



**СТЕРЛИГОВ
ИВАН АНДРЕЕВИЧ**

советник проректора НИУ ВШЭ

РЕЙТИНГИ В НАУКЕ: КТО-ТО ВИНОВАТ?

В вопросах производства и применения рейтингов наука и наукометрия весьма любопытны, однако для полноценного осмысления этого многообразия нужно быть специалистом в психологии, социологии, экономике и менеджменте — и это помимо самой наукометрии. Я худо-бедно разбираюсь разве что в последней, поэтому нижеследующую заметку можно считать любительской: говорить я буду не о показателях как таковых, а о том, почему они применяются. Опираясь придется не только на литературу, но и на свой профессиональный опыт (но предметно высказываться о российской научной политике не буду, извините).

Рейтинги сейчас везде, это характерный признак *evaluation society* с его тотальной измеримостью и сравнимостью (об этом ярко и разносторонне писал Петер Далер-Ларсен¹). В целом их проще всего воспринимать как часть общего движения в сторону рациональной эффективности — или кажимости такой эффективности — во всем подряд, характерного для современного этапа развития человечества с его переносом бизнес-логики на прочие области деятельности, включая науку.

Фактически в большинстве случаев в науке рейтинги выполняют стандартную функцию *инструментов принятия решений* — *judgement devices* в терминологии Люсьена Карпика², чья концептуальная схема используется в том числе в социологических исследованиях роли метрик в научной жизни. Со стороны экономики здесь также наработан целый набор различных теорий об информационной асимметрии, проблеме агента-принципала, сигналинге, скрининге и прочих важных для Уолл-стрит материях, применяющихся в том числе в науковедении. Для науки и практики госуправления R&D рейтинги и аналогичные инструменты вообще являются

¹ Dahler-Larsen P. (2011) *The Evaluation Society*. Stanford: Stanford University Press.

² Karpik L. (2010) *The Economics of Singularities*. Princeton: Princeton University Press.

основой performance-based funding-систем, доминирующих в научной политике современной Европы и Азии. В общем, материала хватает.

При этом складывается парадоксальная ситуация (подозреваю, аналогичная ситуациям в каких-то других областях): очень многие ученые на словах настроены по отношению к метрикам критически, и их демонстративное отрицание сейчас, пожалуй, где-то даже является частью научного этоса. Ключевой документ/процесс здесь — Сан-францисская декларация об оценке науки (DORA), по сути запрещающая оценивать статьи и их авторов по рейтингам журналов. Сайт DORA — хрестоматийный пример академического популизма с красивой заглавной страницей, где изображены тысячи одухотворенных рисерчеров, смотрящих куда-то ввысь, большая яркая кнопка «подписать» и маленькая ссылка «прочитать», которую кликать необязательно.

Нет сомнений, что почти все подписывающиеся, если, конечно, они занимаются реальной наукой, на деле продолжают повседневно использовать журнальные метрики, бренды и рейтинги, то есть речь идет скорее о virtue signalling и консолидации сообщества вокруг декларируемых ценностей. Эта двойственность не сильно смущает, так как воспринимается прежде всего в контексте борьбы с «глупыми чиновниками», якобы одержимыми примитивной наукометрией. Такая трактовка прикладной наукометрии как чего-то наведенного извне зародившимися в недрах тэтчеризма неолиберальными менеджерами и примкнувшими к ним университетскими функционерами, отдающими автономию в обмен на ресурсы, отчасти характерна для рассчитанных на широкую научную аудиторию трудов самих науковедов, например, Ива Женгра.

Подобные представления могут казаться упрощением, потому что рейтинги и метрики нужны прежде всего самим ученым и ими и используются, в том числе без всякого внешнего давления. И будут использоваться все больше и больше, причем в целом ряде сценариев. Но прежде чем перейти к ним, пора уточнить, что же я понимаю под рейтингами.

