

**СОДЕРЖАНИЕ**

<b>Ларивьер В., Лозано Ж.А., Джингра И.</b> Приходят ли элитные журналы в упадок?	3
<b>Лозано Ж.А.</b> Этика использования услуг по редактированию языка в эру цифровой коммуникации и статей с очень большим числом авторов	11
<b>Мартин Т., Болл Б., Каррер Б., Ньюман М.Е. Дж.</b> Соавторство и цитирование в научном издательстве	21
<b>Мискевич Я.</b> Влияние публикаций в трудах на измерение ядра соавторов	31

## РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Академик РАН **Ю.М. Арский** (Российская Федерация) — *главный редактор*, ВИНТИ РАН, 125190, Москва, ул. Усиевича, 20. Телекс 411249

Проф. д-р. **Р.С. Гиляревский** (Российская Федерация) — *заместитель главного редактора*, ВИНТИ РАН, 125190, Москва, ул. Усиевича, 20. Телекс 411249

**С. Дж. Паркер** (Канада) — *заместитель главного редактора*, IDCR, P.O. Box 8500, Ottawa, Ontario K1G 3H9, Canada

**А. Джилкрайст** (Великобритания) — CURA Consortium and GAVEL g.e.i.e, 38 Ship Street, Brighton BN1 1AB, UK

**М. Дрейк** (США) — Технологический институт шт. Джорджия, Библиотечный и информационный центр, 704 Cherry Street, Atlanta, Georgia 30332-0900, USA

**Н. Дюсулье** (Франция) — 192 Chemin du Cannet, 06220 Vallauris, France

**А. де Кемп** (Германия) — Издательство “Springer-Verlag”, Postfach 10 52 80, D-69042 Heidelberg, Germany

Д-р **Т. Кеннон** (Великобритания) — Отдел исследований и разработок Британской библиотеки, 2 Sheraton Street, London W1V 4BH, UK

**М. Миддлтон** (Австралия) — Школа информационных систем, QUT Gardens Point Campus, 2 George Street, Brisbane, 4000 QLD., Australia

**Т. Молвиг** (Норвегия) — Национальное управление по научной информации, вузовским и специальным библиотекам, P.O. Box 2439 Solli, N-0201, Oslo, Norway

**Х. Ринкон Феррейра** (Бразилия) — Бразильский институт информации по науке и технике (IBICT), SAS— Quadra 5, Lote 06, Bloco H, 700-70-000 Brasilia D.F., Brazil

**С. Феррейро** (Чили) — Чилийский университет, Системы информационных и библиотечных служб, Casilla de Correo 10D, Santiago, Chile

Проф. **Ю. Фудзивара** (Япония) — Университет Цукуба, Институт электроники и информатики, Tsukuba-shu, Ibaraki, 305 Japan

Д-р **М. Хименес** (Испания) — Испанское общество по научной документации и информации, Fuencarral, 123-6º dcha., 28010, Madrid, Spain

Д-р **А.И. Чёрный** (Российская Федерация) — ВИНТИ РАН, 125190, Москва, ул. Усиевича, 20. Телекс 411249

## Приходят ли элитные журналы в упадок?\*

**Винсент ЛАРИВЬЕР**  
(Vincent LARIVIÈRE)

Школа библиотекведения  
и информатики, Монреальский универси-  
тет, г. Монреаль, Канада

**Жорж А. ЛОЗАНО**  
(George A. LOZANO)

Эстонский центр эволюционной  
экологии, г. Тарту, Эстония

**Ив ДЖИНГРА**  
(Yves GINGRAS)

Обсерватория наук и технологий,  
Межуниверситетский центр исследований  
по науке и технологии, Университет  
Квебека в Монреале, г. Монреаль, Канада

*Предыдущие исследования показывают, что за период более 20 прошедших лет работы самого высокого качества публикуются в постоянно растущей, разнообразной и большой группе журналов. В этой статье исследуется, влияет ли это многообразие также на небольшое число элитных журналов, которые традиционно считаются самыми лучшими. Мы исследуем модели цитирования за последние 40 лет из 7 продолжающих традиционно выходить элитных журналов и 6 журналов, которые приобрели значимость за последние 20 лет. Чтобы оказаться среди верхних 5% или 1% наиболее цитируемых статей, сегодня статьям необходимо получать почти в два раза больше ссылок, чем 40 лет назад. Начиная с конца 1980-х и до начала 1990-х гг. элитные журналы публикуют снижающуюся долю таких верхних цитируемых статей. Это также относится к двум журналам, «Science» and «Nature», которые традиционно считаются ведущими журналами и часто используются как библиометрические показатели «превосходства». С другой стороны, несколько новых и утвердившихся журналов публикуют растущую долю самых цитируемых статей. Эти изменения приносят новые проблемы и возможности для всех сторон. Журналы могут определять стратегии роста или поддержки своей относительной позиции в журнальной иерархии. Сегодня ученые имеют выбор публиковаться в большем разнообразии журналов, зная, что их работа может все ещё достичь одинаковой аудитории. Наконец, экспертам по оценке и администраторам необходимо знать, что хотя всегда будет иметь место определенный престиж, связанный с публикацией в «элитных» журналах, иерархии журналов находятся в постоянном движении, поэтому включение журналов в эту группу не является долговременным.*

### ВВЕДЕНИЕ

Начиная, по крайней мере, с середины XIX в., научные идеи и открытия распространяются и обсуждаются первоначально через статьи в журналах [1-2]. До недавнего времени эти «работы» состояли буквально из статей, ограниченных рамками номеров и томов, физически распространяемых на регулярной основе подписчикам и учреждениям. Начиная с 1990-х гг., с помощью возможностей Интернета эта старая система начала вытесняться электронным распространением отдельных статей. В этом новом цифровом веке доступ к научным статьям радикально трансформируется посредством сетевых репозиториев препринтов (таких как ArXiv), учрежденческих и личных хранилищ принятых

и опубликованных статей, электронного доступа печатных журналов, только электронных журналов и журналов открытого доступа.

Расселл и Руссо [3] отметили, что Интернет сдвигает акцент от журнала к отдельной статье, и предположили, что эта тенденция может снизить важность импакт фактора. Учитывая, что в цифровом веке ученые имеют доступ к каждой статье отдельно без просмотра целого номера или тома журнала, мы предсказали, что соотношения между импакт фактором и ссылками на статьи будут ослабляться [4]. Мы протестировали эту гипотезу, используя массив данных из более 29 млн. статей и 800 млн. ссылок, и показали, что за период 1902 – 1990 гг. сила соотношения между импакт фактором и ссылками на статьи растет, но, как предсказано, разнообразие долей ссылок в статьях относительно соответствующего импакт фактора их журнала постоянно увеличивается после 1991 г. Кроме того, до 1990 г. соотношение пе-

\* Перевод Larivière V., Lozano G.A., Gingras A. Are elite journals declining? —2013.—  
<http://arxiv.org/ftp/papers/1304/1304.6460.pdf>

редовых (т.е. наиболее цитируемых) статей, опубликованных в ведущих (т.е. имеющих самый высокий импакт фактор) журналах, увеличивается. Однако с 1991 г. эта модель меняется, и мы наблюдаем спад в соотношении верхних цитируемых статей в журналах с самым высоким импакт фактором. Следовательно, самая важная литература публикуется сегодня в постоянно растущих различных источниках [4].

Однако эффект, который мы подтвердили документами на макроуровне, не может применяться к небольшому количеству «элитных» журналов, которые продвигают себя и должны восприниматься как самые важные. Несмотря на изменения в издательской индустрии за последние 20 лет, эта небольшая группа элитных журналов смогла поддерживать публикационную деятельность на уровне стабильного или даже постоянно растущего соотношения наиболее цитируемых статей. Здесь мы сравниваем доли ссылок за 20-летний период до и после 1990 г., произвольно взятого за отсчет времени, когда статьи начали распространяться в цифровом формате. Используя соотношение наиболее цитируемых статей отдельными журналами [5-7], мы рассматриваем модели, найденные в макромасштабе, также встречающиеся в избранном числе элитных журналов, которые затем сравниваем с некоторыми журналами, которые становятся в возрастающей степени важными.

## МЕТОДЫ

Мы использовали ресурс Web of Science издательства Thomson Reuters (включая все стандартные индексы цитирования, SCIE, SSCI и AHCI) с 1970 по 2010 гг. (1970-2012 гг. для полученных ссылок) с общим числом более 27, 8 млн. статей и 784 млн. ссылок.\* Эта база данных основана на источнике данных издательства Thomson Reuters, преобразованном в реляционную базу данных SQL, разработанную для библиометрических анализов. Чтобы иметь со временем совместимые измерения, мы использовали фиксированный интервал цитирования в два года после публикации, к которому добавили ссылки, полученные в течение года издания. Мы также собрали эти данные, используя пятилетний интервал ссылок, следующих за годом публикации, и результаты были почти одинаковыми. Однако пятилетний интервал ссылок будет пренебрегать последними 6 годами из данных, поэтому представлены только двухлетние данные ссылок и анализа. Мы использовали два порога ссылок: верхние 5% наиболее цитируемых статей и верхний 1% наиболее цитируемых. В обоих случаях полученные ссылки включали самоцитирование. Учитывая, что соотношение самоцитирований снижается при наличии большого числа ссылок, самоцитирования не являются фактором при взгляде на верхние цитируемые статьи [8]. Включенными были статьи, заметки и обзоры, а исключались редакционные статьи, так как это типичный случай в библиометрическом анализе [9]. Ссылки, которые они получали, могли происходить из документа любого типа.

Подсчет происходил в три этапа: 1) мы подсчитывали число ссылок, полученных каждой статьей; 2) мы ранжировали эти статьи и отбирали 1% и 5% верхних цитируемых статей за каждый год; 3) мы подсчитывали соотношение этих статей, опубликованных в выбранном массиве журналов. Если это соотношение снижа-

лось для данного журнала, это означало, что его доля верхних цитируемых статей снижается, и наоборот. Важно отличать этот показатель, который говорит нам, какие журналы получают большую часть верхних цитируемых статей ежегодно, от другого показателя, основанного на соотношении ссылок, полученных за годы, к статьям, опубликованным в этих журналах за определенный год. Хотя последний показатель использовался для измерения формы концентрации ссылок со временем [10], он не является надежной мерой, так как подвержен влиянию полураспада ссылок, полученных статьями. Очевидно, что высокоцитируемые статьи, публикуемые в высококачественных журналах, получают свою долю ссылок в рамках специальных журналов, в которых они были опубликованы, накопленных за годы, следующие после их публикации. Что мы хотим измерить, так это то, меняется или нет с годами доля статей, происходящая из элитных журналов и цитируемая в верхних 1% или 5% наиболее цитируемых статей.

Мы провели этот анализ для двух групп журналов, которые будут рассматриваться как «элитные» журналы и «развивающиеся» журналы. «Элитные» журналы были выбраны из самых престижных журналов, которые также имели самый высокий импакт фактор в 2011 г. (среди верхнего 1%). Кроме того, эти журналы не являются обзорными журналами, публикуют большое число статей ежегодно и существуют несколько десятилетий. Эта группа включает три журнала общего характера (*Nature*, *Science* и *Proceedings of the National Academy of Science* - PNAS), четыре медицинских/биомедицинских журнала (*Cell*, *Lancet*, *New England Journal of Medicine* - NEJM и *Journal of the American Medical Association* - JAMA). Журнал PNAS имеет чуть более низкий импакт фактор, чем другие журналы, но был, тем не менее, включен из-за своей репутации и того факта, что в 1980-х гг. он опубликовал самую высокую долю высокоцитируемых статей. С другой стороны, «развивающиеся» журналы были выбраны из тех, которые имели самый высокий рост в своей доле верхних цитируемых статей за последние 40 лет. Эта группа включала один журнал общего характера (*PLoS One*), три журнала по материаловедению/нанотехнологии (*Nano Letters*, *Advanced Materials* и *Nature Materials*), один по медицине (*Journal of Clinical Oncology*) и один по химии (*Chemical Reviews*). *One*, *Nature Materials* и *Nano Letters* являются относительно новыми, созданными в этом веке, а другие три – более старыми. *Chemical Reviews* был основан в 1925 г., а журналы *Journal of Clinical Oncology* и *Advanced Materials* начали издаваться в 1980-х гг.

