

О СЛОЖНОСТИ ОЦЕНИВАНИЯ НАУЧНОЙ ДЕЯТЕЛЬНОСТИ

Л. КЛЕМЕНТ, М.Ч. МОСТЕРД

INCEND Ltd., Дублин, Ирландия

l.klement@incend.eu; m.mosterd@incend.eu

Библиометрия резко набрала силу после введения индекса Хирша в 2005-м году. Управленческие решения в академическом мире всё больше основываются на показателях цитируемости при оценивании научной деятельности. Такое использование данных о научных исследованиях привело к появлению неверных заключений о их значимости и даже фактов мошенничества. Сложность системы научных публикаций сделала библиометрию рычажной точкой управления благодаря наличию косвенной связи между мотивацией учёного и его научной продуктивностью. Преобладающая парадигма управления наукой заставляет учёных максимизировать количественные показатели продуктивности вместо получения и распространения новых знаний. Такая практика чревата потенциальной опасностью неудачного регулирования. Поправить положение можно, внося изменения в структуру измерений качества научных исследований. Расширение набора показателей за счёт характеристик новизны может привести к фундаментальным изменениям в академическом процессе.

Ключевые слова: библиометрия, менеджмент, оценка научной деятельности, индекс Хирша, теория сложности, системное мышление, стимулирование публикуемости, ошибочная практика управления, мошенничество, новизна.

1. ВВЕДЕНИЕ

Академический мир сильно отклонился от своей первоначальной скромной миссии: вести научные исследования. Распределение фондов на исследования, повышение статуса, внутренняя оценка научной деятельности и другие управленческие решения существенно основываются на количественных мерах академической производительности и эффективности. Хотя такие показатели могут быть полезными в совокупности с другими, возрастание роли менеджеров в научно-исследовательских учреждениях усиливает применение цифровых показателей в оценке научной деятельности. В литературе по библиометрическим показателям и их производным преобладают методы, основанные на использовании массива однородных данных. Библиометрия сосредоточена преимущественно на легко вычисляемых факторах, которые сфокусированы на максимизации отчётных показателей, а не на реальном вкладе в науку. Библиометрия и принцип 'publish or perish' («публикуйся или погибай») совместно образовали самоусиливающуюся парадигму. Обнаружилось, что недифференцированное количественное оценивание публикаций заставляет авторов стремиться увеличивать их число (van Kolf Schooten, 2012). С этим связан рост озабоченности фальсификацией результатов исследований и оказанием негативных воздействий на самих исследователей (Tijdkink et al., 2012).

Количественное измерение научной деятельности сильно затрудняется тем, что между измеряемой характеристикой, системой стимулирования и научным результатом существует сложная взаимосвязь. Вдобавок, нелинейность связи входа (характеристик исследования) и выхода (значимости научного результата) ограничивает возможность использования линейных метрик. Большинство современных библиометрических индикаторов основано на двух показателях: числе публикаций и числе цитирований. Учёт цитирований исходит из предположения, что цитирование является эффективным показателем значимости и качества цитируемого результата. Сегодня наиболее употребимым индикатором является индекс Хирша (*h-index*), названный так в честь физика Йорга Хирша, придавшего импульс библиометрии своей публикацией (Jorge E. Hirsch, 2005): «*A scientist has index h if h of his/her N_p papers have at least h citations each, and the other ($N_p - h$) papers have no more than h citations each*». Библиография по индексу Хирша и его вариантам (Cabrerizo, 2012) насчитывает 822 публикации. В 2005 году было 4 публикации о новом индексе (Ball, 2005; Bornmann and Daniel, 2005; Dume, 2005; Frangopol, 2005) и одна о его адаптации к медицинским исследованиям (Vaidya, 2005). К настоящему времени более 190 публикаций посвящено анализу достоинств и недостатков *h*-индекса, сравнению его эффективности с субъективными оцен-

ками, импакт-факторами журналов и финансовыми показателями. Эти изыскания касались в основном применения индекса в разных контекстах, анализа источников данных и проблем учёта цитирований для вычисления h -индекса. Причин огромной популярности этого индикатора много, но чаще остальных упоминаются: простота вычисления, высокая устойчивость против появления статистических «выбросов» в данных, большая работоспособность по сравнению с используемыми ранее индикаторами, соблюдение определённого баланса между учётом объёмов и качества научной продукции.

Разброс мнений об индексе Хирша лежит в диапазоне между признанием его объективной основой для управленческих решений (Costas and Bordons, 2007) и его пригодностью разве что для вычисления размера обуви автора статей (Woit, 2007). Среди наиболее существенных недостатков чаще других упоминаются искажение ранжирования ученых разных возрастов (Kelly and Jennions, 2006) и нечувствительность к многократной публикуемости статьи (Egghe, 2006). Кроме того, имеются такие концептуальные недочёты (Hirsch, 2005), как несопоставимость научных дисциплин, отсутствие учёта числа соавторов и проблемы самоцитирования. По своей природе h -индекс страдает общими провалами анализа цитирования, такими как недостаточная различимость авторов-однофамильцев, позитивных и негативных цитирований, повышенной цитируемости обзоров статей, учёт только журнальных публикаций.

