

ОРГАНИЗАЦИЯ НАУКИ

Покажи мне свой сайт, и я скажу, кто ты

Исследование того, как научная организация представлена в веб-пространстве — это новый способ определения ее эффективности. В числе других он будет использоваться при проверке академических институтов, которую проведет ФАНО. Принципы работы метода были представлены на семинаре в Институте автоматизации и электрометрии СО РАН



«Сегодня сложилась такая ситуация, когда для научных организаций все большее значение приобретают различные рейтинги. Ученым постоянно приходится отчитываться о цитируемости и профессиональных контактах, очень важную роль играет импакт-фактор: персональный, журнала, института. Поскольку эти показатели стали влиять на размер зарплаты, в научном мире активно распространяется такое явление, как «дружественное цитирование» (когда за счет своих публикаций ученый поднимает позиции товарища), — говорит заведующий лабораторией прикладного анализа Института математики им. С.Л. Соболева СО РАН чл.-корр. РАН **Андрей Юрьевич Веснин**. — Что же касается рейтингов, связанных с присутствием научной организации в интернете, то их уже начали активно использовать, но пока еще не пытаются накручивать. Поэтому на сегодняшний день, с моей точки зрения, они дают довольно объективную информацию».

Группа ученых из Института вычислительных технологий СО РАН совместно с коллегами из ИМ СО РАН занимается изучением академического веб-пространства. В последние три года работы проводились в рамках интеграционного проекта Сибирского отделения «Исследование закономерностей и тенденций развития самоорганизующихся систем на примере веб-пространства и биологических сообществ».

Все началось с рейтинга научных организаций, разработанного в 2000-х годах испанскими учеными (сегодня он уже является международным), который изучает сайты, исходя не из количества посещений, а, по словам создателей, принимая во внимание «известность в академических сообществах». Она высчитывается с помощью некоторых параметров, извлекаемых с помощью поисковых систем (google и прочие): числа страниц на сайте, того, сколько на него есть внешних ссылок, количества присоединенных документов в форматах, наиболее часто используемых для публикации научных статей и презентаций, и цитируемости в интернете.

«Первый свой рейтинг мы построили в 2008 году. Сначала он был почти такой же, как у испанских коллег, но потом у нас возникли некоторые вопросы, которые имеющаяся система не решала. В частности, экспертная оценка показала, что получаемые данные соответствуют действительности (информация проверялась с помощью



сторонних программ), но недостаточно точны. Тогда мы решили построить свою программу по исследованию сайтов и ссылок между ними. Она была отлажена на ресурсах Сибирского отделения, и затем с ее помощью было изучено около 300 сайтов научных организаций России, Германии и Сербии», — рассказывает заместитель главного ученого секретаря СО РАН, старший научный сотрудник Института вычислительных технологий СО РАН к.ф.-м.н. **Ольга Анатольевна Клименко**.

Сообщества академических организаций в системе новосибирских ученых представлены в виде графов, на которых показаны сайты (вершины) и существующие ссылки между ними (ребра). Так, на получившихся картинках видно, что институты СО РАН очень тесно друг с другом взаимодействуют. Внутри больших сообществ выделяются маленькие кластеры. Например, граф Сибирского отделения наглядно демонстрирует, что ИАиЭ СО РАН наиболее тесно сотрудничает с ИВТ СО РАН, ИМ СО РАН и Объединенным ученым советом СО РАН по нанотехнологиям и информационным технологиям, а с организациями физического профиля контактов у него оказалось не так много.

Сравнение результатов работы программы с независимой оценкой экспертов показало, что она отражает реальное положение дел. «Когда тот или иной интеграционный научный проект признается удачным, это отражается в веб-пространстве в виде ссылок. Если же сотрудничество оказалось чисто формальным и не пошло, то и в интернете его не наблюдается», — утверждает Ольга Анатольевна. Однако для составления полной картины имеющихся методов оказалось недостаточно, и к проекту привлекли ученых ИМ СО РАН, чтобы они разобрались, как устроена эта система и как ею управлять.

Андрей Юрьевич Веснин рассказал, что сейчас наибольшей популярностью пользуется модель галстука-бабочки. Каждый интернет-граф здесь пытаются разбить на компоненты, где в центре — сильно связанные структуры (это множество вершин, где из одной всегда можно попасть в другую), а по краям — те, из которых легко перейти к первым и те, к которым легко перейти из первых (иногда они совпадают со вторыми, иногда — нет).