На рис. 1 представлены изменение доли всех статей, включенных в верхние 1% и 5% наиболее цитируемых (часть А), и изменение порога ссылок (часть Б). Порог ссылок для включения статей либо в 1% или 5% категорию является целым числом, но отдельные статьи могут иметь это число ссылок, поэтому доля включенных статей всегда выше 1% или 5%. Ссылки еще разбросаны в вершину хвоста распределения ссылок, поэтому для верхнего 1% статей число включенных статей находится на самом деле ближе к действительному 1% статей. Число ссылок, необходимых для включения в число верхних цитируемых статей, строго растет (рис. 1Б). Чтобы быть включенными в верхние 5% или 1% наиболее цитируемых статей, в 2010 г. статьям необходимо иметь почти в два раза больше ссылок, чем 30 или 40 лет назад.

\* Хотя ссылки, сделанные из SSCI и AHCI, журналами включаются, ни один журнал из этих двух баз данных не удовлетворял критериям рассмотрения их как элитных журналов.

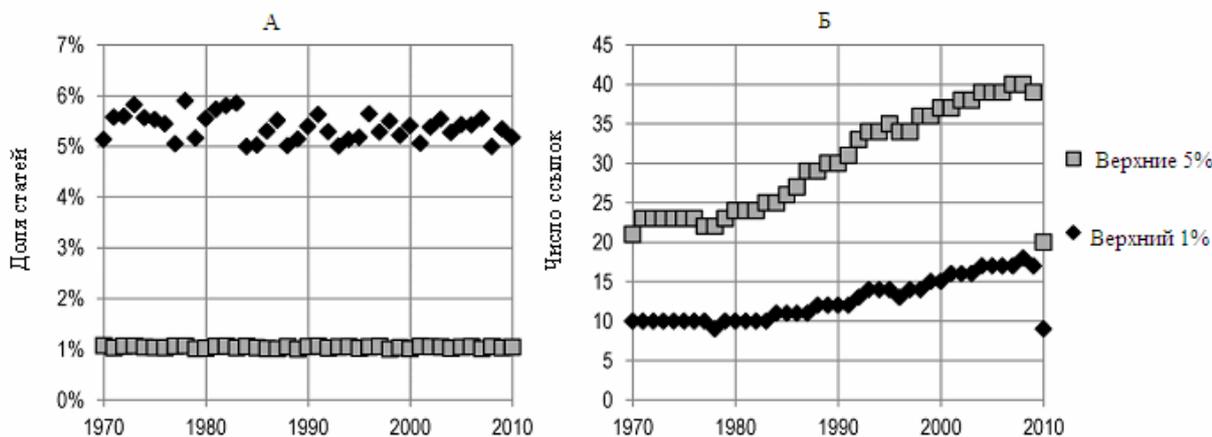


Рис. 1 Часть А. Соотношение статей в верхних 5% и 1% наиболее цитируемых статей. Часть Б. Порог ссылок для включения в верхние 5% и 1% наиболее цитируемых статей, 1970-2010 гг.

## РЕЗУЛЬТАТЫ

Все семь элитных журналов публикуют сегодня меньшую долю верхних цитируемых статей, чем 20-25 лет назад (рис. 2). Во всех случаях эти элитные журналы публикуют большую долю верхних 1% наиболее цитируемых статей, чем долю верхних 5% наиболее цитируемых статей. Эта доля высокоцитируемых статей в трех журналах общего характера журналах (*Science*, *Nature* и *PNAS*) и журнале *Cell* возрастала вплоть до 1990 г. и снижалась после этого. В середине 1980-х гг. журналом *PNAS* было опубликовано почти 9% из верхнего 1% наиболее цитируемых статей и свыше 4% из верхних 5% наиболее цитируемых статей. К 2010 г. эти доли значительно упали до 2,7% и 2,2 %, соответственно. Подобным образом, в своем пике в конце 1980-х гг. и в начале 1990-х гг. журналы *Nature* и *Science* опубликовали около 7% и 6% из верхнего 1% наиболее цитируемых статей, соответственно. Эти доли сегодня составляют около 4% и 3%, соответственно. Похожая тенденция наблюдается для верхних 5% наиболее цитируемых статей, за исключением того, что рост в первые 20 лет был очевиден только для *PNAS* и *Cell*. Среди биомедицинских журналов *Lancet* и *NEJM* следуют чуть отличной модели при рассмотрении верхнего 1% самых цитируемых статей. Доля *NEJM* постоянно снижалась до 1980-х гг., а доля *Lancet* оставалась относительно стабильной. Тем не менее, оба журнала снизили свою долю верхних 5% наиболее цитируемых статей на протяжении всего этого периода с около 1% в 1970-х гг. до почти 0,5% в 2010 г. Наконец, *JAMA* увеличивал свою долю верхних 5% наиболее цитируемых до второй половины 1990-х гг. и верхнего 1% наиболее цитируемых статей вплоть до начала 2000-х гг., и медленно утратил свою позицию после этого.

Развивающиеся журналы были отобраны, поскольку они увеличивали свою долю высокоцитируемых статей (рис. 3). Из трех новых журналов *PLoS One*, междисциплинарный журнал открытого доступа, был основан в 2006 г., в настоящее время он насчитывает 0,6 % от верхнего 1% наиболее цитируемых статей и 0,8% от верхних 5% наиболее цитируемых статей. Подобным образом *Nature Materials* и *Nano Letters* в настоящее время каждый имеют свыше 0,5% и 1% от верхнего 1% наиболее цитируемых статей, соответственно. Трём более ста-

рым журналам *Chemical Reviews*, *Advanced Materials* и *Journal of Clinical Oncology* удалось увеличить свое соотношение верхних цитируемых статей с начала 1980-х гг.

Общее число опубликованных статей увеличивалось в течение последних нескольких десятилетий, и в большинстве случаев число ежегодно опубликованных отдельными журналами статей также росло. Чтобы оценить, относились ли к общему числу опубликованных статей документально подтвержденные выше модели просто в силу роста числа статей, изданных каждым журналом, мы вычислили ежегодный нормализованный верхний 1% показатель для каждого журнала. Этот нормализованный верхний показатель является измерением числа верхних статей (1%, 5% и т.д., здесь мы используем 1% порог), опубликованных в журнале в определенном году, соотносимого с тем, что мы должны были бы ожидать, если бы верхние статьи были случайно распределены по всем журналам. Конкретно, за определенный год он представляется относительным числом верхних статей в определенном журнале (верхние статьи/все статьи), деленным на долю верхних статей, опубликованных во всех журналах. Коэффициент 1 будет показывать, что число верхних статей, опубликованных журналом, является тем, что ожидается случайно. Коэффициент 10 показывает, что журнал опубликовал в 10 раз больше верхних статей, чем ожидалось, чисто случайно. Например, если один миллион статей опубликован в определенном году, то 10 тыс. будет верхних 1%, и если определенный журнал публикует 500 статей в том же году, и 25 находятся среди этих 10 тыс., тогда его нормализованный верхний 1% показатель будет равен 5.

Элитные журналы публикуют гораздо больше верхних статей, чем возможно ожидается, но встречаются различные модели (рис. 4). В первые 20 лет журнал *Cell* переходил от примерно 30 к 70 (значения нормализованно 1% показателя), и сегодня он снова вернулся почти к 50. С 1970 по 1990 гг. журналы *JAMA* и *Lancet* были в однозначных цифрах и низких числах от 13 до 19, соответственно, и затем резко выросли до диапазона от 40 до 50. Журнал *PNAS* держался устойчиво в нижней отметке 20 примерно до 1990 г. и с тех пор медленно снижался почти к 10. Наконец, в последние 20 лет жур-

налы *Nature*, *Science* и *NEJM* медленно росли, но все еще более или менее находятся в том же диапазоне, от приблизительно 40 до почти 50. Интересно, что *Nature* и *Science* имели свой самый устойчивый рост в 1970–1980-х гг. от 10 примерно до 40. Эти модели необязательно отличаются по предметным областям.

Развивающиеся журналы также публикуют более высокий процент верхних 1% наиболее цитируемых статей, чем, возможно, ожидается, учитывая число статей, ими публикуемых (рис. 4). Журналы *Chemical Reviews* и

*Nature Materials* сегодня находятся в диапазоне 40 и 50. Странно, но с 1970 по 1990 гг. *Chemical Reviews* был сильно изменчивым. Журналы *Journal of Clinical Oncology*, *Advanced Materials* и *Nano Letters* сегодня приближаются к 20. Что касается журнала *PLoS One*, этот коэффициент остается в низких однозначных цифрах и только приближается к 1,09 в 2010 г., что показывает, что его рост в доле верхних цитируемых статей (рис. 3) происходит в значительной степени благодаря очень высокому числу публикуемых им статей.

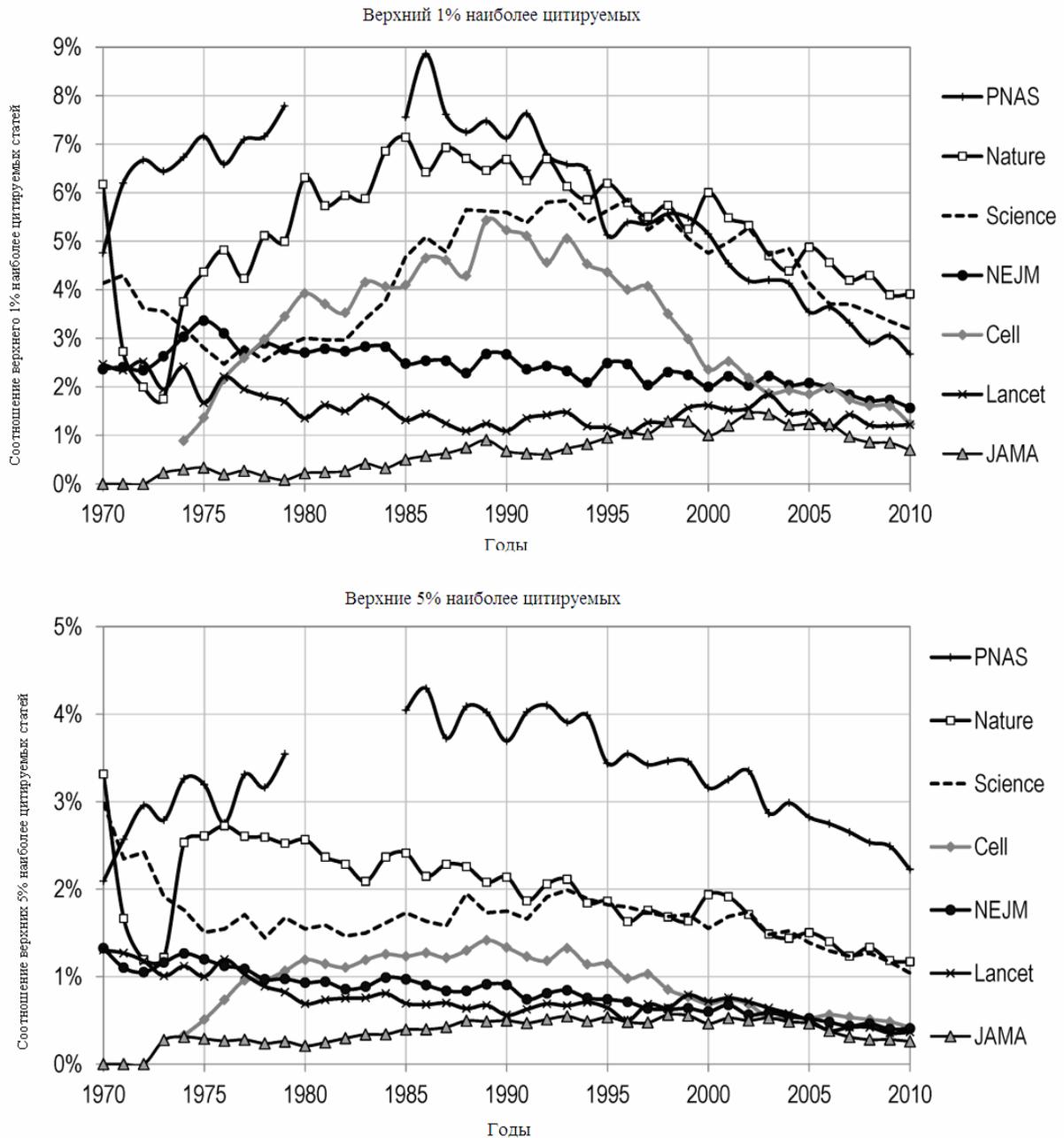


Рис. 2. Соотношение верхних 1% и 5% наиболее цитируемых статей в элитных журналах, 1970-2010 гг. Данные по журналу *PNAS* не были доступны с 1980 по 1984 гг.