Начиная с 2005 года было предложено более 70 модификаций первоначального h -индекса. В статьях обычно используется такой набор входных данных: число публикаций и цитирований, возраст автора, стандарты возрастов для данной области науки, суждения экспертов, импакт-фактор журнала и т.д. Научный вклад в этих статьях в основном фокусируется на преодолении недостатков первоначального h -индекса и других мер значимости и обычно ограничивается математическими поправками и адаптацией к конкретным отраслям науки.

Третий поток литературы по библиометрии нацелен на сравнение и противопоставление наборов показателей. Выявляются и сравниваются их достоинства и недостатки в контексте разных наук. Сравнение эффективности проводится на основе субъективных оценок и статистических методов. Вычисляются корреляции для различных академических сред. В целом, исследования обнаружили высокую корреляцию природы показателей, различаясь в их акцентах на количестве и качестве публикаций и их соотношении. Кроме того, многие публикации обсуждают недостатки статистического оценивания качества научного исследования. Более того, сравниваются выводы из библиографических подсчётов, и находятся лучшие работающие показатели по разным критериям и в разных условиях.

Среди многочисленных публикаций по библиометрии с 2005 года очень немногие рассматривают меры, сосредоточенные на содержании статей. Солер (Soler, 2007) делает особый упор на необходимости оцифровки свойств цитирования. Он предлагает индикатор научной креативности, основанный на сети цитирований в публикациях. Кабанак (Cabanac, 2011) подчёркивает преимущества использования более широкого массива общедоступных метаданных, и предлагает меру сходства между исследователями, основанную на тематических и социальных свидетельствах подобия. Другие авторы (Helbing and Baliatti, 2011) предлагают организовать накопление в базах данных экспертных оценок статей по многомерной шкале читабельности, важности, новизны, спорности.

В данной работе обсуждаются динамика процесса исследования и роль библиометрии. Системный подход иллюстрирует эту динамику и структурирует кейс, описанный Стейпелем и Полдермансом (Diederik Stapel and Don Poldermans). На основе такого анализа вводится смена парадигмы через библиометрическое рассмотрение новизны. Анализируются новые продвижения в библиометрическом оценивании, и обсуждаются последствия дальнейших шагов.

2. ДИНАМИКА ИССЛЕДОВАНИЯ

Концепция научного исследования может быть построена путём определения трёх взаимосвязанных категорий мотивов в профессии учёного: внутреннего побуждения внести вклад в своей области науки, стремления повысить свой личный статус и внешних стимулов карьерного роста, доходов, финансирования исследований. Стремление внести вклад в науку напрямую повышает качество работы исследователя, которое можно определить как обладание научного результата очевидной ясностью, что само по себе может быть одним из главных факторов качества исследования. Репутация считается высшей ценностью в учёных кругах и прямо связана с внешними мотивами. Внешние стимулы обычно определяются в контексте управления на основе количественных и качественных показателей.

Таким образом, библиометрические показатели являются важными критериями в управленческих решениях, особенно если дистанция между управленцем и исследователем велика, а процесс внутреннего рецензирования в организации слаб. Распространённость и доступность библиометрических показателей приводит к их доминированию над другими факторами оценивания, что сильно обедняет получаемые оценки научной деятельности. Такой подход к оцениванию смещает цели исследования в сторону внешних мотивов. Поэтому использование только библиометрических индикаторов опасно и может быть не связанным с качеством ис-

следования. Библиометрические показатели могут служить рычажными точками в системе публикаций. Но никакой из этих показателей не может считаться исчерпывающим индикатором объёма и качества исследования и претендовать на особую важность. Однако даже небольшие изменения в измерениях академической деятельности могут изменить альянс целей исследования, да и всю систему в целом. Это и определяет их важность в формировании более широкой системы опубликования, благодаря их взаимосвязи с целями исследования.

У исследователей, регулярно публикующихся в журналах с высоким импакт-фактором, растёт престиж и повышается конкурентоспособность в получении грантов. И наоборот, исследователям, занимающим престижные позиции в учёной среде, гораздо легче публиковаться в журналах с высоким импакт-фактором (Kerr et al., 1977). Публикация в них расширяет круг читателей, что увеличивает возможное число цитирований. Благодаря такой положительной обратной связи динамика биометрических показателей оказывается самоусиливающейся. Высокие баллы в такой метрике способствуют восприятию высокой значимости работы. Репутация исследователя увеличивает вероятность цитирования и возможности соавторства. То и другое увеличивает библиометрические показатели. Такая замкнутая сеть порождает популярность («эффект Матвея») в академических кругах и сокращает возможности молодых учёных или новых научных направлений в достижении высоких библиометрических показателей.

Важность опубликования в журналах с высоким импакт-фактором делает библиометрию ключевым детерминантом в распространении информации. Авторы предпочитают направлять статьи именно в такие журналы. А со стороны журналов решение о публикации выносится на основании потенциальной цитируемости ради повышения их импакт-фактора. И это может ограничивать распространение ценной информации. Педиатр-пульмонолог Барт Ротье сетовал в интервью (van Kolf Schooten, 2012), что его статьи отвергались широко читаемыми журналами на том основании, что они выглядят слишком узкоспециальными, чтобы иметь высокий потенциал цитирования. Подобным образом такие журналы часто не поощряют оригинальные, инновационные исследования, если они не вписываются в научные тенденции, освещаемые данным журналом (Helbing and Balietti, 2011). Более того, журналы часто не публикуют работу, если она противоречит их профилю (McKittrick, 2011). Для журналов, живущих на доходы от тиража, интересы отрасли могут иметь решающее значение. Отвергаются исследования, не интересующие или идущие вразрез с требованиями основных партнёров в отрасли. Это снижает желание учёных участвовать в таких проектах, заставляет избегать исследований

с рискованными результатами (Stephan, 2012). Это имеет отношение к часто обсуждаемому «смещению опубликования», когда отрицательные или некрупные результаты вообще не публикуются. Выражается даже тревожное мнение (Fanelli, 2011), что мета-анализ литературы ослаблен, и ресурсы тратятся на заведомо провальные исследования.