В английском языке есть два термина: rating («вот тебе яблоко, оцени его вкус по шкале от 1 до 3 баллов») и ranking («вот тебе три яблока, расположи их по вкусоности»). Первый процесс, по сути, занимается отнесением сущностей к категориям либо присвоением параметрических метрик, второй — выстраиванием таблицы, где каждая сущность выше, ниже или — редко — равна другой. Очевидно, процессы эти далеко не всегда противоречат друг другу, как, например, в случае импакт-факторов журналов, но разница в подходах тем не менее принципиальна. В науке самой по себе, как мне кажется, исключительно важны рейтинги первого типа (категориальные), а вот ранжирующие сопоставимой роли не играют. В управлении наукой вторые более важны, но тоже явно уступают первым. Почему так?

Прежде всего, ранжирующие рейтинги по определению довольно антинаучны. Как мы можем строго научно утверждать, что Колмогоров выше или ниже, а то и «круче» Гильберта? По числу ссылок в Web of Science? Сама постановка такой задачи вызовет у математика недоумение. У его начальника может не вызвать, но применение ранжирующих рейтингов для приема на работу («возьмем ее, у нее Хирш выше остальных») сейчас характерно только для стран с низкой исследовательской культурой, а в других чаще используются наборы категориальных метрик («у него пять статей в журналах с импактом выше пяти, из них три последним автором» для биологии или «вот эта хороша, пробилась в *Econometrica*, а еще есть материал в NBER» для экономики³).

Еще одним значимым примером ранжирующих инструментов служат коммерческие рейтинги вузов, но они-то как раз являются сугубо внешним по отношению к науке явлением, предназначены в первую очередь абитуриентам и чиновникам, а науковедами/наукометристами в качестве научных не признаются вовсе. Коммерческим рейтингам в этом выпуске СоциоДиггера посвящен специальный текст.

Категориальные рейтинги гораздо важнее и применяются почти везде из-за характерной для науки информационной асимметрии и «сингулярности» соответствующих сущностей, в суровой экономико-маркетологической терминологии — *experience goods*, то есть товары, ценность которых постигается после приобретения в процессе их использования, а в худшем случае — *credence goods*, товары, которые толком оценить нельзя и после покупки (да простят меня ученые за базарные аналогии).

Как правило, речь идет о разбиении множеств оцениваемых объектов — журналов, ученых, публикаций, организаций, грантовых заявок и т. д. — на несколько, обычно не более пяти, уровней для принятия разнообразнейших решений. Это разбиение может само по себе быть формализованным или не очень, опирающимся на библиометрию, содержательную экспертизу или их комбинацию, запечатленным в официальных документах или витающим неформально, главное здесь — его инструментальная функция.

Есть по крайней мере несколько факторов, позволяющих уверенно прогнозировать дальнейший рост применения такого рода конструкций:

- ⇒ экспоненциальный рост числа научных результатов;
- ⇒ глобализация и повышение доступности этих результатов за счет роста доли английского языка и открытого доступа;

³ Hammarfelt B., Rushforth A. D. (2017) Indicators as Judgment Devices: An Empirical Study of Citizen Bibliometrics in Research Evaluation. *Research Evaluation*. Vol. 26. No. 3. P. 169–180. <https://doi.org/10.1093/reseval/rvx018>.

- ⇒ фокус на прозрачность, соизмеримость и предсказуемость как в управлении организациями и системами, так и в построении личной научной карьеры в условиях репутационной экономики;
- ⇒ кардинально меньшие затраты и большая масштабируемость, а также лучшая сравнимость результатов по сравнению с peer review.

Как видим, многое здесь важно безотносительно наличия чиновников. Возьмем простой пример: вы начали копать совсем новую для вас тему и ищете литературу, а находятся какие-то сотни и сотни работ, прочесть которые вы точно не успеете. Один из лучших *инструментов суждений* — но далеко не единственный! — репутация научных журналов. Издания, широко признаваемые в области как топовые (где-то по импакт-фактору, где-то иначе), скорее всего, — вспомним эмпирический закон Бредфорда⁴ и понятие «ядра», — будут содержать ведущие тексты, и концентрироваться уместно на них.