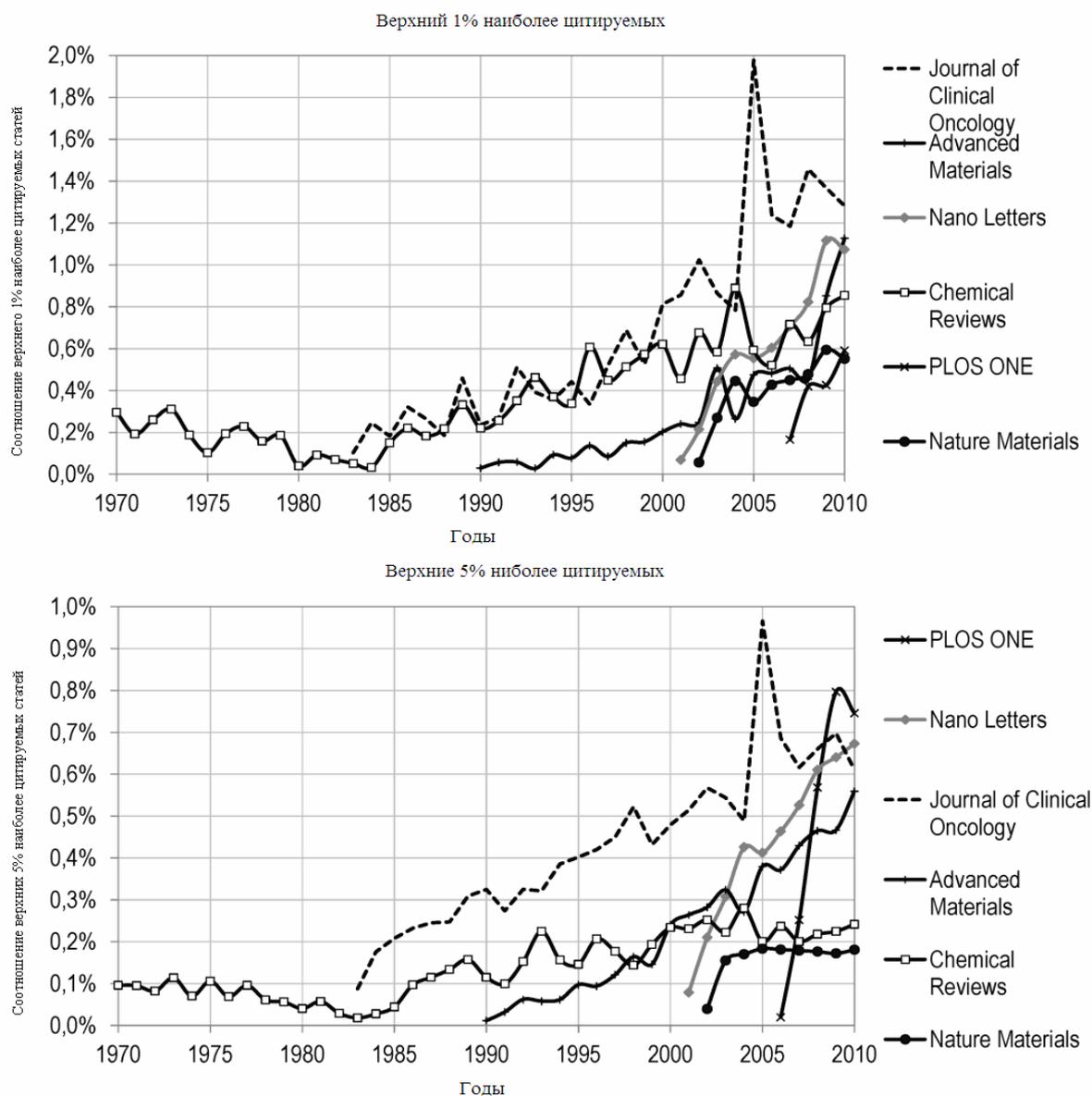


Рис. 3. Соотношение верхних 1% и 5% наиболее цитируемых статей для развивающихся журналов, 1970-2010 гг.

## ОБСУЖДЕНИЕ И ЗАКЛЮЧЕНИЕ

Ранее в крупномасштабном анализе мы документально подтвердили, что доля самых цитируемых статей, опубликованных в журналах с самым высоким импакт фактором, неизменно снижалась с появлением цифрового века [4]. Здесь мы проверяли, справедлива ли эта модель для небольшого числа элитных журналов, которые, как правило, считаются наиболее важными. Наш первоначальный анализ показал, что понятие «высокоцитируемые» меняется. Сегодня публикуется больше статей, и эти статьи имеют более длинные списки литературы, поэтому статьи получают больше ссылок и меньше статей остается вообще не цитируемыми [11]. Следовательно, чтобы быть среди верхнего  $n$ -го перцентиля, сегодня статьям необходимо иметь почти в два раза больше ссылок, чем 30-40 лет назад. Эффект этого вздутия является таким, что ссылки обесцениваются почти наполовину за короткий промежуток времени, т.е. в диапазоне научной карьеры, что может воздействовать на интерпретацию измерений влияния

на основе ссылок на протяжении карьеры отдельного ученого.

С конца 1980-х гг. и начала 1990-х гг. несколько новых и ряд некоторых существующих длительное время журналов становятся более важными, тогда как традиционные элитные журналы, включая *Science* и *Nature*, публикуют снижающуюся долю верхних цитируемых статей. Однако, даже если их доля самых цитируемых статей снижается, элитные журналы все еще «берутся за непосильную работу» и публикуют большую долю верхних цитируемых статей, чем можно было бы ожидать, от их общего числа статей. Это относительное снижение роли элитных журналов по сравнению со всеми журналами заключается фактически в общей тенденции деконцентрации научной деятельности. На уровне стран, как и внутри стран, производство статей больше и больше изменяется географически и, таким образом, меньше концентрируется в небольшом числе ведущих стран или городов [12]. С точки зрения полученных ссылок это распределение сегодня является менее концентрированным, чем в прошлом, так как много статей получают свою долю ссылок [13].

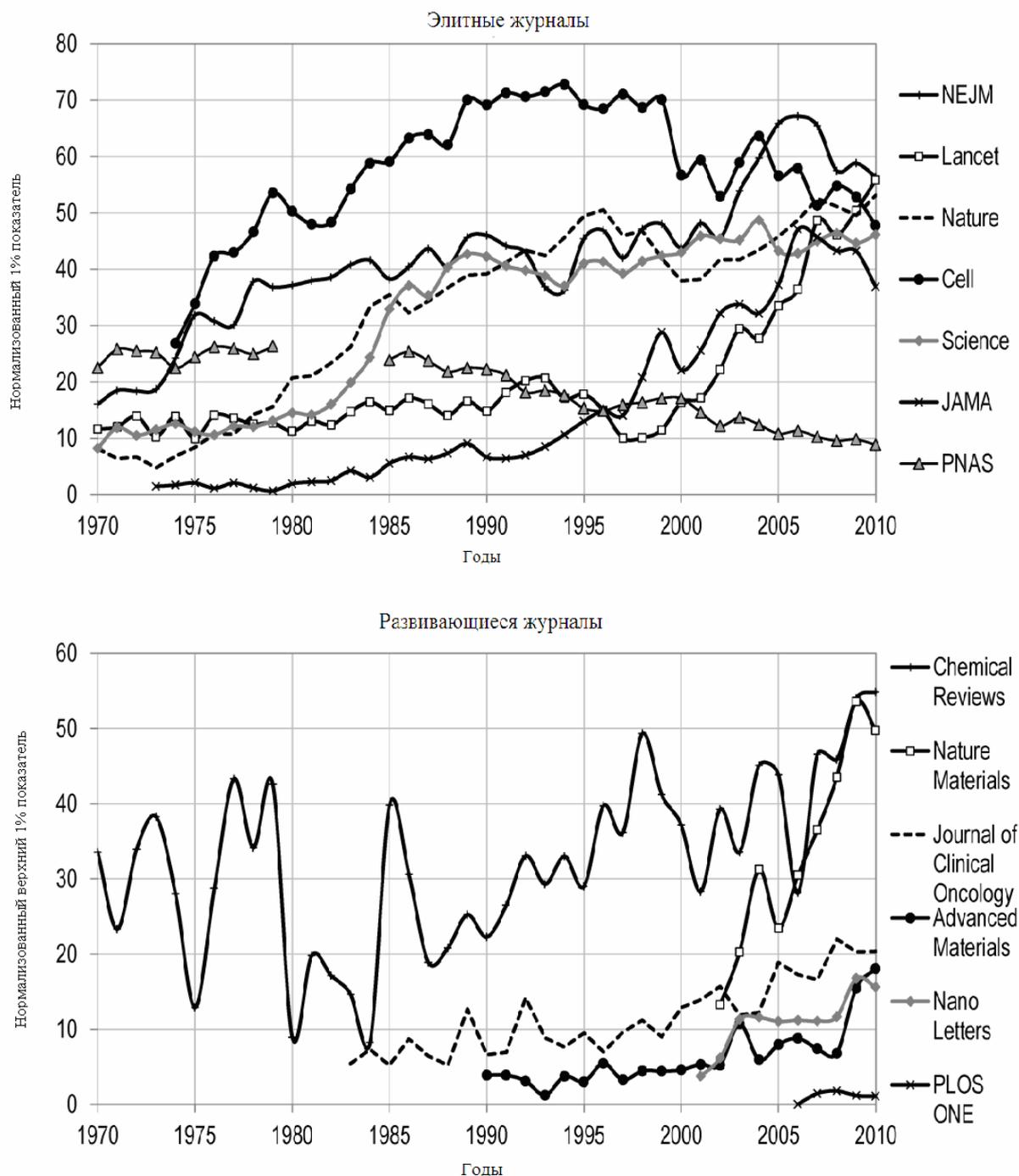


Рис. 4. Нормализованный 1% показатель для элитных и развивающихся журналов, 1970-2010 гг.

Снижение роли элитных журналов было вызвано Интернетом из-за нескольких взаимосвязанных эффектов. Во-первых, цифровой век сделал статьи более независимыми от соответствующих журналов. Цифровой век трансформировал способ нахождения учеными статей. Мы использовали для поиска статей книжные полки библиотеки, но теперь делаем это через Интернет, часто на нашем библиотечном сетевом сайте. Более того, статьи сегодня непосредственно доступны независимо и даже не надо смотреть соответствующий номер или том журнала. Значит, вопрос, становятся ли статьи цитируемыми или нет, является все более независимым от журнала, в котором они появляются.

Во-вторых, цифровой век также облегчил создание новых журналов. Несколько журналов высокого качества появились благодаря созданию новых областей (например, *Nano Letters*, *Advanced Materials*), но часто так происходило просто потому, что цифровые журналы легче и дешевле производить и распространять, чем печатные. Как только эти журналы были созданы, те журналы, которые уже находились в системе, вероятно, должны были потерять долю рынка, а теми, которые подверглись негативному влиянию, были такие, которые имеют самую высокую долю верхних статей. Подобное явление встречалось в доле статей на уровне стран. Когда Китай увеличил свою долю научных статей в мире в течение 1990–2000-х гг., доля самых важ-

ных стран упала, как в случае с США и Японией [14-15]. Дюжины новых издателей и сотни новых журналов появились в последние 20 лет. Неизбежно, что элитные журналы должны были потерять какое-то количество верхних статей в пользу новых конкурентов.

В-третьих, создание некоторых этих новых журналов было мотивировано идеалами «свободного доступа для всех», получившими преимущество легкости, с которой информация может распространяться в цифровом веке (например, *PLoS One*). Свободно доступные статьи имеют большую вероятность быть цитируемыми [16]. Ученым сегодня предоставлен доступ к статьям из большего разнообразия журналов, не только из так называемых передовых журналов в определенной области, или элитных журналов вообще. Технически статья может сегодня быть в любом журнале, и она будет найдена с помощью поисков через Интернет, загружена, если доступна, и цитируема, если покажется релевантной.

