоречиво и может быть как позитивным, так и негативным, что, по мнению С. Ю. Комарова, может исказить картину доступов к сайтам научных библиотек [2]. Вместе с тем согласимся, что «сегодня социальные сети постепенно превращаются в своеобразные поисковые системы, так как многие пользователи с их помощью проводят поиск информации» [3, с. 11]. Для развития сайтов социальный трафик считается наиболее перспективным.

В нашей схеме критериев группировки (табл. 2) его высокий уровень определяет отнесение сайтов к 3-й социальной модели. С учетом влияния этого фактора, а также того, что работа библиотек, а следовательно, и отражающая ее веб-статистика в январе не носят репрезентативного характера, было решено провести группировку сайтов региональных библиотек по более стабильным данным, полученным 30.10.2017 г. (табл. 5).

Таблица 5

**Группировка сайтов научных библиотек Сибири
и Дальнего Востока по данным SimilarWeb от 30.10.2017 г.**

Местонахождение библиотеки [URL]	Страниц на сайте «Яндекс»	Поиск (%)	Ссылки (%)	Соцсети (%)
1) поисковая модель				
Анадырь [<i>library-chukotka.ru</i>]	626	63,78	2,01	2,01
Биробиджан [<i>bounb.eao.ru</i>]	19 092	86,71	5,37	0,10
Магадан [<i>mounb.ru</i>]	1070	76,57	3,05	1,88

Окончание табл. 5

Местонахождение библиотеки [URL]	Страниц на сайте (Яндекс)	Поиск (%)	Ссылки (%)	Соцсети (%)
Петропавловск-Камчатский [kamlib.ru]	1695	79,74	3,72	0,54
Тюмень [tonb.ru]	34 064	91,4	2,76	0,02
Ханты-Мансийск [okrlib.ru]	4980	77,55	3,46	0,90
Чита [zabunb.ru]	1517	76,76	3,43	1,61
2) реферальная модель				
Кемерово [kemrsl.ru]	27 034	54,99	17,73	0,52
Кызыл [lib.rtyva.ru]	265	81,25	18,75	0,00
Новосибирск [ngonb.ru]	18 699	53,89	21,35	2,38
Омск [omsklib.ru]	4929	37,51	26,04	6,13
Салехард [libraries-yanao.ru]	3939	54,44	22,87	1,41
Якутск [nlrs.ru]	14 585	50,50	21,04	1,90
3) социальная модель				
Абакан [nbdrx.ru]	14 186	54,51	13,93	2,95
Барнаул [akunb.altlib.ru]	50 271	80,51	5,85	3,41
Иркутск [irklib.ru]	39 505	52,68	12,6	3,47
Сургут [slib.ru]	2247	73,07	9,06	3,49
Томск [lib.tomsk.ru]	7344	64,16	13,56	2,80
4) переходная модель				
Благовещенск [libamur.ru]	7638	63,37	4,21	1,24
Владивосток [pgpb.ru]	1453	72,44	8,01	1,93
Горно-Алтайск [nbra.ru]	2912	74,19	14,38	0,56
Красноярск [kraslib.ru]	2664	74,54	9,43	1,05
Нижневартовск [mubis.ru]	6924	77,50	11,31	0,19
Улан-Удэ [nbrb.ru]	3817	75,09	6,52	0,36
Хабаровск [fessl.ru]	2711	72,81	14,55	1,80
Южно-Сахалинск [libsakh.ru]	12 749	72,75	6,79	0,78

Из таблицы 5 видно, что для сайтов научных библиотек Сибири и Дальнего Востока не существует правил без исключений. Так, Национальная библиотека Республики Тыва (*lib.rtyva.ru*) попала во вторую (реферальную) группу с необычайно высоким поисковым трафиком, а Омская областная научная библиотека (*omsklib.ru*) в той же самой группе отличилась высоким показателем социального трафика. Следовательно, в отношении региональных сайтов отбор, сделанный по какому-то одному формальному признаку (например, по доминированию одного из источников трафика) совершенно не исключает присутствия аномальных значений по другим параметрам – вопреки схеме, представленной в таблице 2.

В каждой группе можно выделить, условно говоря, типичных представителей своего направления. Для поисковой модели – это Тюменская областная библиотека (*tonb.ru*), для реферальной – Национальная библиотека Республики Саха (Якутия) (*nlsr.ru*), для социальной – Иркутская областная библиотека (*irklib.ru*) и для «смешанного» типа – библиотеки Владивостока (*pgpb.ru*) и Красноярска (*kraslib.ru*).

Сильная зависимость региональных сайтов от организационно-технических условий и административных решений обуславливает большую изменчивость метрик и непрерывную миграцию сайтов из одной группы в другую. При этом генерация новых ресурсов как бы уходит на второй план, не играя решающей роли для развития большинства из региональной выборки. Более того, проведенные замеры обнаружили, что некоторые сайты, видимо, даже не архивируют свой контент и периодически избавляются от части страниц. При очередных замерах размер таких сайтов оказывался меньше предыдущего. Наверное, последующий мониторинг может принести еще немало сюрпризов.

В завершение статьи отметим, что моделирование развития библиотечных сайтов показало возможности вебметрики для создания типологии сайтов, основанной на анализе структуры их трафика. Будущие прогнозы развития сайтов библиотек могут опираться на выявленные тенденции перехода сайтов от простых поисковых и «закрытых» реферальных моделей ко все более сложным и открытым социальным моделям. Выработанный подход и использованные инструменты позволят разрабатывать конкретные меры для развития и продвижения новых библиотечных ресурсов и услуг, достигая при этом наиболее эффективных результатов.

Список литературы

1. *Косяков Д. В., Гуськов А. Е., Быховцев Е. С.* Академические институты России в зеркале вебометрики // Вестн. Рос. акад. наук. 2016. Т. 86. № 11. С. 1015–1025.
2. *Комаров С. Ю.* «Позитивный» и «негативный» внешний трафик из социальных сетей к сайтам научных библиотек // Информационные технологии в гуманитарных исследованиях. Новосибирск, 2017. Выпуск 22: Очерки по науковедению. С. 128–134.
3. *Третьякова О. В., Маковеев В. Н., Чугреев В. Л.* Повышение посещаемости веб-сайтов научной организации и их продвижение в сети Интернет // Социал. пространство. 2016. № 5. С. 1–19.