

Лейденский манифест для наукометрии.

Перевод А. А. Исэрова

Данные наукометрии все чаще используют не только для экспертных оценок, но и для управления научным процессом. По мере приобретения наукометрией все большего значения, все чаще разгораются дискуссии о ней. Поскольку эта проблема актуальна для всего научного сообщества, то мы публикуем перевод статьи коллектива зарубежных коллег во главе с Дайаной Хикс и Паулем Воутерсом, опубликованной в журнале « Nature » в апреле 2015 г.

Nature, April 23, 2015 (vol. 520), pp. 429–431

Лейденский манифест для наукометрии

Для управления оценкой научных исследований следует использовать десять принципов, – предупреждают Дайана Хикс, Пауль Воутерс и их коллеги

Перевод А. А. Исэрова

Точные данные все чаще используют для управления наукой. Оценка научных исследований, которые ранее была индивидуальна и выполнялась силами коллег, теперь стала рутинной и опирается на наукометрию^[1]. Проблема состоит в том, что в основу оценки сегодня положены скорее точные данные, нежели суждения. Распространилась наукометрия: обычно с хорошими намерениями, не всегда хорошо обоснованная, часто плохо применяемая. Поскольку организации все чаще оценивают научную деятельность, мы рискуем испортить систему теми же самыми инструментами, которые были созданы, чтобы ее улучшить.

До 2000 г. на CD-ROM существовал созданный Институтом научной информации (Institute for Scientific Information) Индекс научного цитирования (Science Citation Index), который использовался экспертами для специального анализа. В 2002 г. Thomson Reuters запустила интегрированную веб-платформу, сделав легко доступной базу данных Web of Science. Были созданы конкурирующие индексы цитирования: принадлежащий Elsevier'y Scopus (запущен в 2004 г.) и Google Scholar (бета-версия выпущена в 2004 г.). Возникли такие веб-инструменты, как InCites (на основе Web of Science) и SciVal (на основе Scopus), которые позволили легко сравнивать исследовательскую производительность и научное воздействие (импакт) различных научных организаций, а также программное обеспечение для сопоставления публикаций отдельных исследователей с использованием Google Scholar (Publish or Perish, выпущено в 2007 г.).

В 2005 г. физик из Университета Калифорнии в Сан-Диего Хорхе Хирш создал *h*-индекс, популяризовав подсчет цитирования индивидуальных исследователей. После 1995 г. уверенно рос интерес к импакт-факторам журналов (см. «Одержимость импакт-фактором»).

В последнее время набирает силу наукометрия, связанная с социальными сетями и онлайн-комментариями: в 2002 г. был основан F1000Prime, в 2008 г. – Mendeley, в 2011 г. – Altmetric.com (при поддержке Macmillan Science and Education, которому принадлежит Nature Publishing Group).

Как специалисты по наукометрии, обществоведы и научные управленцы, мы с растущей тревогой наблюдали широко распространившееся ошибочное применение индикаторов к оценке научной деятельности. Вот лишь некоторые из многочисленных примеров. По всему миру университеты охватила одержимость позициями в глобальных рейтингах (таких как Шанхайский рейтинг и список Times Higher Education), хотя эти списки основаны, по нашему мнению, на неточных данных и произвольных индикаторах.

Некоторые работодатели требуют от кандидатов показатели h -индекса. Есть университеты, которые основывают свои решения о карьерном продвижении сотрудников на их пороговых величинах h -индекса и на количестве статей в журналах с «высоким импакт-фактором». Исследователи, особенно в биомедицине, получили возможность хвастаться этими показателями в своих резюме. Повсюду научные руководители просят аспирантов публиковаться в журналах с высоким импакт-фактором и получают на это внешнее финансирование.

В Скандинавии и Китае некоторые университеты распределяют средства или бонусы на основе количественного показателя: к примеру, рассчитывая индивидуальные импакт-факторы, чтобы распределить «ресурсы по показателям научной деятельности» или предоставить исследователю бонус за публикацию в журнале с импакт-фактором выше 15[2].

Во многих случаях исследователи и те, кто их оценивает, все равно приходят к сбалансированному суждению. Однако злоупотребления наукометрией стали слишком распространены, чтобы не обращать на них внимания.