Нарушения профессиональной этики обычно случаются в условиях, когда соблазн получения дохода не сопоставляется с нормами этического поведения или механизмами социальных институтов. Механизмы, регулирующие научную работу, по большей части неформальны. Этика играет основную роль в соблюдении честности в системах соавторства и научного руководства аспирантами (van Kolfchooten, 2012). Предрасположенность к обману или нарушению профессиональной этики обычно преодолевается собственными моральными усилиями личности, однако усиливающееся стимулирование публикуемости часто упоминается как причина неэтичного поведения учёных (Anderson, 2007; Goodstein, 2002). Давление публикуемости, широко наблюдаемое в академической практике, является симптомом преобладания в среде учёных принципа «Публикуйся или погибай». Усиливающаяся конкуренция за получение грантов подстёгивает борьбу за количество публикаций.

«Пресс публикуемости», широко наблюдаемый в сфере научной деятельности, является симптомом преобладания в академическом сообществе принципа «Публикуйся или погибай». Усиливающаяся борьба за получение грантов способствует росту конкуренции в наращивании количества публикаций. Хотя пресс публикуемости увеличивает продуктивность исследователей, он снижает стандарты поведения учёных (Neill, 2008). Проведено подробное исследование влияния прессы публикуемости на медицину Голландии (Tijdink et al., 2012). Опрос 437 профессоров (из 1 100 активно ведущих научные исследования) показал: 54 % считают, что это давление стало чрезмерным; 39 % считают, что оно негативно сказывается на качестве медицинских исследований; а 26 % расценивают его влияние на медицинскую науку как «отвратительный эффект». Продолжение наблюдений за этим процессом показывает, что в среде учёных и практических врачей всё шире распространяется циничное отношение к работе о достоверности медицинской науки.

Максимизация количества публикаций – наиболее частое нарушение профессиональной этики в среде учёных. Исследование, проведённое Британским Медицинским Журналом (Tavare, 2012), обнаружило, что каждый седьмой учёный или врач в Великобритании прибегал к недобросовестному искажению или фальсификации данных ради опубликования статьи. Опрос 2 155 психологов в США (John, 2012) показал, что каждый десятый подделывал данные исследований. Более половины авторов пу-

бликовали данные только успешных экспериментов, не приводили всех измерений, продолжали опыты, пока не получится нужный результат, преподносили неожиданные результаты как планируемые, отсеивали данные, противоречащие гипотезе. Более того, 35 % давали понять, что сами сомневаются в достоверности своих выводов. В подобном исследовании (Fanelli, 2009) был осуществлён мета-анализ 18 обследований фактов мошенничества и нарушения академической этики. Суммарный результат: 1,97 % исследователей сами совершали манипуляции с данными, 72 % наблюдали сомнительные действия коллег. Guus Stelling, профессор гидродинамики, в своём интервью (van Kolf Schooten, 2012) заявил, что это не просто случаи явного мошенничества, а весьма искусные приёмы, наносящие наибольший вред. Наиболее часто используемый приём: из выводов изымаются части, содержащие наименее публикуемые результаты. Это нацелено на нарушение баланса времени, затрачиваемого на получение результатов достаточно новых и ценных для публикации, одновременно минимизируя время на выполнение исследования. Другой приём состоит в том, чтобы публиковать подобные версии одного и того же проекта в двух статьях, одну в узкоспециализированном журнале, другую – в общенаучном. Ещё чаще расширение содержания сводится к добавлению новых данных к ранее опубликованным, чтобы создать иллюзию исследования новой гипотезы. Используя такие приёмы, обычно ищут компромисс между затратой времени на выполнение проекта, тщательностью и качеством, поскольку поощряются более частые публикации и снижается риск отсутствия публикаций в случае получения отрицательных или незначительных результатов. Многочисленные авторы указывают на непоследовательность и неинновативность многих опубликованных работ (Bedeian, 1989; Denning, 1997; de Rond, 2005). Хотя все эти приёмы повышают библиометрические индексы, они в лучшем случае не делают вклада в науку, а в случае мошенничества вредят всему прогрессу науки.

Другое существенное искажение системы публикаций связано с отношением к цитируемости как к высочайшей точности характеристике качества и значимости научного исследования. Значение количества цитирований в значительной степени определяется тем, как научное сообщество воспринимает качество и ценность научного результата. В практике проявили себя шесть основных искажений этого восприятия:

Выбор журнала для публикации статьи существенно определяет читательскую аудиторию, а с ней и потенциал цитируемости. Это означает, что вероятность высокой цитируемости статьи зависит не только от её содержания, но и от того, в каком журнале она опубликована. Хотя очевидна корреляция между цитируемостью статьи и импакт-фактором журнала, сам этот факт ещё не определяет, каким будет число ссылок на данную статью.