Хотя элитные журналы вообще теряют влияние, различные модели, которые мы подтвердили документами, не распадаются явно по предметным областям, таким как общего характера в противоположность биомедицинским. Подобным образом, модели, наблюдаемые в развивающихся журналах, не зависят от области исследования, возраста или новизны журнала. Среди развивающихся журналов, только один *PLoS One*, здесь исследуемый, является полностью «открытым», имел большую долю высокоцитируемых статей, но только потому, что он публикует большое число статей. Следовательно, в среднем его статьи совсем необязательно лучше, но, по крайней мере, они могут быть найдены и прочитаны любым человеком бесплатно. Эта аспектность за последние 20 лет показывает, что цифровой век коснулся каждого журнала в различной степени, вероятно, полагаясь на эффективность его редакционной, рекламной и маркетинговой политики.

Элитные журналы осознают угрозу со стороны новых журналов и стремятся соизмерить их влияние. Например, частная компания Nature Publishing Group вместо простого расширения ведущего журнала стала создавать специализированные журналы в 1990-х гг. (*Nature Genetics*, *Nature Neuroscience*, *Nature Medicine* и т.д.) и обзорные журналы в 2000-х гг. (*Nature Reviews Cancer*, *Nature Reviews Immunology* т.д.). Сегодня она публикует 37 других журналов с включением в название слова «Nature». Эти новые журналы наживают капитал на символическом капитале оригинального журнала *Nature*, и они быстро привлекают высококачественные статьи. Одним возможным объяснением выбора этого пути, явления, которое экономисты называют *версификацией*, является то, что издатель может назначить новые расценки на подписку и рекламу для новых журналов, тогда как трудно заставить потребителей и рекламщиков принять большой рост цен на ведущий журнал, даже если это сопровождается пропорциональным ростом объема журнала. Эти новые журналы распространяются и разрушают влияние ведущего журнала во всей журнальной системе издателя. При объединении всех журналов компании Nature Publishing Group доля верхних 5% статей увеличивается от 3% в 1970 г. до 5% в 2000 г. и остается стабильной до сих пор. С другой стороны, журнал *Science* издается Американской ассоциацией содействия развитию науки, некоммерческой организацией, которая публикует только 3 рецензируемых журнала.

При наличии всех этих новых журналов, элитных и других, ученые сегодня имеют все больше изданий, где они могут представить свои статьи и получить выгоду от обозримости и доступности, которая прежде была

возможна только в наиболее распространяемых журналах, «элитных» журналах. Учитывая высокий процент отказов в публикации со стороны элитных журналов, например, почти 93% для *Science*, ученые могут предпочесть сэкономить время и представить свои статьи в другие журналы, которые в конечном счете достигнут той же аудитории быстрее и, возможно, получат столько же ссылок. В цифровой век относительно легко определить действительный процент ссылок отдельных статей или авторов, поэтому значение репутации журнала сегодня является менее важным. Однако ученые могут все еще предпочитать печататься в элитных журналах. Оправданно или нет, репутация журнала пока имеет некоторое значение для статей в таких журналах благодаря эффекту Матфея [17], в частности, когда исследование рассматривается и оценивается не экспертами.

Наконец, комитеты по продвижению, набору персонала и оценке грантов, администраторы и другие эксперты должны учитывать эти результаты. Качество статей и компетентность ученых должны оцениваться независимо от журнала, в котором работа появилась, по двум причинам. Во-первых, традиционные «элитные» журналы все еще имеют самое высокое влияние ссылок, но другие журналы также публикуют все более высокую долю ведущих цитируемых статей. Во-вторых, даже если эта иерархия журналов была надежной, она не является фиксированной. Действительно, даже если репутации журналов свободно основаны на тенденции журнала публиковать высококачественную работу, репутации обладают внутренней инерцией и всегда есть временной пробел между действительной ценностью журнала и его репутацией. Хотя всегда будет существовать иерархия престижа журнала, эта иерархия является динамичной. Может быть использовано много других, более стабильных критериев для оценки ученых, но если должно быть использовано качество журнала, справедливость требует, чтобы эксперты по оценке были внимательны к изменениям в картине научного издательства.

*Благодарность.* Авторы благодарят Стефани Хостейн и Жана-Пьера Робитайя за полезные комментарии по теме. Ж. Лозано признателен Тартускому университету за беспрепятственный и свободный доступ к его электронным фондам.

## Литература

1. Harmon J.E., Gross A.G. The scientific literature: A guided tour. — Chicago: Chicago University Press, 2007.
2. Meadows A.J. Communication in science. — London: Butterworths, 1974.
3. Russell J.M., Rousseau R. Bibliometrics and institutional evaluation// *Encyclopedia of Life Support Systems (EOLSS)*. Part 19.3 Science and Technology Policy, edited by Rigas Arvantis, 2002. Developed under the auspices of the UNESCO, Eolss Publishers, Oxford, UK. pp. 42-64.
4. Lozano G.A., Larivière V., Gingras Y. The weakening relationship between the impact factor and papers' citation rates in the digital age// *Journal of the American Society for Information Science and Technology*. — 2012. — Vol. 63. — P. 2140-2145.
5. Bornmann L., Leydesdorff L., Mutz R. The use of percentiles and percentile rank classes in the analysis of bibliometric data: Opportunities and limits// *Journal of Informetrics*. — 2013.— Vol. 7, No.1.— P. 158-165.

6. *Waltman L., Schreiber M.* On the calculation of percentile-based bibliometric indicators.— 2012. — <http://arxiv.org/abs/1205.0646>.
7. *Waltman L., Calero-Medina C., Kosten J., Noyons E.C.M., Tijssen R. J. W., van Eck N.J., van Leeuwen Th. N, van Raan A.F.J., Visser M.S., Wouters P.* The Leiden ranking 2011/2012: Data collection, indicators, and interpretation// *Journal of the American Society for Information Science and Technology*. — 2012. — Vol. 63, No.12. — P. 2419-2432.
8. *Glanzel W., Debackere K., Thijs B., Schubert A.* A concise review on the role of author self-citations in information science, bibliometrics and science policy// *Scientometrics*. — 2006. — Vol. 67.—P. 263-277.
9. *Moed H.F.* Citation analysis in research evaluation. — Dordrecht: Springer, 2005.
10. *Barabasi A.L., Song C., Wang D.* Handful of papers dominates citation// *Nature*. — 2012. — Vol. 491, No. 40.
11. *Wallace M.L., Larivière V., Gingras Y.* Modeling a century of citation distributions// *Journal of Informetrics*. — 2009. — Vol. 3, No.4. — P. 296-303.
12. *Grossetti M., Eckert D., Gingras Y., Jégou L., Larivière V., Milard B.* The geographical deconcentration of scientific activities (1987-2007) //Proceedings of the 17th International Conference of Science and Technology Indicators. — 2012. — P. 348-356.
13. *Larivière V., Gingras Y., Archambault É.* The decline in the concentration of citations, 1900-2007// *Journal of the American Society for Information Science and Technology*. — 2009.— Vol. 60, No.4.— P. 858-862.
14. *Leydesdorff L., Wagner W.* Is the United States losing ground in science? A global perspective on the world science system// *Scientometrics*. — 2009. — Vol. 78, No.1. — P. 23-36.
15. *Zhou P., Leydesdorff L.* The emergence of China as a leading nation in science// *Research Policy*. — 2006. — Vol. 35, No.1. — P. 83-104.
16. *Gagouri Y., Hajjem C., Larivière V., Gingras Y., Carr L., Brody T., Harnad S.* Self-selected or mandated, open access increases citation impact for higher quality research//*PLOS ONE*. — 2010. — Vol. 5, No. 10. — e13636.
17. *Larivière V., Gingras Y.* The impact factor's Matthew effect: A natural experiment in bibliometrics// *Journal of the American Society for Information Science and Technology*. — 2010. — Vol. 61, No.2. — P. 424-427.
18. *Garfield E.* Citation indexes for science: A new dimension in documentation through association of ideas// *Science*. — 1955. — Vol. 122. — P. 108-111.
19. *Garfield E.* How can impact factors be improved?// *British Medical Journal*.— 1996.— Vol. 313.— P. 411-413.
20. *Larivière V., Archambault É., Gingras Y., Vignola Gagné É.* The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities// *Journal of the American Society for Information Science and Technology*. — 2006. — No.5. — P. 997-1004.
21. *Spier R.* The history of the peer-review process// *Trends in Biotechnology*. — 2002. — Vol. 20. — P. 357-358.
22. *Vakil N.* The journal impact factor: Judging a book by its cover//*Am. J. Gastroenterol.* — 2005.— Vol. 100. — P. 2436-2437.

## Этика использования услуг по редактированию языка в эру цифровой коммуникации и статей с очень большим числом авторов\*

**Жорж А. ЛОЗАНО**  
(George A. LOZANO)

Эстонский Центр эволюционной экологии, г. Тарту, Эстония

*Ученые многих стран, где английский язык не является основным, обычно используют спектр услуг по подготовке, исправлению или редактированию рукописи. Такая практика открыто одобряется многими журналами и научными учреждениями. Эти услуги очень разнообразны по своему охвату. С одной стороны, это может быть простое чтение корректуры, а с другой – осмысленное и расширенное рецензирование, подготовка предложений, статистический анализ, редактирование на грани переписывания и совместное написание. В данной статье, наряду с руководствами по авторству, рассматриваются разные типы услуг и поднимается вопрос, превзойдут ли высококачественные услуги большинство критериев руководств по авторству. Еще рассматриваются три других фактора. Первый – возможная в эру Интернета легкость сотрудничества позволяет многочисленные взаимодействия (итерации) между авторами и «службой подготовки к печати» (“editing service”), таким образом, по существу, статьи могут быть написаны совместно. Второй – службы подготовки к печати часто предлагают предметных экспертов, которые делают замечания не только в отношении языка, но интерпретируют и улучшают научное содержание (контент). Третий – тенденция в сторону статей с очень большим числом авторов подразумевает, что порог, необходимый для получения авторства, снижается. Неизбежный вывод состоит в том, что на каком-то этапе трудовое участие служб подготовки к печати должно рассматриваться в качестве достаточного, чтобы гарантировать право на авторство. Попытка навязать какие-либо руководства будет, вероятно, бесполезной, но, тем не менее, может быть пришло время снова посмотреть на этику использования некоторых высококачественных услуг по подготовке к печати. На постоянно растущем международном рынке труда признание этой проблемы может в прогрессивном плане оказаться более важным в диспутах относительно авторства, распределения грантов и принятия решений о подборе кадров.*

Английский язык – это язык науки. Чтобы подготовить свои статьи для представления, ученые многих стран, где английский язык не является основным языком, обычно используют разнообразные виды услуг по подготовке рукописи, исправлению и/или редактиро-

ванию. Журналы часто полагают, что лица, для которых английский не является родным языком, используют эти услуги, игнорируя тот факт, что плохое написание научного произведения представляет проблему, которая пересекает такие условные границы. В течение многих десятилетий использование этих услуг считалось стандартной практикой, полностью одобренной и на которой часто даже настаивали некоторые национальные академии и университеты в странах, не говорящих по-английски. Эти услуги очень разнообразны по своей сути. С одной стороны, существует чтение корректуры и

---

\* Перевод Lozano G.A. Ethics of using language editing services in an era of digital communication and heavily multiauthored papers. – <http://arxiv.org/ftp/arxiv/papers/1305/1305.2401.pdf>

литературное редактирование, а с другой – глубокое и расширенное рецензирование, анализ данных, редактирование на грани переписывания текста и совместное написание. Чтение корректуры и литературное редактирование существуют давно, а с появлением Интернета, облегчающего обширное сотрудничество, высококачественные услуги, осуществляемые специалистами, стали в возрастающей степени распространенными. За несколько последних десятилетий очень возросло многоавторство, которое в действительности снизило необходимый для авторства порог. Учитывая эти изменения в манере нашей работы, в данной статье я задаюсь вопросом, существует ли какой-то момент, при котором трудовое участие служб подготовки к печати должно считаться достаточным, чтобы получить право на авторство. Я прихожу к выводу, что, возможно, как раз пришло время пересмотреть этику, стоящую за этой давно существующей практикой.