Рациональное построение репутации и академической карьеры также сложно представить без использования конвенциональных рейтингов. Скажем, гуманитарю для получения постоянной позиции в престижном западном вузе фактически необходимо иметь монографию в издательстве из первого или хотя бы второго эшелона, специалисту в области финансов или менеджмента — статью в журнале из списка Financial Times 50, биологу — статью первым автором в «святой троице» Nature/Science/Cell или хотя бы PNAS или Nature Communications. Свои общепринятые топ-5 есть и у экономистов, и у астрономов, и даже у философов из аналитического мейнстрима — и сформировались они не под гнетом внешних сил.

Те, впрочем, тоже не дремлют: одна из наиболее проработанных и успешных систем государственной оценки науки, распространенная в Скандинавии⁵, строится на разбиении всех публикаций вуза на три-четыре уровня в зависимости от уровня журнала/издательства, и сумма взвешенных соответствующим образом баллов определяет финансирование. Естественно, такая top-down схема протекает на нижние уровни и начинает применяться вузовскими администраторами для оценки сотрудников, а самими сотрудниками — для выбора каналов публикации, формализуя и дополнительно стимулируя использование описанных выше дисциплинарных конвенций — правда, часто скорее американско-глобальных, чем локальных, и скорее из области естественных, чем гуманитарных наук (но драматизировать это «скорее» не стоит: рейтинги составляются преимущественно местными учеными).

⁴ Bradford's Law. URL: https://en.wikipedia.org/wiki/bradford's_law.

⁵ Так называемая Nordic Model, применяется в Норвегии, Финляндии, Дании, Фландрии и ряде других стран, см. подробнее: <https://elsevierscience.ru/info/re-systems-overview/>. Опыт жизни ученого в такой системе излагает известный отечественный политолог Владимир Гельман, работающий в Финляндии: <https://trv-science.ru/2020/10/ne-kvartilem-edinyim/>.

Важно отношение к подобным процессам среди молодых, за которыми будущее. Один из ведущих исследователей роли метрик в науке социолог Бьорн Хаммарфельт, изучавший их восприятие учеными разных стран и возрастов, отмечает, что для молодых это восприятие связано с образом будущего себя (future self), к которому их направляют именно различные метрики, используемые инструментально. При этом если гуманитарии воспринимают использование рейтингов в виде списков «правильных» изданий подчас страдательно, что заставляет исследователей обращаться к конструкциям типа orders of worth Л. Болтански и Л. Тевено⁶, то ученые из более крупных и ресурсоемких областей типа биологии уже относятся к этому совершенно прагматично. Так, биомедик-респондент говорит о своей работе: «If I phrase this in terms of efficiency, in hours per impact factor, this is very well-paid»⁷ («Если говорить в терминах эффективности, в человеко-часах за импакт-фактор зарплата у меня очень хорошая»).

Аналогичный прагматизм не раз и не два встречался и в моей работе: так, физики из ведущего НИИ, обсуждая возможное трудоустройство в университет, оперировали собственными расчетами «цены квартиля» в долларах (речь шла о публикациях в хороших изданиях). Было очевидно, что для них самих логика «вы нам N денег, а мы вам M статей в Physical Review» очень понятна и потому приятна. Этим же, а не столько насилием государства, объясняется вполне триумфальное шествие по естественно-научным НИИ РАН основанной на журнальных рейтингах метрики «Комплексного балла публикационной результативности».

Более того, когда мы вслед за социологами начинаем говорить об «индикативном мышлении» у ученых из естественных наук, оказывается, что оно часто простирается туда, где оно не нужно и где его совсем не ждут: в peer review. Я видел немало рецензий на грантовые заявки, где прямо говорилось: хорошие статьи у человека, потому что импакт-факторы хорошие. Бывало и такое: мы обращаемся к эксперту — уважаемому иностранному профессору по общественным наукам с просьбой оценить несколько журналов в его области и специально указываем, что нам нужна содержательная оценка, так как наукометрии мы не доверяем. В ответ примерно такое письмо: «сейчас глянул метрики Scimago, вот эти три журнала хорошие, удачи вам там». Ведущих западных науковедов подобное поведение приводит в серьезное уныние: Пола Стефан сотоварищи

⁶ Haddow G., Hammarfelt B. (2019) Early Career Academics and Evaluative Metrics: Ambivalence, Resistance and Strategies. In: Cannizzo F., Osbaldiston N. (ed.) *The Social Structures of Global Academia*. P. 125–143. Abingdon, Oxon: Routledge.