Многие библиометрические индексы имеют накопительную природу. Это противоречит отображению значимости, поскольку значимость (impact) научного результата выясняется только со временем. Предложено несколько индексов, пытающихся компенсировать изменение значимости во времени (Egghe, 2007; Jin et al., 2007; Sidiropoulos et al., 2007; Rousseau and Ye, 2008). Эти подходы полезны, особенно в случаях мошенничества или плагиата.

Накопительные библиометрические индексы подвержены систематической ошибке, связанной с возрастом авторов. Нельзя сравнивать учёных разного академического возраста, т.е. времени первой публикации. Старшие исследователи имеют просто больше времени для накопления количества публикаций и их цитирований и уже только поэтому имеют более высокие значения индексов (Kelly and Jennions, 2007). Это создаёт барьер для молодых исследователей. Обладание низким импакт-фактором снижает их возможности публиковаться в авторитетных журналах, получать финансирование из научных фондов, продвигаться в карьере. Некоторые индексы вносят поправку на научный стаж автора (AWCR, AWCRpA, AW index inspired by Jin et al., 2007; Rousseau and Jin, 2008). Однако выдвигается возражение этому (Hirsch, 2005): учёт возраста может быть неэффективным, поскольку некоторые учёные публикуют мало статей в начале научной карьеры, прежде чем развернуться в полную силу как полноценный исследователь.

Учёт количества цитирований несёт ущерб от самоцитирования как для авторов, так и для и со стороны журналов. В некоторых случаях самоцитирование оправдано, поскольку автор естественно наращивает компетентность в своей области науки. Однако излишнее самоцитирование обесценивает библиометрические показатели автора и импакт-фактор журнала (Falagas and Alexiou, 2008). Самоцитирование завышает их (Fowler and Aksnes, 2007). Многие базы данных (например, Thomson Reuters Web of Knowledge) дают вариант, в котором самоцитирование не учитывается. Журналы тоже вовлечены в самоцитирование ради повышения своего импакт-фактора: от авторов требуют цитировать публикации из предыдущих выпусков журнала (Van Kolfshooten, 2012). Честность редакторов является единственным препятствием для такого раздувания цитирования.

Часто наблюдается взаимное предпочтение авторов при одобрительном цитировании. Это практически не влияет на показатели авторов с высоким уровнем цитирования, но существенно искажает показатели цитируемых редко.

Во многих случаях соавторство, особенно многочисленное, искажает показатели авторов. Наличие нескольких авторов, работавших над одним

проектом, в ряде отношений имеет преимущества. Это ограничивает искажения индивидуальных показателей, повышает продуктивность исследователя, способствует междисциплинарности исследований. Однако соавторство может использоваться ради увеличения индекса цитируемости автора за счёт публикации незначительного результата в составе общей статьи. Влияние соавторства на библиометрические показатели было рассмотрено в ряде работ (Batista et al., 2006; Egghe, 2008; Schreiber, 2008), где предложено раздельно подсчитывать цитируемость соавторов. Особенно существенным это оказалось при оценивании обзорных статей и статей по тем областям науки, в которых нормальным является соавторство более 40 человек.

Приведённые ограничения на анализ цитирования иллюстрируют недостаточность учёта только цитируемости для оценивания качества научного результата. Библиометрия является лишь одной из рычажных точек для управления процессом научных исследований. Она косвенно способствует ненамеренным и намеренным нарушениям профессиональной этики: от ненамеренных искажений оценок до сознательного мошенничества.

3. КОГДА НАУЧНОЕ ИССЛЕДОВАНИЕ СТАНОВИТСЯ ОБМАНОМ

В последние годы большое внимание привлекает тема обмана в научных исследованиях. Большую известность получили дела двух датских учёных, Diederik Stapel and Don Poldermans. Если социопсихолог Дидерик Стапель совершил открытое мошенничество, то кардиолог Дон Полдерманс проявил подсознательное нарушение профессиональной этики. Оба случая потрясли академическое сообщество тем, что этические и институциональные инструменты поддержания равновесия не сработали. В приведённом выше анализе упоминались инструменты на всех уровнях методологического «айсберга»: уровнях событий, структур, трендов и исходной парадигмы.

Дидерик Стапель был прославленным социопсихологом и деканом в Tilburg School of Social and Behavioural Sciences. В 2011 году он был уволен за мошенничество в научном исследовании и подделку экспериментальных данных, ошеломительно обнаруженных в исследованиях, ведущихся с 2004 года. Главными причинами своего поведения Стапель назвал (Stapel, 2012) стремление снискать славу, создав обманное впечатление о своих исследованиях. В знаменитом и широко обсуждаемом интервью он подробно разбирает, что обусловило его поступок (Anon, 2011):

«Я не устоял против давления высоких показателей публикуемости, требований постоянного повышения своего статуса. Я хотел слишком

многого слишком быстро. В системе, где контроль слаб и мало противовесов, я выбрал неверный путь».

Полагают, что масштаб мошенничества наиболее велик в международных исследованиях. Заключительный отчёт следственного комитета указывает на существование десятков фальшивых массивов данных. Стапель признан виновным в подделке данных в 55 статьях и 10 диссертациях, выполненных под его руководством. Такое возмутительное поведение нанесло огромный вред науке вообще и социальной психологии в частности. Нобелевский лауреат Дэниел Канеман подчёркивает этот вред в открытом письме психологическому сообществу (Kahneman, 2012):

«Ваша наука стала рекламным образцом объекта для сомнений в честности всех психологических исследований. Перед вами проблема не только с теми немногими, кто подверг категорическому сомнению правильность ряда основных результатов. Проблема с гораздо более многочисленными коллегами, которые ранее приняли ваши поразительные результаты за факты, когда это было опубликовано. Эти люди поставили под вопрос всю вашу науку, и теперь вы отвечаете за то, чтобы этот вопрос снять».