## **УСЛУГИ ПО ПОДГОТОВКЕ, ИСПРАВЛЕНИЮ И РЕДАКТИРОВАНИЮ**

Доступен широкий спектр работ по подготовке, исправлению текста и/или использованию услуг по редактированию. Самой основной услугой является чтение корректур. Корректор является последним звеном в линии защиты от незначительных ошибок, которые не были замечены авторами, рецензентами, редакторами или помощниками редактора. Корректор исправляет явные ошибки в пунктуации, дефисном написании, письме прописными буквами, пробелах между словами и строками, порядке слов, грамматике и правописании и ставит вопросы авторам относительно других незначительных потенциальных проблем. В издательской индустрии корректор также тщательно сравнивает одобренную версию рукописи с корректурой в гранках, которая является точной копией того, что должно быть напечатано в больших количествах. Многие журналы все еще содержат в штате корректоров, но они имеют дело со статьями после их принятия, на конечных стадиях перед опубликованием. Этот аспект работы корректора не обеспечивается службами подготовки к печати. Работа служб подготовки к печати происходит до того, как статья представлена в журнал.

Еще одним шагом по улучшению рукописи является литературное редактирование. Редактор исправляет некоторые из тех же самых спорных моментов, которые позже снова будут учтены корректором, но, помимо того, литературный редактор рукописи изучает и исправляет структуру предложений, целостность, согласование времен, жаргон и принимает во внимание характерные для журнала требования. Хотя литературный редактор рукописи имеет в запасе больше времени, чем корректор, работа литературного редактора заключается в том, чтобы улучшить текст без осуществления каких-либо совсем необязательных изменений. Следовательно, предложения могут переставляться для обеспечения лучшего потока идей, но стиль написания не будет очень меняться. Для выяснения чего-то более сложного литературный редактор рукописи задает вопросы автору, например, если имеются сегменты, содержащие неясность и требующие разъяснения. Литературный редактор делает текст технически правильным, возможно, даже совершенным в структурном плане, но не вносит каких-либо значительных изменений в содержание. Большинство журналов имеют в штате литературных редакторов, но они работают со статьями только после их принятия.

На более высоком уровне существует редактирование «содержания» или «редактирование по сути». То есть там, где значения этих терминов, используемых службами подготовки к печати, начинают отклоняться от стандартных значений этих терминов. Традиционно термин «редактор содержания (контента)» относится к тому, кто на самом деле не рассматривает вопросы, которые позже могут быть охвачены литературным редактором и корректором, а концентрируется на более значимых проблемах, следовательно, на «существенных». Например, редакторы издаваемых книг являются редакторами содержания в традиционном смысле использования данного термина; они просто дают множество советов. Например, редактор содержания может предложить, чтобы «Введение» было более целенаправленным, «Методы» более подробными, «Результаты» оформлены иначе, а «Обсуждение» более тесно связано с «Введением» и «Результатами». Советы редакторов содержания могут стать более определенными, выделяющими сегменты, которые являются неясными, параграфы, которые следует поменять местами или убрать, соответствующие темы, которые надо включить, и т.д. Однако редакторы содержания, на самом деле, не пишут и не переписывают рукопись. Тем не менее, среди служб подготовки к печати существенное (по сути содержания) редактирование в таком смысле встречается сравнительно редко.

Редакторы содержания, работающие в сфере служб подготовки к печати, дают рекомендации не только относительно того, что уже есть, но также исправляют, переделывают, расширяют или убирают предложения, параграфы и целые разделы. Фактически рукопись не отредактирована, а скорее частично или полностью переписана. Степень переписывания зависит от оказываемой услуги, ожиданий, оригинального состояния статьи, желания лица, работающего над статьей, более глубоко вникнуть в нее и, конечно, от стоимости услуг. Как правило, происходит несколько взаимодействий между службой подготовки к печати и автором, начиная с экстенсивных переписываний как со стороны редактора, так и со стороны автора, и заканчивая литературным редактированием и чтением корректуры. Однако большинство служб подготовки к печати не используют термин «переписывание», когда характеризуют и рекламируют свои услуги, но, тем не менее, их так называемые услуги по «существенному редактированию» на самом деле являются услугами «переписывания заново».

Другие службы не беспокоятся о таких незначительных моментах семантики. Например, на своем премиум-уровне (“premium level”) какая-то служба обеспечивает «такое преобразование рукописи, чтобы каждый раздел включал необходимое содержание (т.е. результаты приводятся отдельно от методов и интерпретаций; все общие сведения даются скорее в разделе «Введение», а не в разделе «Обсуждение» и т.д.)» [1]. Таким образом, по существу, они сами рассматривают основные принципы написания научной статьи, аналогично тому, как поступает преподаватель при обучении студента тому, как писать свои первые статьи. Кроме того, на высшем уровне (“superior level”) подготовки рукописи редактор потратит дополнительное время, необходимое для большего переписывания, преобразования, форматирования... Это достаточный выбор для клиентов, которые хотят отправить относительно сырой черновой материал и которые нуждаются в существенной помощи по организации материала и имеют большое количество рисунков и таблиц, не соответствующих принятому для

публикации качеству» [1]. В этих разделах приводятся выдержки, чтобы ликвидировать неясности. Отметим формулировку «клиенты, которые хотят отправить относительно сырой черновик». Это уже выходит за рамки того, чему многие преподаватели учат своих аспирантов.

Высококачественное «редактирование содержания» иногда называют «эволюционным, связанным с развитием, редактированием». Эта «услуга идет значительно дальше, нежели простое редактирование языка – специалист с докторской степенью, редактор с опытом редактирования высококачественных журналов и профессионализмом в вашей области будет тщательно работать над вашей рукописью» [2]. Такой редактор рассматривает «вопрос или цель, которые обосновывают исследование, структурные изменения, необходимые в рукописи, согласованность и поток аргументов в статье, неясные моменты, требующие разъяснения, явные логические несовместимости в заявленных гипотезах, потенциальные аналитические или методологические слабости, насколько правильно выводы подкрепляются результатами и т.д.» [2]. Следовательно, такой редактор является экспертом в данной области, главным образом коллегой, чей вклад в подготовку рукописи выходит за рамки простого редактирования языка, и эффективно переписывает или пишет статью вместе с автором. Как и в любом другом многоавторском сотрудничестве, предложения по его составу к окончательной версии находятся на рассмотрении первого и/или старшего автора.

Ряд служб предлагает услугу по отклонению и вторичному представлению рукописи, для чего, например, «руководитель проекта сначала оценит комментарии рецензента относительно того, в какой доработке нуждается ваша рукопись, а затем закажет пакет услуг по подготовке к печати, чтобы представить рукопись, которая будет отвечать ожиданиям журнала. Повторное представление рукописи для журнала включает многократные этапы редактирования и тесную переписку между редактором и автором» [3]. Таким образом, они вместе пишут рукопись в соответствии с требованиями рецензента и дают подробные ответы на комментарии рецензента.

Большинство высококачественных служб также предлагают услуги, называемые «рецензирование коллегами» и «предложение журнала относительно опубликования». Рецензирование коллегами не относится к рецензированию, которое происходит после представления статьи, а скорее относится к рецензированию, которое имеет место до ее представления, т.е. типу рецензирования, которое часто осуществляется коллегами внутри сообщества, до представления рукописи к публикации. Эксперт в данной области посмотрит (отрецензирует) статью и даст полный отчет, указывая сильные и слабые стороны и делая множество других предложений, но в этом случае статья не будет переписываться. Эта работа схожа с работой редактора содержания или с рецензированием, предшествующим представлению статьи или осуществляемым после ее представления в традиционных значениях терминов. Предложение журналов относительно публикации статьи относится к услуге, когда эксперты в данной области изучают статью и рекомендуют для нее наиболее подходящие журналы в зависимости от намерений и предпочтений автора. Единственной причиной оплачивать какую-либо из этих услуг является то, что считается, что научное знание журналов и области у экспертов больше, чем такое автор и ближайших коллег.

Другие услуги начинаются не со статьи, а с научного предложения. Можно провести обзоры литературы, которые примут во внимание ваши непосредственные исследовательские задачи и проблему, синтезируют результаты в базу знания, четко определяя настоящее положение доступного знания и неизвестной информации, определяют области полемики в литературе и сформулируют вопросы, требующие дальнейшего исследования. Специалисты, осуществляющие эти услуги, работают с материалами, представленными клиентами и полученными нашими осуществляющими написание статьи лицами, чтобы после тщательного изучения подготовить подробное, убедительное, согласованное, четкое и привлекающее внимание научное предложение/использование гранта. Они убеждают потенциальных клиентов снабжать нас приблизительной темой своего исследования; наши осуществляющие написание статьи лица и редакторы напишут для вас привлекающее внимание и убедительное научное предложение [4]. Здесь нет никакой двусмысленности: представьте приблизительную тему, а пишущие лица и редакторы сделают все остальное.

Та же самая компания под своими заказными услугами подразумевает три типа услуг: заказное написание статьи, анализ и интерпретация данных и специальное консультирование. В заказном написании написание статьи выполняется на основе чернового материала или предварительного наброска, предложенного клиентом. Наши редакторы [так!] и осуществляющие написание статьи лица проводят тщательный поиск литературы, используют материалы, методы и результаты/необработанные данные, предоставленные клиентом, для написания выразительно представленного текста. В анализе и интерпретации текста они выполняют статистический анализ и интерпретацию сырых данных, представленных клиентами. Это относится не столько к статистическим тестам, запрашиваемым клиентом, а скорее к полному анализу, начиная с сырых данных и проведения всех предварительных анализов, выбора и проведения соответствующих статистических тестов и интерпретаций результата. И, наконец, услуга специального консультирования предназначена для клиентов, которым необходима помощь в отношении концепций и концептуального содержания. Наши редакторы могут объяснить научные концепции и обеспечить обратную связь в отношении того, подходят ли ваши идеи современно установленному научному корпусу знания [5]. При наличии этих услуг очевидно, что глубокое понимание требующей расширения науки не всегда является предпосылкой для того, чтобы стать плодотворным ученым.

Это всего лишь просто примеры некоторых многогранных служб подготовки к печати. Как заявляют сами службы, некоторые из этих услуг выходят далеко за рамки простой коррекции языка. Они производятся специалистами в рамках узкой области экспертизы, главным образом, анонимными коллегами, которые делают экстенсивный вклад в: 1) концепцию и разработку исследований на стадии предложения, 2) статистический анализ вытекающих данных, 3) интерпретацию в контексте современного знания и 4) написание и повторное написание конечного продукта. Цель здесь состоит не в том, чтобы одобрять или критиковать любую из этих услуг, выбранных в качестве примера. Дюжины и даже сотни подобных услуг можно легко найти в режиме онлайн. Они просто предоставляют услугу и не являются теми, кто отвечает за статью. Дело в том, что где-то в последовательности многообразия этих услуг, от про-

стого чтения корректуры до услуг высокого уровня, может быть есть линия пересечения. Не ясно, где находится эта линия, но некоторые из этих высококачественных услуг четко осуществляют работу, которая традиционно считалась работой, заслуживающей права авторства.

## ЧТО СОБОЙ ПРЕДСТАВЛЯЕТ АВТОРСТВО?

Авторство научных работ обычно представляло собой простое понятие. Со времени создания первых научных журналов и до начала XX в. большинство статей имело одного автора. Было достаточно легко определить, кто является ответственным и кому достанется доверие. Сейчас это уже не так. С ростом числа авторов на статью само авторство становится высоко дискутируемой областью исследования [6-14].

Во всех руководствах, касающихся авторства, всегда имеется предостережение, что условия относительно авторства разнятся среди дисциплин. Международный комитет редакторов медицинских журналов (International Committee of Medical Journal Editors – ICMJE) [15] полагает, что значение авторства должно основываться на 1) реальных вкладах в концепцию и разработку, получение данных или анализ и интерпретацию данных; 2) подготовке чернового варианта или его критическом просмотре для важного интеллектуального содержания (контента) и 3) окончательном одобрении версии, которая будет опубликована. Авторы должны отвечать условиям 1, 2 и 3 (*выделено автором*). Союз «и» является моментом спора [16], поскольку, если бы можно было придать строгую силу этим руководствам, тогда многие статьи вообще перестали бы иметь автора. Люди, работающие на службы подготовки к печати, редко одобряют окончательную версию [17], и утверждается, что третье требование даст техническую возможность придерживаться руководств, в то же время исключая лиц, которые иным образом внесли значительный вклад в исследование и саму рукопись [18]. Союз «или» будет более реалистичным в том случае, когда какая-то работа высокого уровня, проведенная службами подготовки к печати, будет несомненно отвечать за авторство по пунктам 1 и 2.