⁷ Müller R., de Rijcke S. (2017) Thinking with Indicators. Exploring the Epistemic Impacts of Academic Performance Indicators in the Life Sciences. *Research Evaluation*. Vol. 26. No. 3. P. 157–168. <https://doi.org/10.1093/reseval/rvx023>.

даже написала в Nature письмо «Рецензенты зашорены наукометрией»⁸, где в том числе описывается, как рейтинги и метрики применяют панели экспертов-ученых даже там, где это прямо запрещено правилами. По сути же это говорит о том, насколько глубоко метрики проникли в ткань науки как сверху, так и снизу.

Хорошо это или плохо, непонятно. Где-то точно хорошо, где-то точно плохо. На фоне взрывного роста науки и технологий последних десятилетий можно лишь осторожно заключить, что кардинального урона сектор не несет, а возможности сопоставимого расширения реер review даже теоретически ограничены крайней ресурсоемкостью и другими проблемами.

Среди потенциально негативных аспектов упоминаются два, и они взаимосвязаны. Во-первых, это проблема *evaluation gap*, то есть разницы между критериями и инструментами внешней оценки (в том числе категориальными рейтингами) и характером и набором целей, которые ставит себе сам исследователь. Как видно, разрыв очень разный в разных дисциплинах, соответственно, где-то *goal displacement* больше, где-то меньше, но он не всегда негативен, ведь пользу можно увидеть даже в насильном переводе гуманитариев на англоязычные журналы и издательства (и это мужественно отмечают респонденты того же Хаммарфельта).

Во-вторых, это проблема *risk aversion*⁹, заставляющая рационально действующих ученых выбирать более надежные с точки зрения продуцирования нужных результатов темы и форматы работы. Для наблюдателя изнутри и извне науки картинка тут тоже может представляться разной. Так, ведущий науковед трактует нижеследующую ситуацию негативно: «a famous physicist, who spends his Saturday mornings reviewing physics papers for the journal Science, once explained to me that one must produce a steady flow of publishable papers in order to survive, so that one can do „the real thing“ (risky and adventurous research that has no guaranteed outcome) in the time that is left»¹⁰ («Знаменитый физик, обычно проводящий субботнее утро за рецензированием рукописей для Science, однажды объяснил мне, что для выживания необходимо постоянно продуцировать поток публикуемых текстов, чтобы в оставшееся время заниматься „настоящим делом“ (рискованными исследованиями без гарантированного результата)»). С точки зрения чиновников

⁸ Stephan P., Veugelers R., Wang J. (2017) Reviewers Are Blinkered by Bibliometrics. *Nature*. Vol. 544. P. 411–412. <https://doi.org/10.1038/544411a>.

⁹ Franzoni C., Stephan P. (2021) Uncertainty and Risk-Taking in Science: Meaning, Measurement and Management. *National Bureau of Economic Research*. Working Paper 28562. <https://doi.org/10.3386/w28562>.

¹⁰ Wouters P. (2017) Bridging the Evaluation Gap. *Engaging Science, Technology, and Society*. Vol. 3. No. 108–118. <https://doi.org/10.17351/ests2017.115>.

и по крайней мере части обывателей, особенно знающих о выживании больше профессора-физика на постоянной ставке, драмы тут нет.

Стоит ли продолжать воспринимать метрики и рейтинги как что-то внешнее по отношению к науке или их лучше воспринимать как более сложное явление на стыке взаимоотношений разных акторов¹¹, и прежде всего самих ученых — вопрос открытый. Ясно одно: тема для социодиггеров тут интересная, многогранная, важная — и публикуемая в высокорейтинговых журналах, чего уж там.



¹¹ Leydesdorff L., Wouters P., Bornmann L. (2016) Professional and Citizen Bibliometrics: Complementarities and Ambivalences in the Development and Use of Indicators — a State-of-the-Art Report. *Scientometrics*. Vol. 109. No. 3. P. 2129–2150. <https://doi.org/10.1007/s11192-016-2150-8>.