Стапель сформулировал инновационные цели исследования, разработал изощрённые методики эксперимента и основательную теорию. Его коллеги и аспиранты строили гипотезы и готовились к масштабным экспериментам. Стапель предложил своим сотрудникам построить таблицы для заполнения их данными предстоящих опытов по проверке гипотез. После этих тщательных приготовлений Стапель оставил за собой проведение опытов. Он сфабриковал большую часть данных и передал их коллегам под видом результатов реальных экспериментов. Переданные им данные превосходно подтверждали определённые гипотезы.

Второе дело относится к Дону Полтермансу, известному исследователю сердечно-сосудистой системы, работавшему в the Erasmus Medical Centre. Согласно данным Thomson Reuters Web of Science, он был автором 702 статей в 1993 – 2011 гг. В самом продуктивном 2006 году он опубликовал 88 статей. Это служит примером того, каких огромных результатов публикуемости можно достичь за счёт соавторства. Участие во многих исследованиях дало почти экспоненциальный рост количества цитирований, достигнувшего сегодня величины в 9 292 ссылок. Это иллюстрирует связь между количеством публикаций и вероятностью цитирования.

В мае 2011 года Полтерманс оказался в центре внимания работ по проекту Decrease VI, изучающим состояние научных исследований, в ходе которого возникли возражения против данных о пациентах. Его методы проведения работ вызвали озабоченность, и был образован следственный комитет. Были тщательно рассмотрены его методы исследований и 479

случаев соавторства. Выдвинутые обвинения включали манипулирование данными, использование недокументированных и противоречивых данных и неструктурированных рабочих процедур. Его отправили в отпуск в июле 2011 г., и продолжили расследование прошлых работ. Заключительный отчет следственного комитета от 16 ноября 2011 г. сообщает об «обнаружении серьезных недостатков в предъявленной процедуре одного исследования и погрешностей в процедуре второго». Комитет заявил также, что обнаружены факты фабрикации и неподобающего подбора данных. 17 ноября 2011 г. the Erasmus Medical Centre уволил Полтерманса за несоблюдение научной честности и нарушение профессиональной этики.

В публичном заявлении Дон Полтерманс заявил, что ему, как опытному исследователю, следовало быть более аккуратным и что он согласен с заключениями комитета (O’Riordan, 2011). Однако он подчеркнул, что его проступки были ненамеренными. Дело Дона Полтерманса возбудило в академических кругах горячие дискуссии о роли этики в научной работе и эффективности её мониторинга. Хотя обвинение в мошенничестве осталось не вполне доказанным, сама многочисленность публикаций в соавторстве вызывает сомнения в пригодности этого показателя для оценивания качества работы учёного (на уровне событий).

На уровне структур и трендов растущее давление публикуемости тоже способствует применению нечестных приёмов. Оба дела свидетельствуют о бедственном характере этого давления. Если Дон Полтерманс ссылается на неослабную максимизацию количества публикаций, то Дидерик Стапель считает (Stapel, 2011), что фабриковать данные его вынудили ориентация академического сообщества на повышение статуса и жёсткая конкуренция в борьбе за ограниченные средства. Budd et al. (2011) обнаружили, что 55 % дезавуированных статей по биомедицинской тематике созданы с нарушениями научной этики, что выросло с 40 % в 1999 г. (Budd et al., 1999). И что более тревожно, цитирование таких статей продолжается.

Выясняется, что структура, поддерживающая такое положение, возникла из-за того, что библиометрическое оценивание основано на рассогласованных показателях количества публикаций и цитирований. Внешнее стимулирование авторов производится с помощью индекса Хирша, а успешность журналов определяется через импакт-фактор. Это приводит к тому, что со стороны журналов деформируется отбор статей, а исследователи понуждаются к производству большего числа поразительных результатов. Публикация недостаточно подтверждённых выводов подталкивает некоторых исследователей к понижению требований к обоснованности представляемых результатов. Подробные экспериментальные

методики, основательные теории и признанный статус учёного часто побуждают коллег, рецензентов и другие регуляторы качества смириться с понижением требований.

Действующими силами являются глубокие человеческие мотивы, связанные с исследовательским процессом. Учёные направляют свои усилия на достижение целей карьеры, доходов, ассигнования исследований, репутации, внутреннего стремления внести вклад в науку. В условиях существующей системы стимулирования эти силы содержат и потенциал побуждения исследователя к отклонению от норм научной практики и мошенничеству. И вероятность этого возрастает, если внутренняя мотивация низка, этика слаба и регулирующие механизмы не срабатывают. Изменение системы, особенно в части способа оценивания качества научной работы, может уравновесить действующие силы и сократить позы к обману. Таким образом, парадигма может быть продвинута в сторону более гармоничного состояния, нацеленного на производство высококачественных научных результатов.