Совет научных редакторов (Council of Science Editors – CSE) предлагает схожие руководства, но взамен, в первую очередь, концентрируется на наборе общих принципов, указывая, например, что определение авторства не является компетенцией редакторов, что все лица, внесшие значительный вклад, должны быть указаны в качестве авторов и что авторы должны одобрить статью перед публикацией [19]. Чтобы действительно определить, кто классифицируется как автор, они прежде всего прибегают к руководствам ICMJE, а затем предлагают свои собственные: «Авторы – это лица, определенные исследовательской группой как внесшие значительный вклад в представленную работу и согласившиеся отвечать за этот вклад. Кроме того, чтобы быть ответственным за часть работы, автор должен быть способным определить, кто из соавторов ответственен за другие конкретные части работы. Помимо этого, автор должен быть уверен в целостности вкладов своих соавторов. Все авторы должны просмотреть и одобрить окончательную рукопись». Кроме пункта о конечном одобрении, эти руководства значительно отличаются от таковых ICMJE и накладывают различную нагрузку на каждого автора.

CSE вносит дополнение относительно того, что авторство не касается «профессионалов, осуществляющих написание, которые участвовали только в редактирова-

нии рукописи и не играли какую-либо роль в разработке или проведении исследования или интерпретации результатов». Вероятно, цель здесь заключается в том, чтобы предупредить людей, работающих на службы подготовки к печати о том, что они не будут включены в качестве авторов. Однако, учитывая действительный характер отношения между службами подготовки к печати и их клиентами, в ряде случаев это руководство будет иметь обратный эффект. Как описано выше, служба подготовки к печати иногда не только пишет статьи, но так же играет решающую роль в замысле исследования и интерпретации результатов. В конце концов, CSE обуславливает, что авторство не касается лиц, которые только дают рекомендации, обеспечивают область исследования или финансовую поддержку, что дисквалифицирует многих руководителей диссертаций, ведущих исследователей, руководителей групп и главных исследователей. Эти руководства могут иметь хорошее намерение, но они не соответствуют реальности ситуации.

Национальный институт здравоохранения США [20] расширил и упростил руководства: «авторство должно основываться как на значительном вкладе в концептуализацию, подачу материала, выполнение и/или интерпретацию научного исследования, так и на редактировании или существенном рецензировании или просмотре (исправлении) научной статьи, и на желании взять на себя ответственность за исследование». В этом случае работа, проведенная профессиональными службами подготовки к печати, определенно включается, если они (службы) хотят нести ответственность за исследование. Просто, вероятно, их никогда об этом не спрашивают.

Многие научные учреждения, национальные академии, профессиональные общества и журналы также имеют свои собственные руководства относительно авторства. Шестьдесят процентов биомедицинских журналов [21], 53% профильных научных журналов, 32% журналов по общественным наукам и 6% журналов по искусству и гуманитарным наукам [22]. Для сравнения, только 11% профессиональных обществ имеют руководства по авторству в своих сводах этических правил [22]. Исследование и написание статьи являются основными требованиями для авторства в журналах, но две трети сводов этических правил включают исследование в качестве единственного критерия [22]. Несмотря на это разногласие, обзор и метанализ проблем авторства показывает, что вообще замысел исследования и/или подача материала исследования и написание рукописи являются наиболее важными критериями для авторства [23]. Руководства не только различаются, но и применяются только к соответствующим исследовательским учреждениям, национальным академиям, профессиональным обществам или журналам, таким образом, даже если бы они могли быть строго навязаны, их сфера полномочий была бы ограниченной.

Существуют другие формы авторства, которые, как правило, считаются неэтичными, но, тем не менее, довольно распространенными. Имеются две стороны монеты; примеры, когда отдельные лица, соответствующие критериям авторства, не приводятся в качестве авторов, и примеры, когда авторство дается лицам, которые не отвечают таким критериям [16, 24-28].

«Приглашенное», «почетное» или «дарственное» авторство относится к примерам, когда кто-то, кто обычно не соответствует критериям для авторства, тем не менее, включается как автор. Например, при опросе химиков

США около 20% сообщили, что они обнаружили себя в качестве авторов статьи только после того, как она была напечатана [26]. Причины для включения разные. **Приглашенное авторство** касается включения старших авторов, основанного на ожидании того, что их включение повысит статус статьи в процессе рецензирования и/или после опубликования, что похоже на приход на встречу со знаменитостью. На самом деле, это подразумеваемая, но сильная критика авторами редакторов, рецензентов и возможных читателей, которые, как предполагается, находятся под впечатлением фамилий авторов, а не качества статьи (*auctoritas, non veritas, facit legem* – авторитет, а не истина, делает закон). **Дарственное авторство** предоставляется как часть соглашения по дарственному обмену (*quid pro quo – услуга за услугу*), где коллеги и молодые авторы, которых, самое большее, должны упоминать в благодарностях, повышаются до уровня соавторов. **Почетное авторство** дается *ex officio* (по должности) старшим преподавателям или заведующим кафедрами (*honoris causa – за заслуги*). Не ясно, является ли здесь слово «почетный» оригинальным или вынужденным, и если имеется в виду последнее, то является ли оно явным или подразумеваемым. Тридцать два процента радиологов при опросе сообщают, что их попросили включить почетного автора в свои статьи [29]. За период 1978 – 1998 гг. число авторов значительно возросло, в основном из-за роста требований старших преподавателей и заведующих кафедрами, претендующих на соавторство [30].

С другой стороны, несостоятельность относительно включения в качестве авторов лиц, которые соответствуют критериям авторства, приписывается так называемому безымянному авторству (*ghost authorship*). Безымянное авторство имеет место, когда индивидуумы охотно или неохотно исключаются из авторов. Преднамеренное исключение может происходить, когда предполагается, что включение определенных авторов может снизить степень доверия к статье. Например, могут проводиться медицинские исследования и будут писаться статьи служащими фармацевтическими или осуществляющими выпуск инвентария компаний, но авторство может быть ограничено лицами без каких-либо явных столкновений интересов. Проблемы и этические нормы, связанные с этой формой безымянного авторства подробно обсуждались в разных источниках [16, 18, 27, 31-33]. Второй тип безымянного авторства встречается, когда легитимные авторы неохотно или непреднамеренно исключаются как авторы. При опросе химиков США 50% отметили, что они не получили заслуженного авторства и не получили должного признания [26]. Проблема потенциальных авторов, которые были непреднамеренно исключены, продолжает существовать, так как большинство протоколов диспутов относительно авторства требуют того, чтобы их вернули обратно в учреждение, где проводилось исследование, место, где, как правило, работают старшие авторы и где спорный вопрос рассматривается коллегами старшего автора. Следовательно, справедливость будет восстановлена. Высококачественная работа ряда служб подготовки к печати, вероятно, располагается где-то между этими двумя формами безымянного авторства. С одной стороны, допущение того, что службы подготовки к печати очень сильно вовлечены в производство рукописи, может понизить уровень доверия к статье и репутацию других авторов. С другой стороны, предложение авторства приписанным ученым, работающим на службе подготовки к печати, редко является правом выбора,

но уже обсуждалось, что пишущие для служб подготовки к печати люди иногда должны включаться в качестве авторов [18, 34].

Чтобы покончить с этим затруднительным положением, предлагается, что журналы должны перейти от приведения фамилий авторов к указанию фамилий тех, кто внес свой непосредственный вклад [35, 36]. Эта модель сотрудничества с учетом внесения вклада (*contributorship*) похожа на модель, принятую в кино, где каждый, кто внес вклад в конечную работу, упоминается в титрах в соответствии со своей определенной ролью. Многие журналы просят авторов определить тех, кто внес вклад в статью, при ее представлении, но эта информация не включается в саму статью при публикации. Другие журналы включают эту информацию в статью, но только в дополнение к традиционному списку авторов. Таким образом, сотрудничество с учетом непосредственного вклада когда-нибудь должно заменить авторство. Проблема, связанная с таким сотрудничеством, состоит в том, что даже допуская, что те, кто вносит в него свой вклад, являются абсолютно честными людьми, они нежелательны; при вторичном опросе только 30% авторов медицинских журналов постоянно сообщали о тех же самых сотрудниках в своих статьях [37]. Тем не менее, даже используя систему сотрудничества с учетом вклада, профессиональная работа, проводимая службами подготовки к печати, должна быть, по всей вероятности, показана с учетом ряда ее известных ролей.

#### **БЫСТРОЕ РАЗРАСТАНИЕ СТАТЕЙ С ОЧЕНЬ БОЛЬШИМ ЧИСЛОМ АВТОРОВ**

Несмотря на численный рост и известность руководства по авторству, они явно не работают [29, 36, 38]. Это происходит из-за того, что руководства - это просто руководства, а не правила, и принуждение следовать им здесь невозможно. Одним из принципов, одобренным большинством руководств по авторству, является то, что решение вопроса об авторстве не входит в компетенцию редакторов, журналов или финансирующих организаций, а скорее касается людей, непосредственно выполняющих работу. Однако соавторы сходятся во взглядах только в 30% случаев относительно вкладов друг друга [37]. Более того, как упоминалось выше, ICMJE требует, чтобы авторы отвечали всем трем критериям авторства, однако на условиях анонимности 60% радиологов сообщали, что ряд авторов **их собственных статей** не отвечали ни одному критерию авторства ICMJE [29]. Следовательно, кроме их коллективной честности, которая должна противопоставляться возможности внутренней борьбы или тайного соглашения между авторами, имеется мало того, что предостерегает всех авторов от поддержки требований каждого относительно «достаточно значимых» вкладов.

Проблема состоит в том, что обычно нет реальных материальных затрат для добровольного или заслуженного добавления большого числа авторов. При включении в качестве автора, как правило, выгоду от включения получает тот, кого включили, а что касается других авторов, то затраты на включение еще одного автора являются незначительными. Пока каждый автор может претендовать на каждую статью и на каждую ссылку как на свою собственную, статьи и ссылки не делятся авторами, а скорее магическим образом множатся посредством числа авторов [6, 39, 40]. Побочная выгода заключается в том, что, если статья когда-то была отклонена, а в ней много авторов, то легче избежать ответственности [10]. Следовательно, при современном положении дел суще-

ствующая система вознаграждает растущие статьи с очень большим числом авторов. В результате за последние 50 лет число авторов на статью увеличилось [6, 7, 9, 11, 12, 30]. Пятьдесят лет назад одиночное авторство было нормой, но в наши дни это является весьма удивительным. Увеличение числа авторов происходит не только из-за роста работ многодисциплинарного характера и научной сложности [41].

Стоимость включения большего числа авторов минимальна, даже когда дело касается нематериальных затрат, таких как желание, чтобы ссылались на фамилию. Статьи с одним или двумя авторами, что в настоящее время редкость, рассматриваются с использованием фамилий обоих авторов, пока здесь все хорошо. Некоторые авторы могут захотеть быть первыми, но на самом деле это не имеет значения. Мы можем иметь в виду классиков в какой-либо области [42], и, как правило, никто из них не рассматривается в качестве первого автора, а обычно сама работа воспринимается как истинное и равное сотрудничество. Как в случае традиционного использования термина «автор» в литературе, доверие и ответственность приписываются в равной степени и, что более важно, оба автора приводятся по фамилии. В случае трех или более авторов только первый автор упоминается по фамилии, а остальные просто приводятся под сокращением “*et al.*” («и др.»). Некоторый конфликт может существовать в отношении точного порядка соавторства, но, тем не менее, если однажды достигнут порог, отражающий трех авторов, добавление большего числа авторов влияет на то, упоминаются ли конкретно другие авторы по фамилии.

Когда порог “*et al.*” («и др.») увеличивается от третьего до седьмого автора, то опять становится очевидным эгоизм (эффект «моего я»). Ванкуверский стиль упоминания авторов, принятый ICMJE, требует, чтобы все авторы приводились по фамилиям в статьях с 6 или меньшим числом авторов, но для статей с числом авторов более шести только 6 первых авторов приводятся по фамилии, а остальные заменяются сокращением “*et al.*” («и др.»). В биомедицинских журналах частота распространения статей, с учетом числа авторов, быстро растет от одного до шести авторов, но затем резко падает на семи авторах и выше [43]. Ванкуверские руководства делают это похожим на то, как если бы биомедицинские статьи внезапно стали иметь новое оптимальное число авторов – 6. Реальность заключается в том, что все авторы хотят, чтобы их фамилии были отражены в списке, когда статья цитируется, поэтому хорошо иметь больше авторов, но в этом случае максимум только шесть.

