

ОБЗОР ВЕБОМЕТРИЧЕСКИХ ПОДХОДОВ К ИЗУЧЕНИЮ ДЕЯТЕЛЬНОСТИ УНИВЕРСИТЕТОВ

А.Д. Кузнецова

*Национальный Исследовательский Университет «Высшая Школа Экономики»
Санкт-Петербург*

Сегодня Интернет дает неограниченные возможности, в том числе и способность получать информацию об объектах без конкретной интеракции с ними. Университеты, будучи активными участниками сети, могут также быть оценены через призму Интернета. Многие рейтинги университетов, действующие как инструмент выбора и оценки университетов, вызывают значительный объем критики, который зачастую сводится к выбору показателей, составляющих рейтинг, их весов и способов агрегации. Альтернативные подходы, такие как альтметрики и вебометрики, основывающиеся на измерении веб-присутствия вузов, вызывают растущий интерес и на сегодня являются одними из популярных показателей для анализа и рейтингования деятельности университетов.

Инфометрическое направление исследований является сравнительно новым, что затрудняет его четкую операционализацию. Tague-Sutcliffe [1] дает следующее определение инфометрии: "Инфометрия — изучение количественных аспектов информации в любой ее форме, а не только в виде записей или библиографии, причем в любой социальной группе, а не только ученых. Egghe [2] пишет об инфометрии как "широкий термин, охватывающий все метрики исследований, которые относятся к информационным наукам, включающих библиометрию, наукометрию, вебометрию.

Библиометрический подход включает в себя анализ библиометрических показателей, индексов цитирования. Это позволяет оценить научно-академическую деятельность университетов. Ввиду этого мы используем библиометрические подходы к изучению активности университетов через анализ библиометрических данных и данных цитирования в WoS и Scopus.

Вебометрики — это способ анализа веб-присутствия, который является под частью киберметрики [3]. Ядром данного метода является Фактор Веб-Влияния (Web Impact Factor WIF) — нормализованное количество уникальных входящих ссылок на домен университета [4]. Суть в том, что общее количество ссылок на страницу сайта и среднее количество ссылок на страницу могут служить индикаторами ценности и важности информации на этих страницах. В соответствии с исследованием Smith и Thelwall [5] Фактор Веб-Влияния позитивно коррелирует с реальной (оффлайн) деятельностью университета, что увеличивает возможности анализа и в некотором смысле упрощает его. Это своего рода попытка оценить реальную (оффлайновую) деятельность университетов на основании онлайн-данных: ссылок на домен университета (Web Impact Factor), результатов выдачи поисковых движков. Исследования показывают, что данный метод способен достаточно точно оценить размах научной деятельности университета, не используя напрямую данные из библиометрических баз WoS и Scopus. Однако необходимо отметить, что вебометрики являются скорее комплементарными по отношению к уже сложившимся библиометрическим показателям.

В то время как методологии веб-ранжирования, например, заявленные в Cybermetrics Lab [6], уделяют большое внимание анализу доменов университетов, мы предполагаем, что ключевая информация об университетах в целом принадлежит другим агентам, представленным также и в сети: индустриальным фирмам и новостным СМИ. Альтметрики и вебометрики, основывающиеся на измерении веб-присутствия вузов, вызывают растущий интерес, однако, не учитывают контекст и специфику веб-присутствия. Таким образом, возникает необходимость разработки способов анализа веб-присутствия и научной деятельности вузов, дающие достаточно степеней свободы для учета их специфики, но также и обеспечивающие возможность сопоставлений. Введение локального контекста — изучение ссылок из ключевых федеральных и региональных СМИ и топа работодателей и партнеров — позволяет произвести более подробную оценку деятельности университетов.

Проблема необходимости создания новых индикаторов для ранжирования университетов стоит довольно остро [7]. Она во многом определяется желанием университетов занимать более высокие позиции в рейтингах, что сподвигает многие университеты на обман: приведение неправильной информации о данных приема, выпускниках, баллах ЕГЭ и т.д. [8]. Также данные количественные показатели не учитывают специфику деятельности университета и контекст его активности — сухой численный подсчет не дает полного понимания процессов, происходящих в и вокруг университета.

Существуют альтернативные способы, близкие к рассматриваемым в работе, описывающие деятельность университетов, используя современные онлайн-данные. Некоторые исследователи используют семантический анализ текстов СМИ, в которых упоминаются университеты, с целью понимания общественного отношения и оценки вуза. Семантический анализ коротких сообщений длиной до 140 сообщений в сети Twitter, в которых упоминаются университеты, был проведен Gunduz, Demirhan, и

Sagiroglu [9]. В результате их исследования оказалось, что связь между эмоциональной окраской, используемой лексикой, описывающей вуз, и его положением в рейтинге отсутствует.

