

Рейтинг сайтов научных организаций СО РАН

Ю. И. ШОКИН, О. А. КЛИМЕНКО, Е. В. РЫЧКОВА, И. В. ШАБАЛЬНИКОВ
Институт вычислительных технологий СО РАН, Новосибирск, Россия
e-mail: ict@ict.nsc.ru

In the present work we analyze the websites of the SB RAS Academic Institutions using webometrics techniques. Webometrics arised as an extension of bibliometrics into the web space. Only official domains of scientific organizations with independent domain have been analyzed. In order to determine the quantitative characteristics of the web sites, we use Yandex, Google and Yahoo search engines since these tools comprehensively cover the Russian language based segment of the Internet, have relevant services for conducting research and continuously update the collected information.

Введение

В работе предлагается метод исследования сайтов академических институтов методами вебометрики. Вебометрика (webometrics) возникла после расширения понятия библиометрики на веб-пространство. Библиометрика представляет собой набор методов для изучения текстов и информации, включая анализ цитирования и контентный анализ. Термин *вебометрика* ввели Т. Almind и Р. Ingwersen в 1997 г. [1]. В статье [2] дано определение вебометрики как изучение количественных аспектов создания и использования информационных ресурсов, структур и технологий в веб-пространстве, выросшее из подходов библиометрики и информатики.

Веб-пространство — новый тип информационной системы без централизованного контроля, без координированного индексирования содержания. Однако веб-пространство — не хаотическая система, в нем происходит самоорганизация в кластеры. Веб-пространство организуется путем взаимодействия большого числа индивидуумов и групп. Связи между разными документами устанавливаются с помощью ссылок. Изучение этих ссылок с помощью методов вебометрики позволяет определять устройство веб-пространства.

Для исследования веб-сайта используются несколько количественных параметров. В 1998 г. Р. Ingwersen [3] ввел понятие Web Impact Factor (WIF). WIF — это отношение числа входящих ссылок на сайт V к общему числу страниц сайта S . Однако для определения качества сайта бывает недостаточно использовать WIF. Параметр S определяется поисковыми системами и может существенно отличаться от реального числа страниц сайта. В статье А. Noguzi [4] отмечено, что значение параметра WIF зависит от используемой поисковой системы, от языка сайта и других факторов.

Регулярные исследования университетской и академической среды ведутся лабораторией Cybermetrics Lab, которая принадлежит Consejo Superior de Investigaciones Cientificas — крупнейшему исследовательскому центру Испании. Первоначально проект Webometrics Ranking of World Universities (далее — проект Webometrics) [5] был

направлен на стимулирование веб-публикаций. Рейтинг составлялся исходя не из числа визитов и дизайна сайта — он основывался на полноте представления университета в сети. Со временем рейтинг университетов приобрел популярность, а в число оцениваемых организаций были включены академические институты и научные центры. В январе 2008 г. в рейтинг 1000 научных организаций [6] вошли Портал СО РАН [7] (54-е место) и Портал РАН [8] (91-е место).

Методология исследования сайтов

Предметом анализа является официальный домен организации, поэтому в рейтинг включены только научные организации, имеющие независимый домен. Если организация имеет более одного сетевого адреса, информация обо всех доменах изучается в совокупности.

Для оценки сайтов используются параметры S , V , R , Sc .

Параметр S — размер сайта, т. е. количество страниц, определяемое поисковыми системами. Участники проекта Webometrics работали с Google, Yahoo, Live Search и Exalead. Авторами для определения размера сайта S использовались системы Яндекс, Google, Yahoo, так как именно эти поисковые системы наиболее полно индексируют русскоязычную часть Интернета, имеют соответствующие сервисы для проведения исследований и постоянно актуализируют собранную информацию. При этом в Яндексе была обнаружена такая особенность: если у института сайт относительно “новый” (не более трех лет) или же адрес сайта существенно отличается от реального доменного имени, то Яндекс вообще не находит ни одной страницы на таком сайте. Вероятно, это может быть связано с тем, что такой сайт еще не проиндексирован Яндексом.