Пользуясь знаниями о том, как выстраиваются отношения между компонентами графа, можно поднять отдельно взятый сайт в том или ином рейтин-

ге. Первый способ: наладить связь с самым сильным элементом системы. Как правило, когда человек ходит по интернету, он продвигается по ссылкам не более чем на шесть шагов (здесь действует знаменитый «принцип шести рукопожатий»). Чем больше у лидера ссылок на вас, тем чаще пользователи, заходящие на его страницу (а поскольку он популярен, их очень много), будут попадать по ссылке к вам.

Второй способ: сговориться с аутсайдером. Например, существует некоторый граф из 106 вершин. Группа из пяти элементов задалась целью поднять свои позиции. В таком случае план действий таков: сначала каждый ее член делает много ссылок на каждого. Когда кто-то новый приходит на один из «сговорившихся» сайтов, он попадает в этот круг и какое-то время в нем прокручивается, тем самым повышая рейтинг всех пяти участников. Здесь имеет значение именно замкнутость, небольшая величина группы, ведь если она будет слишком крупной, такого эффекта не получится. Дополнительный бонус возникает, если кому-то из участников удастся наладить связь с лидером — так он автоматически повышает позиции дружественных сайтов.

Третий способ: самим сделать ссылки на всех. Таким образом, ваши пользователи будут заходить на множество других вершин графа, и в программе вы будете выглядеть как сильносвязанный компонент.

Известны случаи, когда, зная параметры рейтинга, научно-образовательные организации пытались продвинуться в нем с помощью специальных ухищрений. Один московский университет загрузил на свой веб-ресурс всю информацию, накопленную за предыдущие годы. Так им удалось сделать его очень «тяжелым» и на какое-то время поднять позиции. Другой, уже новосибирский, вуз на своем сайте предоставил студентам возможность поздравлять друг друга с днем рождения, и — очевидно — посещаемость резко возросла.

По словам исследователей, поскольку все организации с разной степенью серьезности подходят к наполнению своих веб-ресурсов, исходные данные являются недостаточно точными. В результате и рейтинг получается примерным. Первые 20 сайтов иногда немного меняются друг с другом местами, но остаются лидерами. Так что надо ставить себе задачу не занять среди них какое-то определенное место, а хотя бы попасть в их число.

Также ученые отмечают, что любые уловки по накрутке помогут лишь на время. Исследования показали: реальная академическая репутация организации все равно даст о себе знать. «Лидерство в науке и в веб-пространстве коррелируются. Если институт действительно сильный, то его сайт, даже будучи совершенно старым и «замшелым», будет иметь очень высокий рейтинг, потому что там содержится важная информация, на которую часто ссылаются», — комментирует Ольга Анатольевна.

Диана Хомякова
Фото автора

СО РАН в рейтингах журнала Nature

На сайте журнала Nature опубликован обновленный рейтинг публикационной активности научных организаций — Nature Index 2014 Global. Эта база данных компилируется Nature Publishing Group в сотрудничестве с Digital Science и находится в открытом доступе. Данные в рейтинге обновляются ежемесячно, сообщается в пресс-релизе Отделения ГПНТБ СО РАН.

Рейтинг включает 500 институтов и университетов по всему миру и по умолчанию ранжирован по «взвешенной доле авторов».

При анализе учитываются только исследовательские статьи из 68 самых высокорейтинговых (по Web of Science) журналов. Основным показателем, используемым при анализе базы Nature Index — это число статей (AC — Articles Count). Например, если в статье есть один или несколько соавторов из России, то России присваивается AC=1. Соответственно по единице присваивается научной организации, которую указал автор в качестве аффилиации.

Используется также дробный показатель (FC — Fractional Count), который показывает вклад института или страны в создание статьи. По умолчанию считается, что авторы внесли равный вклад в статью, поэтому FC всей статьи равен 1. Если у статьи пять авторов, то каждый получает по FC=0,2. Соответственно, чем больше авторов у статьи, тем меньше вклад каждого.

Третий показатель — взвешенный дробный показатель (WFC — Weighted Fractional Count) — стал необходим, когда в базе статей обнаружился уклон в астрономию. Nature Index содержит в пять раз больше статей по астрономии и астрофизике, чем по всем другим дисциплинам. WFC получается умножением FC на 0,2.

Российская академия наук находится на 21 месте за период с 1 декабря 2013 по 30 ноября 2014 г. Из 139 институтов РАН — 33 института СО РАН. НГУ занимает 461 место в рейтинге.

Описание, руководство и комментарии к рейтингу опубликованы в Nature, 2014, Vol. 515 No. 7526 supp.

Виктор Иванов, ГПНТБ СО РАН

nature
INDEX 2014
GLOBAL