Тенденция в сторону большего числа авторов не замедляется [14, 28, 44], и это не из-за работ более многодисциплинарного характера и научной сложности [41]. Ван Лун констатирует, что не существует обоснованной причины того, почему какая-либо статья должна иметь более 3 авторов, и предлагает радикальное решение, что журналы должны предупреждать авторов о том, что статьи с числом авторов более трех менее вероятно будут приняты. Это пока не происходит, и, вероятно, не случится в ближайшее время. Менее экстремистское и более практическое решение будет состоять в том, чтобы разделить каждую статью и каждую ссылку по числу авторов, таким образом, доверие и влияние будут разделены между авторами, не будучи увеличенными числом авторов [39, 40, 46]. Например, используя этот метод для статьи журнала *Nature* 2001 г. о первоначальных результатах и анализе генома человека, которая насчитывала 244 автора и получила 5 968 ссылок на конец 2007 г., мы

имеем: вместо того, что каждый автор получит 5 968 ссылок, он (автор) должен получить свою заслуженную долю уважения, 5 968/244 или примерно 24 ссылки [47]. Тенденция в сторону статей с очень большим числом авторов является самоусиливающейся; поскольку число авторов растет, то как соответствующий вклад каждого автора, так и порог для авторства снижаются. В какой-то момент вклад каждого автора настолько мал, что соответственно вклад служб подготовки к печати может считаться достаточным для авторства.

## ИНТЕРНЕТ

Инфляции (раздуванию) авторства способствовала технология, которая упрощает коммуникацию и сотрудничество. До того, как фотокопировальные аппараты стали широко доступными в 1960-х гг., было даже невозможно осуществлять рецензирование коллегами [48]. Вместо этого редакторы просто выбирали самые лучшие статьи, возможно, просили внести небольшие изменения, но в основном принимали статьи в том виде, в каком они представлялись. Авторы просто не могли сотрудничать так, как это возможно сегодня. Даже после того, как стали использоваться фотокопировальные аппараты, надо было сделать настоящие физические копии и отправить каждому соавтору, который затем мог внести новые предложения, добавить новые параграфы или переписать существующие. Главным образом, тот, кто осуществляет совместное написание, был ограничен тем, что мог сделать только рукописные заметки, которые основной автор затем включит в текст или окончательную версию рукописи. Времена изменились. Теперь даже возможно, чтобы статья одновременно совместно писалась несколькими авторами, находящимися в разных частях мира. Следовательно, до появления обработки текста и электронной почты службы подготовки к печати были ограничены исправлением языка, литературным редактированием и чтением корректуры; направленные на высокий конечный результат современные службы были просто невозможны. Изменение вполне подходящих этических руководств с целью замены устаревшего содержания новыми технологиями не является чем-то необычным. В этом случае этические руководства относительно использования служб подготовки к печати просто не справляются с легкостью осуществления коммуникации и сотрудничества, принесенной Интернетом.

## ГДЕ ГРАНИЦА?

В разных культурах этика различается. Английский язык – это современный язык науки, и ученые из стран, где английский язык не является основным, естественно, вероятнее всего, должны использовать службы подготовки к печати. Сначала это могли быть службы только по исправлению английского языка, но со временем, с появлением новой технологии и облегчающего процесс Интернета, эти службы расширились до такой степени, что сейчас их можно использовать как службы подготовки к печати для выполнения всего, кроме сбора данных.

Тип безымянного написания статьи (ghostwriting), где безымянные авторы охотно и сознательно исключены как авторы, является более известным в медицине, чем в других областях. В медицине, где на карту непосредственно поставлены человеческие жизни, часто доступно больше денег, и, может быть, более важно не только быть беспристрастными, но и поддерживать ви-

димость беспристрастности. Доступность финансов также дает возможность использовать более дорогие службы подготовки к печати и превратить безымянное написание статей в прибыльную и стабильную карьеру, где те, кто пишет, могут быть удовлетворены своей зарплатой и не будут беспокоиться о получении авторства за свою работу [17]. Такие службы подготовки к печати могут быть дорогими. Даже обычное чтение корректур и литературное редактирование могут стоить около 10 евро за страницу (т.е. 250-300 слов). Высококачественные услуги, описанные выше, могут стоить несколько тысяч евро за статью. Следовательно, всего несколько статей могут добавить значительное финансовое бремя на большинство исследовательских бюджетов, стоимость может быть препятствием для всех исследовательских программ, кроме финансируемых самым лучшим способом. Высокая стоимость означает, что вне рамок медицины, использование служб подготовки к печати традиционно ограничено странами западной и северной Европы и Японией. В последние несколько десятилетий ученые из других растущих экономик, таких как Китай, Южная Корея и Бразилия, начали пользоваться преимуществами таких служб подготовки к печати.

Однако, поскольку наука становится более интернациональной и междисциплинарной, лингвистические, финансовые и культурные различия должны прийти в столкновение. Комитеты по набору персонала и агентства по распределению грантов могут быть заинтересованы в ученых с обширными списками публикаций и ссылок, но они не всегда учитывают, что кандидаты имеют различное отношение к службам подготовки к печати и используют их в различной степени. Финансирующие агентства в ряде законодательств даже не могут позволить себе тратить финансовые средства на службы подготовки к печати. Следовательно, на международном рынке труда исследователи, которые полагаются на службы подготовки к печати, могут считать свою продуктивность «мертвым грузом», если они попадают под законодательства, где службы подготовки к печати не используются или не разрешены. Дополнительная стоимость также означает, что высокозатратная по стоимости продуктивность и влияние [49, 50] будут ниже, даже если службы подготовки к печати представляют собой допустимую статью расхода. В любом случае, где-то на протяженности от корректирования до совместного написания статьи, эта линия (граница), вероятно, должна пересекаться. Точное положение этой границы может различаться среди культур, но некоторые факторы, которые следует учитывать, - это число авторов, степень взаимодействия между автором и службой подготовки к печати и редакторский уровень экспертизы в определенной области.

Поскольку число авторов растет, ответственность, доверие и суть авторства становятся слабее [12, 16, 46]. Так как соответствующий вклад каждого автора снижается, то растущие небольшие вклады становятся достойными авторства. В ряде областей и культур, несмотря на все руководства, авторство распространяется и на знающих свое дело технических сотрудников, «внутренних» редакторов и рецензентов, старших ученых, предоставляющих лабораторные помещения, статистиков, приглашенных и почетных соавторов. Этот процесс может быть самоусиливающимся; поскольку добавляется больше авторов, порог для авторства даже снижается и становится проще оправдать добавление еще одного автора. Возьмем статью с количеством авторов  $n$ , приведенных в порядке соответствующих им вкладов.

Поскольку  $n$  растет, в определенный момент тип далеко идущего исследования, аналитическая работа и работа по написанию статьи, осуществляемая службами подготовки к печати, качественно и количественно может превысить вклад  $n$ -го автора, и, следовательно, может рассматриваться как достаточная для того, чтобы претендовать на авторство.

Вторым фактором, который надо принять во внимание, является степень взаимодействия между службой подготовки к печати и автором. До того, как широкое распространение получила электронная коммуникация, повторные циклы написания статьи, исправления и переписывания текста, имевшие место между авторами и службами подготовки к печати, были просто невозможны. Литературные редакторы и корректоры, когда они использовались до представления материала, только исправляли текст; они ничего в него не добавляли и не делали никаких существенных изменений. Редакторы содержания (контента) использовались в традиционном смысле и предлагали общие замечания без фактического переписывания каких-либо основных разделов рукописи. Однако новая технология облегчила написание статей с помощью итераций. Трудно определить соавторство, но любое истинное сотрудничество характеризуется повторными взаимодействиями между соавторами, когда они работают над рукописью [22, 51]. Как уже описывалось выше, за исключением сбора данных, службы подготовки к печати могут быть серьезно включены во все аспекты производства статьи, от начала и до конца, начиная от предложения (плана) и заканчивая литературным редактированием и чтением корректуры. Ясно, что, по крайней мере, в ряде случаев служба подготовки к печати и автор(ы) по существу пишут статью вместе.

Третьим моментом, который требует внимания, является степень компетенции людей, работающих на службы подготовки к печати. Странно, что лица с ученой степенью доктора философии, доктора медицинских наук или кто-то еще со схожим продвинутым образованием, как правило, не представляют собой лучший выбор для чтения корректуры или литературного редактирования статьи. Это вызвано двумя причинами. Во-первых, это было бы простой тратой их образования и компетенции. Во-вторых, чтение корректуры особенно, и литературное редактирование в меньшей степени, требуют, помимо всего прочего, чрезвычайного внимания к деталям. От этих лиц требуется способность видеть отдельные деревья, не отвлекаясь на лес в целом. Нельзя сказать, что чтение корректуры не может быть выполнено кем-то еще с продвинутой степенью образования, но такой человек всегда будет подвергаться соблазну предложить или сделать самостоятельно более существенные изменения. Чтение корректуры и литературное редактирование являются навыками и услугами, требующими зоркого глаза к деталям и знания языка, но они не требуют наличия ученой степени.

Кроме того, даже для более исчерпывающих услуг, таких как редактирование содержания, нет необходимости в использовании эксперта в определенной подобласти. Нет необходимости в глубоком понимании того, что говорится, чтобы знать, правильно ли это сказано на оригинальном английском языке. Если цель состоит в том, чтобы сделать исправления только в языке, то нет необходимости, чтобы кто-то сделал замечания по научному содержанию. Действительно, тот, кто имеет только общее представление об области, более вероятно, может распознать ненужный жаргон, определить

разделы, которые не совсем понятны, и сделать текст более читабельным для обычной аудитории. Даже на уровне редактирования содержания (контента) работа, выполненная службами подготовки к печати, похожа на ту, которая осуществляется руководителями диссертаций, когда аспиранты, готовящиеся получить ученую степень по медицине или философии, прежде всего учатся писать научные статьи, и руководители, по установившейся практике, включаются в качестве соавторов, иногда даже при наличии небольшого вклада.

Некоторые высококачественные службы подготовки к печати, описанные выше, действительно гарантируют, что работа будет выполнена специалистом в определенной подобласти, имеющим не только опыт опубликования в данной подобласти, но и опыт опубликования высококачественной работы в элитных журналах. Использование такого лица полностью изменяет представление относительно простого чтения корректуры или литературного редактирования статьи, что обсуждалось выше, и на самом деле не является необходимым для редактирования основного содержания. Однако, как четко объясняют некоторые эти службы, такой человек лучше подходит для того, чтобы осуществлять в рукописи изменения другого типа, такие как рассмотрение литературы, разъяснение целесообразности исследования, комментарии относительно недавно разработанных методик, проведение статистического анализа и обсуждение применений исследования в определенной области. По многим стандартам эти типы вкладов в статью четко отвечают большинству критериев авторства и даже превышают их.

## ЗАКЛЮЧЕНИЕ

Было бы бессмысленно одобрять только определенные компании или только определенные уровни обслуживания. Что касается проблемы авторства или многоавторства, то нет способа оказывать давление на такие руководства. Национальные финансирующие агентства могут препятствовать использованию служб подготовки к печати, которые при длительном применении могут улучшить навыки написания статей, но при краткосрочном использовании могут только мешать исследователям в своей стране. Финансирующие агентства могут установить лимиты на то, какой бюджет может быть выделен на статью для использования служб подготовки к печати, но это будет трудно регулировать. Следовательно, руководства относительно использования служб подготовки к печати будут настолько же эффективными, как и руководства по авторству: полностью зависимыми от желания им следовать. Тем не менее, первый шаг состоит в том, чтобы признать существование потенциальной проблемы.

Службы подготовки к печати могут внести существенный вклад в интеллектуальное содержание статьи. Стратегии, позволяющие использование служб подготовки к печати, но отказывающиеся им (службам) в авторстве, являются несовместимыми со стандартным определением авторства. Такие стратегии на самом деле одобряют безымянное авторство, оба типа безымянного авторства. В эру, когда наука все в большей степени имеет международный и многоавторский характер, диспуты в отношении культуры и этики должны непременно возникать. Однако руководства, именно руководства, а не правила, и не только они, не должны навязываться, но также и ограничиваться в своих соответствующих законодательствах и областях. Кроме того, руководства журналов, учреждений, профессиональных обществ и

национальных академий относительно авторства часто различаются и не совместимы друг с другом, таким образом, невозможны единые решения для устранения этих конфликтов. В случае возникновения споров относительно включения в состав авторов службы подготовки к печати, они (споры) должны быть изучены в индивидуальном порядке, с учетом контекста и этических руководств, релевантных всем рассматриваемым сторонам: журнал, соответствующие учреждения и национальные регулирующие органы, финансирующие агентства, редакторы и научные культуры современных и возможных авторов. Наконец, комитеты по найму сотрудников и финансированию должны быть осведомлены о том, в какой степени международные кандидаты полагаются и будут продолжать полагаться на службы подготовки к печати. В настоящее время пространственный документ для публикации больше не основывается на способности писать.