Параметр V — видимость сайта — означает число внешних ссылок на ресурс, которое определяется поисковыми системами. Участники проекта Webometrics использовали поисковые системы Yahoo, Live Search и Exalead. Авторы статьи определяли этот параметр с помощью Яндекса, Google, Yahoo.

Параметр R — число “мощных” файлов — означает суммарное количество файлов форматов Adobe Acrobat (.pdf), Microsoft Word (.doc) и Microsoft Powerpoint (.ppt), представленных на сайте. Информацию о таких файлах также предоставляют Яндекс, Google и Yahoo. Участники проекта Webometrics сочли необходимым учитывать данный параметр в связи с тем, что объем доступной на сайте информации не ограничивается только форматом веб-страниц. В свободном доступе на сайте должны быть документы (доклады, презентации, статьи и т. п.) в других форматах.

Параметр Sc — индекс цитирования — определялся путем суммирования документов, в которых приводятся цитаты из докладов, статей и других научных материалов. Участники проекта Webometrics использовали сведения из системы Google Scholar. Авторами статьи использовался также индекс цитирования Яндекса. Следуя методике Webometrics, авторы брали среднее арифметическое от количества страниц на сайте, выдаваемое разными поисковыми системами, аналогично — среднее арифметическое от числа внешних ссылок на сайт и среднее арифметическое от числа “мощных файлов” на сайте, выдаваемое разными поисковыми системами. Затем рейтинг подсчитывался по формуле:

$$W = 4V + 2S + R + Sc,$$

где

$$V = \log_{10} [(V_{\text{Яндекс}} + V_{\text{Google}} + V_{\text{Yahoo}})/3]; \quad S = \log_{10} [(S_{\text{Яндекс}} + S_{\text{Google}} + S_{\text{Yahoo}})/3];$$

$$R = \log_{10} [(R_{\text{Яндекс}} + R_{\text{Google}} + R_{\text{Yahoo}})/3]; \quad Sc = (Sc_{\text{Яндекс}} + Sc_{\text{Google}})/2.$$

Адреса сайтов организаций СО РАН были взяты из информационной системы “Организации и сотрудники СО РАН” [9], размещенной на Портале СО РАН.

Выводы

В таблице приведены результаты исследования 80 сайтов организаций СО РАН. Размер сайтов варьируется от нескольких десятков страниц до 30 тыс. страниц, при этом 19 организаций имеют сайты более 1000 страниц, от 100 страниц до 1000 имеют 30 организаций, а остальные — менее 100 страниц, т. е. около 40 % организаций имеют сайты с малым числом страниц.

Более 1000 внешних ссылок на сайт есть у 13 организаций, меньше 100 ссылок — у 20 организаций. То есть 25 % сайтов имеют мало внешних ссылок.

Итоговый рейтинг сайтов организаций СО РАН от 8 апреля 2008 г.

№	Организация	Веб-адрес	V	S	R	Sc	W
1	Портал СО РАН	www.sbras.ru	4.89	5.26	4.08	6536	6570
2	Институт вычислительных технологий СО РАН	www.ict.nsc.ru	4.15	4.43	2.55	3945	3972
3	Институт ядерной физики им. Г.И. Будкера СО РАН	www.inp.nsk.su	3.44	3.91	3.04	1239	1263
4	Институт цитологии и генетики СО РАН	www.bionet.nsc.ru	3.38	3.88	2.77	1210	1234
5	Государственная публичная научно-техническая библиотека	www.spsl.nsc.ru	3.55	3.99	2.15	1129.5	1153
6	Институт математики им. С.Л. Соболева СО РАН	www.math.nsc.ru	3.00	3.96	3.08	1065	1088
7	Институт солнечно-земной физики СО РАН	www.iszf.irk.ru	2.20	2.11	1.85	937.5	952
8	Институт систем информатики им. А.П. Ершова СО РАН	www.iis.nsk.su	4.07	3.45	2.22	866	891
9	Президиум СО РАН	www.sbras.nsc.ru	3.77	4.33	2.95	700.5	727
10	Институт философии и права СО РАН	www.philosophy.nsc.ru	2.92	3.68	2.40	668	689
11	Институт вычислительной математики и математической геофизики СО РАН	www.sccc.ru	2.84	2.73	1.39	566	584
12	Институт проблем освоения Севера СО РАН	www.ipdn.ru	1.73	3.73	2.10	565	581
13	Институт неорганической химии им. А.В. Николаева СО РАН	www.che.nsk.su	2.67	2.90	1.38	558	575