Таким образом, мы представляем Лейденский манифест, названный так по итогам конференции, на которой он был разработан (см. <http://sti2014.cwts.nl>). Его десять принципов не станут новостью для занимающихся наукометрией, хотя никто из нас не смог бы изложить их во всей полноте, поскольку до сего времени такая кодификация отсутствовала. Светила этой дисциплины, такие, как основатель ISI Юджин Гарфилд (Eugene Garfield) уже провозглашали некоторые из этих принципов[3]. Но их не принимают во внимание, когда специалисты по оценке научной деятельности отчитываются перед университетскими управленцами, которые не являются экспертами в соответствующей методологии. Ученые, занятые поиском литературы, чтобы с ее помощью оспорить ту или иную оценку, находят материалы в разрозненных и – с их точки зрения – неизвестных журналах, доступ к которым у них ограничен.

Мы предлагаем основные принципы в оценке исследовательской деятельности, основанной на наукометрии, с тем, чтобы ученые могли бы проверять тех, кто их оценивает, а «оценщики» могли бы проверять свои индикаторы.

Десять принципов

1. **Количественная оценка должна дополнять качественную, экспертную оценку.** Количественные измерения могут уравновесить возможное предубеждение перед экспертным рецензированием (peer review) и упростить обсуждение. Они должны усиливать экспертное рецензирование, поскольку трудно судить коллег, не владея спектром необходимых сведений. Тем не менее, специалисты, проводящие оценку научной деятельности, не должны следовать соблазну переложить принятие решений на числа. Индикаторы – не замена информированному суждению. Каждый сохраняет ответственность за свою оценку.

2. **Сопоставляйте научную деятельность с исследовательскими задачами организации, группы или ученого.** Цели исследовательской программы должны быть описаны в начале работы, и индикаторы, используемые для оценки научной деятельности, должны четко соответствовать этим целям. Выбор индикаторов и пути их использования должны принимать во внимание широкий социально-экономический и культурный контекст. У ученых разные научные задачи. Исследование, сдвигающее границы научного знания, отличается от исследования, сосредоточенного на поиске решений общественных проблем. Экспертная оценка может быть основана не только на академических идеях о научных достижениях, но и принимать во внимание достоинства, важные для политических решений, промышленности или общества. Ни одна модель оценки не применима ко всем контекстам.
3. **Отстаивайте научное качество в исследованиях, важных для того или иного региона.** Во многих частях мира высокое качество научного исследования приравнено к публикациям на английском языке. Испанское законодательство, к примеру, утверждает желательность публикаций испанских ученых в журналах с высоким импакт-фактором. Импакт-фактор рассчитывается по расположенной в США и все еще в основном англоязычной базе Web of Science. Такого рода предубеждение создает особые проблемы в общественных и гуманитарных науках, где исследования в большей степени регионально и национально обусловлены. Многие другие дисциплины также имеют национальное или региональное измерение – например, эпидемиология ВИЧ в Африке южнее Сахары.

Этот плюрализм и общественная значимость могут подавляться в пользу написания текстов, которые бы представляли интерес для «сторожей» высокого импакт-фактора – англоязычных журналов. В Web of Science широко цитируют тех испанских социологов, которые работают на абстрактных моделях или изучают данные по США. Теряются характерные черты работы тех социологов, чьи испаноязычные статьи имеют высокий импакт-фактор, с такими темами, как местное рабочее законодательство, здравоохранение для пожилых семей или занятость иммигрантов[4]. Наукометрия, основанная на высококачественной не-англоязычной литературе, поможет определить и вознаградить высокий научный уровень в исследованиях, значимых для конкретных регионов.