Альтернативным подходом к рейтингованию университетов является вебметрический. Проект Webometrics Ranking of World Universities [10] основан на ранжировании на базе объема веб-контента (количества страниц и файлов) и видимого влияния публикаций по числу цитирований в социальных медиа. Это крупнейший академический рейтинг высших учебных заведений, предоставляющий надежную и многомерную информацию о работе университетах всего мира. Основная цель проекта — стимулирование вузов к академическому “присутствию” в сети: открытый доступ к научным и культурным знаниям. Как уже ранее отмечалось, рейтинги оказывают огромное влияние на деятельность университетов. Следовательно, академические рейтинги будут способствовать развитию академической/научной деятельности вузов. С развитием сети Интернет, пространство сети стало “зеркалом” деятельности университетов и научным инструментом коммуникации. Ключевым моментом рейтинга Webometrics является то, что он анализирует не образовательную деятельность в целом, а представленность вуза в интернет пространстве путем сравнения научно-исследовательских достижений их сайтов.

Методология рейтингования Webometrics основана на индексе Ingwersen [4] — Web Impact Factor.

$WIF = L/W$, где L — ссылки (links) и W (webpages). Исследователи из Webometrics добавляют веса к этим составляющим: $WIF = aL/bW$

С помощью основных поисковых систем получаются 4 количественных показателя:

- Размер (S) — количество страниц сайта, извлеченных из: Google, Yahoo и Bing Search;
- Видимость (V) — общее число уникальных внешних ссылок в Yahoo Search, Live Search, Exalead;
- Ценные файлы (R) — число выложенных на сайте документов в формате: Adobe Acrobat (.pdf), Adobe PostScript (.ps), Microsoft Word (.doc) и Microsoft Powerpoint (.ppt);
- Scholar (Sc) — число публикаций и цитат, найденных в Google Scholar.

Тем не менее, процесс оценивания объектов должен обладать несколькими измерениями с опорой на различные количественные показатели, способные дать описание разнообразных сторон деятельности университетов. Сочетание наукометрических показателей, являющихся также частью инфометрии, и вебметрик позволит глубже понять и оценить как публичную, так и научную деятельность вузов. Релевантными для России индексам цитирования, в качестве наукометрических показателей, считаются данные из таких баз данных как Web of Science Core Collection, Russian Science Citation Index, Scopus, РИНЦ. Вебметриками же выступают показатели представленности вузов в Интернете, на сайтах ключевых компаний-работодателей и партнеров и СМИ — а именно, количество внешних уникальных ссылок на домен университета; количество упоминаний на сайтах компаний и СМИ.

Статья подготовлена в ходе проведения исследования 17-05-0024 в рамках Программы «Научный фонд Национального исследовательского университета „Высшая школа экономики“ (НИУ ВШЭ)» в 2016 – 2017 гг. и в рамках государственной поддержки ведущих университетов Российской Федерации «5-100».

ЛИТЕРАТУРА

1. Tague-Sutcliffe J. An introduction to informetrics // Information processing & management. 1992. Vol. 28. №. 1. Pp. 1-3.
2. Egghe L. Expansion of the field of informetrics: Origins and consequences // Information Processing & Management. 2005. Vol. 41. №. 6. Pp. 1311-1316.
3. Ingwersen P., Björneborn L. Methodological issues of webometric studies // Handbook of quantitative science and technology research. – Springer Netherlands, 2004. Pp. 339-369.
4. Ingwersen P. The calculation of web impact factors // Journal of documentation. 1998. Vol. 54. №. 2. Pp. 236-243.
5. Smith A., Thelwall M. Web impact factors for Australasian universities // Scientometrics. 2002. Vol. 54. №. 3. Pp. 363-380.
6. Aguillo I. F., Ortega J. L., Fernández M. Webometric ranking of world universities: Introduction, methodology, and future developments // Higher education in Europe. 2008. Vol. 33. №. 2-3. Pp. 233-244.
7. Daraio C., Bonaccorsi A. Beyond university rankings? Generating new indicators on universities by linking data in open platforms // Journal of the Association for Information Science and Technology. 2016.
8. Espeland W. N., Sauder M. Rankings and reactivity: How public measures recreate social worlds1 // American journal of sociology. 2007. Vol. 113. №. 1. Pp. 1-40.
9. Gunduz S., Demirhan F., Sagiroglu S. Investigating sentimental relation between social media presence and academic success of Turkish universities // Machine Learning and Applications (ICMLA), 2014 13th International Conference on. IEEE, 2014. Pp. 574-579.
10. RANKING WEB OF UNIVERSITIES. <http://www.webometrics.info/en>.