4. СДВИГ ПАРАДИГМЫ

Современные подходы к изменению оценивания качества исследований основываются на использовании возможностей Интернета и средств опубликования научных результатов. Недавно предложенные индикаторы связаны с тенденциями употребления веб-технологий в академическом мире. Всё большее число учёных обмениваются и обсуждают свои результаты в социальных сетях типа Twitter, ResearchGate, и Mendeley. Такая деятельность на этих платформах породила «альтернативные» метрики, так называемые *altmetrics*. Например, социальная сеть ResearchGate создала альтметрику «RG score», учитывающую вклад пользователя в сеть и его активность во взаимодействиях в сети.

Данные о пользователе потенциально содержат информацию о ценности его научного исследования, которую можно использовать при подсчёте показателей. Можно использовать такие факторы, как число просмотров, число скачиваний, комментарии и рейтинги на социальных страницах, аннотации, соотношение прямых и косвенных поисков. Однако использование этого комплекса данных должно вестись с учётом трёх моментов: характера поведения исследователя в сети, точности данных и произвола при задании весов факторов в итоговой оценке. Изучение поведения обнаружило тренд роста читаемости статей (Tenopir et al., 2007, 2009). При этом среднее время прочтения статьи сократилось с 48 минут в 1977 г. до 31 минуты в 2004–05 гг., по данным 898 респондентов (Tenopir et al., 2009). Поэтому число просмотров не связано с ценностью работы, а отражает лишь поисковую активность учёных при просмотре

специальной литературы и темпы появления новых результатов в данной области. Точность данных о пользователе сильно страдает от шумов в разнородных данных и от широты каналов доступа. Растущее число репозиторий (зависящее и от предметной области, и от её оргструктуры) и других средств доступа затрудняет вовлечение данных в исследование. Отсутствует стандартизация представления данных, что мешает использованию данных из разных каналов. Вдобавок, является практически невозможным применение единой метрики к разным стадиям представления результатов работы (авторская рукопись, её редакционная обработка, опубликованная версия). Наконец, далеко не просто определить пропорции вклада каждого типа данных в формуле итоговой суммарной оценки. Для построения действительно эффективной оценки качества научного исследования учёного по его библиометрическим показателям необходимы дальнейшие исследования значимости (информативности) каждого показателя. Более того, необходимо провести прозрачное для научного сообщества экспериментальное тестирование каждого индикатора, прежде чем вводить его в управленческую практику.

Несмотря на прогресс альтметрики, современные подходы не учитывают внутреннюю мотивацию ученого вносить свой вклад в науку. Динамика системы научных публикаций может существенно измениться, если добавить к учитываемым количественным показателям характеристики новизны результата.

Новизна результата в данной области может быть определена как прибавочная информация к содержанию ранее опубликованных работ. Хотя часть этого качества определяется тем, что именно будет опубликовано, остальная часть характеризуется потенциальным цитированием и признанием новизны. Это изменяет существующий баланс между старанием сократить время на получение максимума публикаций по данному исследованию и стремлением получить значимые новые научные результаты. Так можно смягчить наводнение научного мира публикациями, многие из которых порождены только целью опубликования и максимизации цитируемости. Изменение библиографических оценок мало влияет, если учёный внутренне мотивирован, но создаёт его приспособительное поведение, если внутренняя мотивация не акцентируется, а академическая этика слаба и системы профессионального контроля качества неэффективны. Важно, чтобы давление на авторов сместилось с количественных показателей на показатели инновативности. Это будет поощрять повышение качества исследований, поскольку баланс между качеством публикации и количеством публикаций станет менее прибыльным. Такое изменение парадигмы укрепит значимость внутренних эмоциональных аспектов научной работы.

Кроме пересмотра весовых приоритетов мотивов научной деятель-

ности, изменение библиографического оценивания будет иметь и другие последствия, по типу эффекта бабочки. Во-первых, интенсифицируются экспериментальные исследования, благодаря изменению соотношения риска и отдачи. Во-вторых, журналы станут отдавать предпочтение публикации инновационных работ, чтобы не прослыть ретроградами. В-третьих, несправедливое предпочтение обзорных статей статьям оригинальным будет исправлено верным учётом количества цитирований. Наконец, будет скорректировано различие в подсчёте индексов качества работы старых и молодых исследователей, и лица, принимающие управленческие решения, получат более адекватный инструмент для оценивания и сравнения.

5. ЗАКЛЮЧЕНИЕ

Мошенничество в научной среде случается редко, но нарушения профессиональной этики встречаются часто. Существующая библиометрическая система оценивания делает выгодным поведение, расходящееся с духом науки. Анализ динамики этой системы указывает на необходимость введения более целостных метрик. Действующая система не охватывает всей сложности связей процесса исследования и представления его результатов. Обратные связи в существующей системе фактически поощряют гонку за количеством публикаций, подделку экспериментальных данных, снижение требований к качеству научной работы. Негативное влияние внешних стимулов на процесс исследований наглядно проявилось в историях Дедерика Стапеля и Дона Полдерманса. Эти примеры говорят о трендах усиления пресса публикуемости и увеличения количества некачественных публикаций.

Преодоление этих трендов потребует расширения библиометрической системы, которая является рычажной точкой для управления системой опубликования научных результатов. Включение индикаторов новизны и ценности результата может усилить эту рычажную точку. Такое изменение приведёт к нелинейным эффектам во всей системе публикаций через разветвлённые взаимодействия многих её компонентов. Более того, будущая библиометрия должна иметь динамический характер, поскольку обнаружение единичного факта мошенничества может обесценить весь объём предыдущего научного исследования. Для обеспечения динамичности потребуется расширить набор измеряемых характеристик и ввести дополнительные индикаторы, что позволит отслеживать качество исследования на всех его этапах и тем самым улучшать весь процесс научной деятельности.