Итоговый рейтинг сайтов организаций СО РАН от 8 апреля 2008 г. (продолжение)

№	Организация	Веб-адрес	<i>V</i>	<i>S</i>	<i>R</i>	<i>Sc</i>	W
14	Институт оптики атмосферы им. В.Е. Зуева СО РАН	www.iao.ru	2.41	2.14	0.00	553	566
15	Новосибирский институт органической химии им. Н.Н. Ворожцова СО РАН	www.nioch.nsc.ru	2.70	3.72	3.15	518.5	539
16	Институт автоматизации и электрометрии СО РАН	www.iae.nsk.su	2.75	3.05	1.95	512.5	531
17	Институт химической кинетики и горения СО РАН	www.kinetics.nsc.ru	2.57	3.17	2.28	485	503
18	Институт физики им. Л.В. Киренского СО РАН	www.kirensky.ru	2.44	3.36	2.50	394	412
19	Институт экономики и организации промышленного производства СО РАН	ieie.nsc.ru	3.32	3.45	1.54	387.5	409
20	Институт вычислительного моделирования СО РАН	icm.krasn.ru	3.28	3.49	0.88	351	371
21	Институт физики полупроводников им. А.В. Ржанова СО РАН	www.isp.nsc.ru	2.75	2.75	1.43	351	368
22	Институт леса им. В.Н. Сукачева СО РАН	forest.akadem.ru	2.30	3.32	2.52	315.5	333
23	Институт космофизических исследований и аэронавтики им. Ю.Г. Шафера СО РАН	ikfia.ysn.ru	2.87	2.08	1.28	313.5	330
24	Институт земной коры СО РАН	www.crust.irk.ru	2.75	3.17	1.51	251.5	270
25	Институт археологии и этнографии СО РАН	www.archaeology.nsc.ru	2.53	2.54	1.14	252.5	268
26	Институт катализа им. Г.К. Борескова СО РАН	www.catalysis.ru	2.61	3.15	1.81	238	256
27	Институт динамики систем и теории управления СО РАН	www.idstu.irk.ru	1.90	1.72	1.00	238	250
28	Институт систем энергетики им. Л.А. Мелентьева СО РАН	www.sei.irk.ru	3.41	2.64	1.88	216	236
29	Институт теоретической и прикладной механики им. С.А. Христиановича СО РАН	www.itam.nsc.ru	2.68	3.60	2.28	212	232
30	Институт биофизики СО РАН	www.ibp.ru	2.41	2.77	1.98	214.5	231
31	Институт химии нефти СО РАН	www.ipc.tsc.ru	2.44	2.69	2.10	202	219

Итоговый рейтинг сайтов организаций СО РАН от 8 апреля 2008 г. (продолжение)