1. **Сохраняйте сбор данных и аналитические процессы открытыми, прозрачными и простыми.** Создание баз данных, требуемых для оценки, должно четко следовать определенным правилам, установленным до завершения оцениваемого исследования. В последние десятилетия это правило было обычной практикой среди академических и коммерческих групп, создававших методологию библиометрической оценки. Эти группы публиковали свои протоколы в рецензируемой литературе. Такая прозрачность делала возможным тщательную проверку. Например, в 2010 г. общественное обсуждение технических качеств одного важного индикатора, используемого одной из этих групп (Центром изучения науки и технологии в Лейденском университете в Нидерландах) привело к пересмотру расчета этого индикатора[5]. Коммерческие организации, которые недавно начали заниматься подобной деятельностью, должны отвечать тем же стандартам; никто не должен соглашаться с существованием «машины для оценки» в черном ящике.

Простота – достоинство для индикатора, поскольку расширяет прозрачность. Но грубо упрощенная наукометрия может исказить результаты (см. принцип 7). Эксперты, занимающейся оценкой, должны стремиться к балансу – простым индикаторам, соответствующим сложности исследовательского процесса.

1. **Позволяйте оцениваемым исследователям проверять данные и анализ.** Чтобы обеспечить качество данных, все исследователи, включенные в библиометрические подсчеты, должны получить возможность проверки верного определения их научных результатов. Каждый, кто руководит процессами оценки или занимается ими, должен обеспечивать правильность данных через самопроверку или проверку третьей стороной.

Университеты могут реализовывать этот принцип в своих информационно-исследовательских системах, и именно он должен стать основным в отборе поставщиков этих систем. Чтобы получить точные, высококачественные данные, нужны время и деньги. Отведите на это средства.

2. **Дисциплины отличаются друг от друга по практике публикаций и цитирования.** Лучше всего составлять ряд возможных индикаторов и позволять дисциплинам выбирать среди них. Несколько лет назад группа европейских историков получила относительно низкий рейтинг в оценке, поскольку они писали больше книг, а не статей для журналов, индексируемых Web of Science. Этим историкам не повезло – они работали на департаменте психологии. Историкам и обществоведам важно, чтобы в подсчет их публикаций входили книги и литература на национальных языках; специалистам по информатике нужно, чтобы учитывались доклады на конференциях.

Количество цитат зависит от дисциплины: математические журналы с наивысшим рейтингом имеют импакт-фактор около 3, а такие же журналы по клеточной биологии – около 30. Требуются нормализованные индикаторы, и наиболее убедительный метод нормализации основан на процентилях: каждый текст оценивается на основе процентиля, к которому он принадлежит в распределении цитат в своей дисциплине (к примеру, верхние 1%, 10%, 20%). Одна высокоцитируемая публикация несколько улучшает положение университета в рейтинге, основанном на перцентильных индикаторах, но может вознести университет с середины на самый верх рейтинга, разработанного на средних показателях цитирования[6].

1. **Основывайте оценку отдельных исследователей на качественной оценке их резюме.** Чем вы старше, тем выше ваш h -индекс, даже если вы больше ничего не публикуете. H -индекс отличается по дисциплинам: максимум у ученых в области наук о жизни составляет около 200, у физиков – 100, обществоведов – 20–30[7]. Он зависит от базы данных: есть исследователи, чей h -индекс составляет около 10 в Web of Science, но 20–30 в Google Scholar[8]. Чтение и оценка работы исследователя куда важнее, чем опора только на один показатель. Даже в сопоставлении больших групп ученых наилучшим подходом будет тот, где принимается во внимание больше сведений об уровне знаний, опыте, деятельности и влиянии отдельного исследователя.
2. **Избегайте неуместной конкретности и ложной точности.** Научно-технологические индикаторы подвержены концептуальной двусмысленности и неопределенности, так что требуют четких постулатов, с которыми не все соглашаются. К примеру, долго обсуждалось значение подсчетов цитат. Итак, лучше всего использовать разнообразные индикаторы, чтобы обеспечить более убедительную и плюралистичную картину. Если погрешности и ошибки можно квантифицировать, например, через показатель величины ошибки («усы»), то такие сведения должны сопровождать публикацию показателей индикаторов. Если это не возможно, то те, кто подсчитывают индикаторы, должны по крайней мере избегать ложной точности. К примеру, журнальный импакт-фактор публикуется с тремя десятичными знаками, чтобы избежать совпадения показателя. Тем не менее, если принять во внимание концептуальную двойственность и случайную изменчивость подсчетов цитат, нет смысла различать журналы на основе очень маленького различия в импакт-факторе. Избегайте ложной точности: гарантирован только один десятичный знак.
3. **Признавайте системное воздействие оценки и индикаторов.** Индикаторы меняют систему через те стимулы, которые они устанавливают. Это воздействие нужно предугадывать. Это означает, что набор индикаторов всегда предпочтителен – использование лишь одного создаст дух азартной игры и сместит цель (целью станет показатель). Например, в 1990-е гг. Австралия финансировала университетские исследования на основе формулы, в основном построенной на количестве материалов, опубликованных организацией. Университеты могут подсчитать «стоимость» материала в рецензируемом журнале; в 2000 г. она составляла 800 австралийских долларов (около 480 долларов США в том году) в финансировании исследований. Предсказуемо, что число материалов, публикуемых австралийскими учеными, выросло, но размещаться они стали в