Структурные изменения библиометрического оценивания усилят гармонию согласования целей научной работы и приведут к сдвигу

парадигмы. Стимулирование будет направлено на повышение качества научной продукции, и этим будет препятствовать соблазну нарушить профессиональные нормы поведения. Динамические индикаторы новизны и ценности образуют петлю положительной обратной связи с главными аспектами мотивации учёного, позволят подвергать практической проверке временем эффективность управленческих решений в сфере науки. Включение оценки новизны в библиометрическое оценивание затрудняется рядом обстоятельств: слабой разработанностью самого понятия новизны и оригинальности в данной области; трудностью определения адекватной количественной меры для некоторых факторов; да и самим существованием соблазна сфабриковать данные ради получения сенсационного открытия.

В альтметрии наблюдается появление инновационных подходов к библиометрическому оцениванию. Более прозрачными должны стать показатели авторства, требуется разработать методы определения эффективности новых социальных факторов в оценке. Перспективным выглядит предложение создать средства неформального взаимного рецензирования в академических социальных сетях и на веб-страницах издателей. Однако реальная значимость этих средств будет сильно зависеть от степени активности в них. Предлагаемые подходы могут сильно повлиять на будущее. Развитие семантических технологий содержит потенциал умножения академических приложений, в том числе для измерения новизны. Это может быть выполнено путём уточнения анализа входящих и исходящих цитирований, решения проблемы положительного/отрицательного цитирования, анализа процесса генерирования тематических репликаций. Исследования в области семантических технологий могут существенно улучшить библиометрию и научный процесс в целом.

При всей важности библиометрического оценивания его эффективность зависит от прозрачности и этичности в академическом сообществе. Нормы академической этики остаются наиболее сильным регулятором процесса опубликования научных результатов. Эти нормы поддерживают правильное обращение с данными, объективность рецензирования, отсутствие нарушений в поведении профессионалов, честность издателей. Этичность может быть усилена как на институциональном, так и на личном уровне. На случаи ослабления этики и внутренней мотивации будущая библиометрия должна быть способной к адаптации, чтобы сохранять академическую порядочность.

ЛИТЕРАТУРА

1. *Anderson, M.S. et al. (2007) 'The Perverse Effects of Competition on Scientists' Work and Relationships', Science and Engineering Ethics, Vol. 13*

No. 4, pp.437–461.

2. Ball P. (2005) 'Index aims for fair ranking of scientists', *Nature*, Vol. 436, No. 7053. P. 900.

3. Batista P.D., Campitelli M.G., Kinouchi O., Martinez A.S. (2006) 'Is it possible to compare researchers with different scientific interests?', *Scientometrics*, Vol. 68, No. 1, pp.179–189.

4. Bedeian A.G. (1989) 'Totems and taboos: Undercurrents in the management discipline', *Academy of Management News*, Vol. 19, No. 4, pp.1–6.

5. Bornmann L., Daniel H.D. (2005) 'Does the h-index for ranking of scientists really work?' *Scientometrics*, Vol. 65, No. 3, pp.391–392.

6. Budd J.M., Sievert M., Schultz T.M., Scoville C. (1999) 'Effects of article retraction on citation and practice in medicine', *Bulletin of the Medical Library Association*, Vol. 87, No. 4, p.437.

7. Budd J.M., Coble Z.C., Anderson K.M. (2011) 'Retracted publications in biomedicine: cause for concern', pp.390–395.

8. Cabanac G. (2011) 'Accuracy of inter-researcher similarity measures based on topical social clues', *Scientometrics*, Vol. 87, No. 3, pp.597–620.

9. Cabrerizo F.J. (2012) 'h-index Bibliography' [online] <http://sci2s.ugr.es/hindex/biblio.php> (Accessed 19 October 2012).

10. Costas R., Bordons M. (2007) 'The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level', *Journal of Informetrics*, Vol. 1, No. 3, pp.193–203.

11. De Rond M. (2005) 'Publish or Perish: Bane or Boon of Academic Life?', *Journal of Management Inquiry*, Vol. 14, No. 4, pp.321–329.

12. Denning P.J. (1997) 'A new social contract for research', *Communications of the ACM*, Vol. 40, No. 2, pp.132–134.

13. Dume B. (2005) 'How high is your h-index?', *Physics World*, Vol. 18, No. 9, pp.7.

14. Egghe L. (2006) 'Theory and practise of the g-index', *Scientometrics*, Vol. 69, No. 1, pp.131–152.

15. Egghe L. (2007) 'Dynamic h-index: The Hirsch index in function of time', *Journal of the American Society for Information Science and Technology*, Vol. 58, No. 3, pp.452–454.

16. Egghe L. (2008) 'Mathematical theory of the h- and g-index in case of fractional counting of authorship', *Journal of the American Society for Information Science and Technology*, Vol. 59, No. 10, pp.1608–1616.

17. Epstein D. (2007) 'Impact factor manipulation', *The Journal of the European Medical Writers Association*, Vol. 16, No. 3, pp.133–134.

18. Falagas M.E., Alexiou V.G. (2008) 'The top-ten in journal impact factor manipulation', *Archivum Immunologiae et Therapiae Experimenta-*

lis, Vol. 56, No. 4, pp.223–226.