№	Организация	Веб-адрес	<i>V</i>	<i>S</i>	<i>R</i>	<i>Sc</i>	W
32	Томский научный центр СО РАН	www.tsc.ru	2.88	2.34	0.12	201	217
33	Красноярский научный центр СО РАН	www.krasn.ru	2.19	1.94	0.88	181,5	195
34	Центральный сибирский бо- танический сад СО РАН	www.csbg.narod.ru	3.25	2.08	1.05	162.5	180
35	Институт криосферы Земли СО РАН	www.ikz.ru	3.17	3.41	2.04	127	148
36	Геологический институт СО РАН	geo.buryatia.ru	2.32	2.56	1.44	106.5	122
37	Институт теплофизики им. С.С. Кутателадзе СО РАН	www.itp.nsc.ru	2.50	2.98	2.30	104	122
38	Кемеровский научный центр СО РАН	www.kemsc.ru	2.41	2.23	0.87	106.5	121
39	Институт горного дела СО РАН	www.misd.nsc.ru	2.59	3.04	1.31	90.5	108
40	Институт геохимии им. А.П. Виноградова СО РАН	www.igc.irk.ru	1.97	1.84	1.48	93.5	106
41	Институт химии твердого тела и механохимии СО РАН	www.solid.nsc.ru	2.29	3.06	2.79	85.5	103
42	Институт физики прочности и материаловедения СО РАН	www.ispms.ru	2.26	3.57	1.89	85.5	103
43	Институт гидродинамики им. М.А. Лаврентьева СО РАН	www.hydro.nsc.ru	2.01	1.88	0.80	85	97
44	Институт водных и экологи- ческих проблем СО РАН	iwep.asu.ru	2.50	2.65	2.13	78	95
45	Лимнологический институт СО РАН	www.lin.irk.ru	2.07	2.22	1.03	82	95
46	Институт систематики и экологии животных СО РАН	www.eco.nsc.ru	2.08	1.90	0.95	70.5	83
47	Институт химической био- логии и фундаментальной медицины СО РАН	www.niboch.nsc.ru	2.24	2.36	2.08	62.5	78
48	Институт географии им. В.Б. Сочавы СО РАН	irigs.irk.ru	2.47	2.83	1.09	61	77
49	Институт истории СО РАН	www.history.nsc.ru	2.18	2.92	1.24	60	75
50	Иркутский институт химии им. А.Е. Фаворского СО РАН	www.inchemistry.irk.ru	1.89	2.25	1.25	61.5	74

Итоговый рейтинг сайтов организаций СО РАН от 8 апреля 2008 г. (продолжение)

№	Организация	Веб-адрес	V	S	R	Sc	W
51	Институт лазерной физики СО РАН	www.laser.nsc.ru	2.00	2.28	0.67	59.5	72
52	Конструкторско-технологический институт научного приборостроения СО РАН	www.tdisie.nsc.ru	2.36	2.20	1.49	52.5	67
53	Иркутский научный центр СО РАН	www.isc.irk.ru	2.14	2.16	0.30	50.5	63
54	Институт химии и химической технологии СО РАН	www.icct.ru	1.65	2.60	1.75	46	59
55	Конструкторско-технологический институт вычислительной техники СО РАН	www.kti.nsc.ru	2.05	2.25	0.48	45	58
56	Геофизическая служба СО РАН	www.gs.nsc.ru	1.88	1.88	0.00	40.5	51
57	Сибирский институт физиологии и биохимии растений СО РАН	www.sifibr.irk.ru	1.29	1.74	1.20	42	51
58	Институт "Международный томографический центр" СО РАН	www.tomo.nsc.ru	2.11	2.47	1.96	35	50
59	Новосибирский филиал Института водных и экологических проблем СО РАН	iwep.ad-sbras.nsc.ru	1.82	0.56	0.00	42	50
60	Институт мерзлотоведения им. П.И. Мельникова СО РАН	mpi.ysn.ru	2.06	1.27	0.00	38.5	49
61	Институт почвоведения и агрохимии СО РАН	www.issa.nsc.ru	2.21	1.68	0.00	25	37
62	Конструкторско-технологический филиал института гидродинамики им. Лаврентьева СО РАН	www.kti-git.nsc.ru	2.22	1.56	0.00	25	37
63	Институт нефтегазовой геологии и геофизики им. А.А. Трофимука СО РАН	www.ipgg.nsc.ru	1.80	3.00	1.92	20	35
64	Омский научный центр СО РАН	osc.okno.ru	2.21	2.13	1.31	20	34
65	Омский филиал Института математики им. С.Л. Соболева СО РАН	ofim.okno.ru	1.92	2.31	1.46	17.5	31
66	Институт мониторинга климатических и экологических систем СО РАН	www.imces.ru	1.90	2.26	0.22	15.5	27
67	Институт сильноточной электроники СО РАН	www.hcei.tsc.ru	2.51	2.70	1.71	6	23
68	Институт физико-технических проблем Севера СО РАН	iptpn.ysn.ru	1.97	1.44	0.00	12.5	23
69	Тюменский научный центр СО РАН	www.tmnsc.ru	2.11	2.04	0.80	10	23
70	Институт проблем химико-энергетических технологий СО РАН	www.ipcet.ru	1.41	1.46	0.85	10	19