менее цитируемых журналах, что позволяет сделать предположение о падении качества статей[9].

4. **Регулярно подвергайте индикаторы тщательной проверке и пересмотру.**

Исследовательские задачи и цели оценки меняются, и с ними развивается и исследовательская система. Когда-то полезная наукометрия становится неадекватной, но возникает новая. Системы индикаторов нужно пересматривать и порой менять. Поняв последствия своей упрощенной формулы, Австралия в 2010 г. ввела более сложный показатель – инициативу «Исследовательское превосходство для Австралии», которая делает акцент на качестве.

Следующие шаги

При соблюдении этих десяти принципов оценка исследовательской деятельности может играть важную роль в развитии науки и ее взаимодействии с обществом. Наукометрия может обеспечить ключевую информацию, которую было бы сложно собрать или понять средствами индивидуальной экспертизы. Но нельзя позволять этой количественной информации превратиться из инструмента в самоцель.

Лучшие решения принимаются, когда надежная статистика сочетается с вниманием к целям и природе исследования, которое подвергается оценке. Требуются как количественные, так и качественные данные; и те и другие по-своему объективны. Принятие решений в науке должно быть основано на высококачественных процессах, основанных на данных высочайшего качества.

Авторы:

Дайана Хикс (Diana Hicks) – профессор публичной политики, Джорджийский технологическом институте (Атланта, Джорджия, США)

Пауль Воутерс (Paul Wouters) – профессор наукометрии и директор Центра изучения науки и технологии в Лейденском университете (Нидерланды)

Лудо Валтман (Ludo Waltman) – исследователь, Центр изучения науки и технологии в Лейденском университете (Нидерланды)

Сара де Рийке (Sarah de Rijcke) – ассистент, Центр изучения науки и технологии в Лейденском университете (Нидерланды)

Исмаэль Рафолс (Ismael Rafols) – исследователь научной политики, Испанский национальный совет по исследованиям, Политехнический университет Валенсии (Испания)

diana.hicks@pubpolicy.gatech.edu

[1] Wouters, P., in *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact* (eds. Cronin, B., Sigimoto, C.), pp. 47–66 (MIT Press, 2014)

[2] Shao, J., Shen, H. *Learned Publ.* 24, 95–97 (2011).

[3] Segien, P.O. *Br. Med J.* 314, 498–502 (1997); Garfield, E. *J. Am. Med. Assoc.* 295, 90–93 (2006).

[4] López Piñeiro, C., Hicks, D. *Res. Eval.* 24, 78–89 (2015).

[5] van Raan, A.F.J., van Leeuwen, T.N., Visser, M.S., van Eck, N.J., Waltman, L. *J. Infometrics* 4, 431–435 (2010).

[6] Waltman, L. et al. *J. Am. Soc. Inf. Sci. Technol.* 63, 2419–2432 (2012).

[7] Hirsch, J.E. *Proc. Natl Acad. Sci. USA.* 102, 16659–16572 (2005).

[8] Bar-Ilan, J. *Scientometrics* 74, 257–271 (2008).

[9] Butler, L. *Res. Policy* 32, 143–155 (2003).

01.06.2015, 2447 просмотров.

Атрибуты

Файл1-тестовое свойство [Скачать файл](#)