19. Fanelli D. (2009) 'How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data T. Tregenza, ed. *PLoS One*, 4(5) [online] <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0005738> (Accessed 9 August 2012).

20. Fanelli D. (2011) 'Negative results are disappearing from most disciplines and countries', *Scientometrics*, Vol. 90, No. 3, pp.891–904.

21. Fowler J.H., Aksnes D.W. (2007) 'Does self-citation pay?' *Scientometrics*, Vol. 72, No. 3, pp.427–437.

22. Frangopol P.T. (2005) 'The Hirsch index - a new scientometric indicator for the evaluation of the results of a scientific researcher', *Revista de Chimie*, Vol. 56, No. 12 pp.1279–1281.

23. Goodstein D. (2002) 'Scientific misconduct', *Academe*, Vol. 88, No. 1, pp.28–31.

24. Helbing D., Baliotti S. (2011) 'How to create an innovation accelerator', *European Physical Journal-Special Topics*, Vol. 195, No. 1, pp.101–136.

25. Hirsch J.E. (2005) 'An index to quantify an individual's scientific research output', *Proceedings of the National Academy of Sciences*, Vol. 102, No. 2005, pp.16569–16572.

26. Jin B.H., Liang L.M., Rousseau R., Egghe L. (2007) 'The R- and AR-indices: Complementing the h-index', *Chinese Science Bulletin*, Vol. 52, No. 6, pp.855–863.

27. John L.K., Loewenstein G., Prelec D. (2012) 'Measuring the prevalence of questionable research practices with incentives for truth telling', *Psychological Science*, Vol. 23, No. 5, pp.524–532.

28. Kahneman D. (2012) 'A proposal to deal with questions about priming effects' [online] http://www.nature.com/polopoly_fs/7.6716.1349271308!/suppinfoFile/Kahneman%20Letter.pdf (Accessed 3 February 2013)

29. Kelly C.D., Jennions M.D. (2007) 'h-index: age and sex make it unreliable', *Nature*, Vol. 449, No. 7161, pp.403–403.

30. Kerr S., Tolliver J., Petree D. (1977) 'Manuscript characteristics which influence acceptance for management and social science journals', *Academy of Management Journal*, Vol. 20, No. 1, pp.132–41.

31. McKittrick R. (2011) 'Bias in the Peer Review Process: A Cautionary and Personal Account', in Michaels, P.J. (Ed.), *Climate Coup*, Cato Institute, Washington DC, pp.71–98.

32. O'Riordan M. (2011) 'Erasmus MC fires Poldermans; ESC reviews his work' *theheart.org* [online] 23 November. <http://www.theheart.org/article/1315171.do> (26 July 2012).

33. Rousseau R., Jin B. (2008) 'The Age-Dependent h-Type AR(2)-Index: Basic Properties and a Case Study', *Journal of the American Society for*

Information Science and Technology, Vol. 59, No. 14, pp.2305–2311.

34. Rousseau R., Ye F.Y. (2008) 'A proposal for a dynamic h-type index', *Journal of the American Society for Information Science and Technology*, Vol. 59, No. 11, pp.1853–1855.

35. Schreiber M. (2008) 'To share the fame in a fair way, h(m) modifies h for multi-authored manuscripts', *New Journal of Physics*, Vol. 10, pp.1–9.

36. Sidiropoulos A., Katsaros D., Manolopoulos Y. (2007) 'Generalized Hirsch h-index for disclosing latent facts in citation networks', *Scientometrics*, Vol. 72, No. 2, pp.253–280.

37. Soler J.M. (2007) 'A rational indicator of scientific creativity', *Journal of Informetrics*, Vol. 1, No. 2, pp.123–130.

38. Stapel D.A. (2011) 'Diederik Stapel: 'De druk is mij te veel geworden'' *Volkskrant* [online] 31 October. http://www.volkskrant.nl/vk/nl/3184/opinie/article/detail/3004921/2011/10/31/Diederik-Stapel-De-druk-is-mij-te-veel-geworden.dhtml?utm_source=scherm1&utm_medium=button&utm_campaign=Cookiecheck (Accessed 29 July 2012)

39. Stapel D.A. (2012) *Ontsporing*, Uitgeverij Prometheus, Amsterdam.

40. Stephan P. (2012) *How Economics Shapes Science*, Harvard University Press, Cambridge.

41. Tavaré A. (2012) 'Where are we with transparency over performance of doctors and institutions?' *British Medical Journal*, Vol. 345, No. e4464.

42. Tjldink J., Vergouwen A.C.M., Smulders Y. 'Publication culture and burn out among Dutch medical professors', <http://www.socialsciences.leiden.edu/cwts/news/cwts-seminar-20120127.html> (Accessed 26 July 2012).

43. Vaidya J.S. (2005) 'V-index: A fairer index to quantify an individual's research output capacity', *British Medical Journal* [online] <http://www.bmj.com/rapid-response/2011/10/31/v-index-fairer-index-quantify-individual-s-research-output-capacity> (Accessed 15 October 2012).

44. van Kolf Schooten F. (2012) *Ontspoorde wetenschap*, Uitgeverij De Kring, Amsterdam.

45. Woit P. (2007) 'My H-index: 33' *A Quantum Diaries Survivor* [online] 11 April. <http://dorigo.wordpress.com/2007/04/11/my-h-index-33/> (Accessed 27 July 2012).