Итоговый рейтинг сайтов организаций СО РАН от 8 апреля 2008 г. (продолжение)

№	Организация	Веб-адрес	<i>V</i>	<i>S</i>	<i>R</i>	<i>Sc</i>	W
71	Институт природных ресурсов, экологии и криологии СО РАН	inrec.chita.ru	2.00	2.59	0.60	5	18
72	Байкальский институт природопользования СО РАН	www.binm.ru	1.71	2.55	0.67	5	17
73	Институт геологии алмаза и благородных металлов СО РАН	www.diamond.ysn.ru	1.48	2.06	1.15	5	16
74	Институт геологии и минералогии им. В.С. Соболева СО РАН	www.igm.nsc.ru	1.54	2.19	1.00	5	16
75	Институт горного дела Севера им. Н.В. Черского СО РАН	igds.ysn.ru	1.79	1.74	0.00	5	15
76	Якутский научный центр СО РАН	www.ysn.ru/services	2.21	0.00	1.81	5	15
77	Институт угля и углехимии СО РАН	www.kemsc.ru/ICC/	0.60	2.02	0.56	5	12
78	Тюменский филиал Института теоретической и прикладной механики им. С.А. Христиановича СО РАН	www.timms.tmnsc.ru	0.78	1.39	0.00	5	10
79	Иркутский филиал Института лазерной физики СО РАН	www.isc.irk.ru/ifilf.htm	0.94	0.00	0.22	5	8
80	Омский филиал Института физики полупроводников СО РАН	www.obisp.okno.ru	0.37	0.00	0.00	5	6

Число “мощных” файлов на сайте больше у тех институтов, которые предоставляют в свободный доступ полные тексты документов (статей, докладов, отчетов, инструкций и т. п.). Только пять организаций имеют сайты с числом “мощных” файлов более 1000.

Индекс цитирования более 1000 определяется в системе Google Scholar у трех сайтов. Следует заметить, что у большинства сайтов институтов СО РАН индекс цитирования в Google Scholar низкий из-за того, что сайты содержат мало информации на английском языке. Индекс цитирования в Яндексе у 13 сайтов более 1000, индекс цитирования от 100 до 1000 имеют 37 организаций.

Итоговый рейтинг показывает, что не более 15 организаций СО РАН имеют хорошо развитые сайты, которые известны в веб-пространстве.

Список литературы

- [1] ALMIND T., INGWERSEN P. Infometric analyses on the World Wide Web: methodological approaches to 'webometrics' // J. of Documentation. 1997. Vol. 53, N 4. P. 404–426.

- [2] BJORNERBORN L., INGWERSEN P. Perspectives of webometrics // Scientometrics. 2001. Vol. 50, N 1. P. 65–82.
- [3] INGWERSEN P. The calculation of Web Impact Factors // J. of Documentation. 1998. Vol. 54, N 2. P. 236–243.
- [4] NORUZI A. The Web Impact Factor: A critical review // The Electronic Library. 2006. Vol. 24, N 4. P. 490–500.
- [5] ПРОЕКТ “Webometrics Ranking of World Universities”. <http://www.webometrics.info/>
- [6] РЕЙТИНГ научных организаций Webometrics.
http://www.webometrics.info/top1000_r&d.asp
- [7] ПОРТАЛ СО РАН. <http://www.sbras.ru>
- [8] ПОРТАЛ РАН. <http://www.ras.ru>
- [9] ИНФОРМАЦИОННАЯ система “Организации и сотрудники СО РАН”.
<http://www.sbras.ru/sbras/db/>

Поступила в редакцию 12 мая 2008 г.