*Благодарность.* Я благодарю коллег, которые проявили большую доброжелательность при обсуждении данных проблем. В подготовке статьи не использовались никакие услуги по чтению корректуры, литературному редактированию или совместному написанию. Однако помощь рецензентов, редакторов и литературных редакторов признается и оценивается. Я благодарю Библиотеку Тартуского университета за предоставление доступа к ее электронным фондам. Эта публикация значится под номером 1305 Эстонского центра эволюционной экологии – ECEE (рег. 80355697).

## Литература

1. *Write science right.* Editing service levels, 2012. - [http://www.writesciencerright.com/editing\\_service\\_levels.html](http://www.writesciencerright.com/editing_service_levels.html).
2. *Macmillan science communication.* About the service, 2012.- <http://www.mscediting.com/service/index.html>.
3. *Enago.* Rejection/resubmission editing package, 2012. - <http://www.enago.com/resubmissionediting.htm>.
4. *Manuscriptedit.* Literature review, 2012. - <http://www.manuscriptedit.com/literature>.
5. *Manuscriptedit.* Customized Services, 2012. - <http://www.manuscriptedit.com/customizedservices>.
6. *Lindsey D.* Production and citation measures in the sociology of science: The problem of multiple authorship // *Social Studies of Science.* – 1980. – Vol. 10. – P. 145-162.
7. *Regalado A.* Multiauthor papers on the rise // *Science.* – 1995. – Vol. 268, No. 5207. – P. 25.
8. *Drenth J.P.H.* Proliferation of authors on research reports in medicine // *Science.* – 1996. – Vol. 2, No. 4. – P.469-480.
9. *Howard M.O., Walker R.D.* Multiple authorship: Trends over 50 years in the *Journal of Studies on Alcohol* // *Journal of Studies on Alcohol.* – 1996. – Vol. 57, No. 1. – P. 105-106.
10. *Erlen J.A., Siminoff L.A., Sereika S.M., Sutton L.B.* Multiple authorship: Issues and recommendations // *Journal of Professional Nursing.* – 1997. – Vol. 13, No.4. – P. 262-270.
11. *Quencer R.M.* Creeping authorship: Where do we draw the line? // *American Journal of Neuroradiology.* – 1998. – Vol. 19, No. 3. – P. 589.

12. Cronin B. Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices? // *Journal of the American Society for Information Science and Technology*. – 2001. – Vol. 52, No. 7. – P. 558-569.
13. Hama Y., Kusano S. Geographic variation in the number of authors on scientific abstracts // *Canadian Association of Radiologists Journal*. – 2001. – Vol. 52. – P. 25-28.
14. Rabman L., Muirhead-Allwood S.K. How many orthopedic surgeons does it take to write a research article? 50 years of authorship proliferation in and internationalization of the orthopedic surgery literature // *Orthopedics*. – 2010. – Vol. 33, No. 7. – P. 478. - doi:10.3928/01477447-20100526-06.
15. *International Committee of Medical Journal Editors*. Uniform requirements for manuscripts submitted to biomedical journals: Ethical considerations in the conduct and reporting of research: Authorship and contributorship, 2012. - [http://www.icmje.org/ethical\\_1author.html](http://www.icmje.org/ethical_1author.html).
16. Bennett D.M., Taylor D.M. Unethical practices in authorship of scientific papers// *Emergency Medicine*. – 2003. – Vol. 15, No. 3. – P. 263-270.
17. Logdberg L. Being the ghost in the machine: A medical ghostwriter's personal view// *PLoS Med*. – 2011. – Vol. 8, No. 8. - e1001071.
18. Matheson A. How industry uses the ICMJE guidelines to manipulate authorship –and how they should be revised // *PLoS Med*. – 2011. – Vol. 8, No.8. - e1001072.
19. *Council of Science Editors*. White paper on promoting integrity in scientific journal publications, 2012 updated (3rd revised ed.). - Wheat Ridge, Colorado, USA, 2012.
20. *National institutes of health*. Guidelines for the conduct of research in the intramural research program at NIH, 2007. - <http://sourcebook.od.nih.gov/ethicconduct/Conduct%20Research%206-11-07.pdf>.
21. Wager E. Do medical journals provide clear and consistent guidelines on authorship?// *Medscape General Medicine*. – 2007. – Vol. 9, No.3. - article number 16.
22. Bošnjak L., Marušić A. Prescribed practices of authorship: Review of codes of ethics from professional bodies and journal guidelines across disciplines // *Scientometrics*. – 2012. – Vol. 93, No. 3. – P. 751-763.
23. Marušić A., Bošnjak L., Jerončić A. A systematic review of research on the meaning, ethics and practices of authorship across scholarly disciplines // *PLoS One*. – 2011. – Vol. 6, No. 9. - e23477.
24. Rennie D., Flanagan A. Authorship! Authorship! Guests, ghosts, grafters, and the two-sided coin // *Journal of the American Medical Association*. – 1994. – Vol. 271, No. 6. – P. 469-471.
25. House M.C., Seeman J.I. Credit and authorship practices: Educational and environmental influences// *Accountability in Research*. – 2010. – Vol. 17, No. 5. – P. 223-256.
26. Seeman J.I., House M.C. Influences on authorship issues: An evaluation of receiving, not receiving, and rejecting credit // *Accountability in Research*. – 2010. – Vol. 17, No.4. – P. 176-197.
27. Wislar J.S., Flanagan A., Fontanarosa P.B., DeAngelis C.D. Honorary and ghost authorship in high impact biomedical journals: A cross sectional survey // *British Medical Journal*. – 2011. – Vol. 343, No. 7835. - d6128.
28. Smith E., Williams-Jones B. Authorship and responsibility in health sciences research: A review of procedures for fairly allocating authorship in multi-author studies // *Science and Engineering Ethics*. – 2012. – Vol. 18, No. 2. – P. 199-212.
29. Eisenberg R.L., Ngo L., Boiselle P.M., Bankier A.A. Honorary authorship in radiologic research articles: Assessment of frequency and associated factors// *Radiology*. – 2011. – Vol. 259, No.2. – P. 479-486. - doi:10.1148/radiol.11101500.
30. Drenth J.P.H. Multiple authorship. The contribution of senior authors// *Journal of the American medical association*. – 1998. – Vol. 280, No. 3. – P. 219-221.
31. Lacasse J.R., Leo J. Ghostwriting at elite academic medical centers in the United States// *PloS Medicine*. - 2010. – Vol. 7, No. 2. - e1000230. - doi:10.1371/journal.pmed.1000230.
32. Dance A. Authorship: Who's on first?// *Nature*. - 2012. – Vol. 489. – P. 591-593. - doi:10.1038/nj7417- 591a.
33. Lexchin J. Those who have the gold make the evidence: How the pharmaceutical industry biases the outcomes of clinical trials of medications // *Science and Engineering Ethics*. – 2012. – Vol. 18, No.2. – P. 247-261.
34. Jacobs A., Wager E. European Medical Writers Association (EMWA) guidelines on the role of medical writers in developing peer-reviewed publications// *Current Medical Research and Opinion*. – 2005. – Vol. 21, No. 2. – P. 317-321.
35. Rennie D., Yank V., Emanuel L. When authorship fails: A proposal to make contributors accountable // *Journal of the American Medical Association*. – 1997. – Vol. 278, No. 7. – P. 579-585.
36. Smith R. Authorship: Time for a paradigm shift? The authorship system is broken and may need a radical solution // *British Medical Journal*. – 1997. – Vol. 314, No. 7086. – P. 992.
37. Ilakovic V., Fister K., Marusic M., Marusic A. Reliability of disclosure forms of authors' contributions // *Canadian Medical Association Journal*. – 2007. – Vol. 176, No. 1. – P. 41-46.
38. Goodman N.W. Survey of fulfilment of criteria of authorship in published medical research // *British Medical Journal*. – 1994. – Vol. 309, No. 1482.
39. Harzing A.W. The publish or perish book: Your guide to effective and responsible citation analysis (1ed.). - Melbourne: Tarma Software Research Pty Ltd., 2010.
40. Pöder E. Let's correct that small mistake // *Journal of the American Society for Information Science and Technology*. – 2019. – Vol. 61, No. 12. – P. 2593-2594.
41. Papatheodory S.I., Trikalinos T.A., Ioannidis J.P.A. Inflated numbers of authors over time have not been just due to increasing research complexity // *Journal of Clinical Epidemiology*. – 2008. – Vol. 61, No. 6. – P. 546-551.

42. *Watson J.D., Crick F.H.C.* A structure for deoxyribose nucleic acid // *Nature*. - 1953. - Vol. 17, No. 4356. - P. 737-738.
43. *Epstein R.J.* Six authors in search of a citation: Villains or victims of the Vancouver convention? // *British Medical Journal*. - 1993. - Vol. 306, No. 6880. - P. 765-767.
44. *Bebeau M.J., Monson V.* Authorship and publication practices in the social sciences: Historical reflections on current practices // *Science and Engineering Ethics*. - 2011. - Vol. 17, No. 2. - P. 365-388.
45. van Loon A. J. Pseudo-authorship // *Nature*. - 1997. - Vol. 389, No. 4. - P. 11.
46. *Engelder T.* The coupling between devaluation of writing in scientific authorship and inflation of citation indices // *GSA Today*. - 2007. - Vol. 17, No. 7. - P. 44-45.
47. *Ioannidis J.P.A.* Measuring co-authorship and networking-adjusted scientific impact // *PloS One*. - 2008. - Vol. 3, No. 7. - art no. e2778.
48. *Spier R.* The history of the peer-review process // *Trends in Biotechnology*. - 2002. - Vol. 20, No. 8. - P. 357-358. - doi:10.1016/s0167-7799(02)01985-6.
49. *Lozano G.A.* A new criterion for allocating research funds: "Impact per dollar" // *Current Science*. - 2010. - Vol. 99, No. 9. - P. 1187-1188.
50. *Zhao S.X., Ye F.Y.* h-efficiency: Measuring input-output performance of research funds // *Current Science*. - 2011. - Vol. 101, No. 1. - P. 21-22.
51. *Eggert L.D.* Best practices for allocating appropriate credit and responsibility to authors of multi-authored articles // *Frontiers in Psychology*. - 2011. - Vol. 2. - P. 196. - doi:10.3389/fpsyg.2011.00196.

# Соавторство и цитирование в научном издательстве\*

**Тревис МАРТИН**

**(Travis MARTIN)**

Отделение электротехники и вычислительной техники, Мичиганский университет, шт. Мичиган, г. Анн-Арбор, США

**Брайан БОЛЛ**

**(Brian BALL),**

**Брайан КАРРЕР**

**(Brian KARRER)**

Отделение физики, Мичиганский университет, шт. Мичиган, г. Анн-Арбор, США

**М. Е. Дж. НЬЮМАН**

**(M. E. J. NEWMAN)**

Центр исследования сложных систем, Мичиганский университет, шт. Мичиган, г. Анн-Арбор, США

*Большое количество опубликованных исследований изучало свойства либо сетей цитирования среди научных статей, либо сетей соавторства среди ученых. В этой статье, используя расширенный массив данных, охватывающий статьи по физике, опубликованные в журналах Physical Review более чем за век, изучается гибридная сеть соавторства/цитирования, содержащая оба компонента, которые анализируются, чтобы получить понимание корреляций и взаимодействий между авторством и цитированием. Среди прочего рассматривается степень, с которой индивидуумы больше стремятся ссылаться на самих себя или своих коллег, чем на других лиц, степень, с которой они быстрее цитируют самих себя или своих коллег после публикации, и степень, с которой они стремятся выразить благодарность посредством ссылки на другого ученого.*

## ВВЕДЕНИЕ

Сети цитирования [1] и сети соавторства [2-4] являются отдельными сетевыми представлениями совокупности научной литературы, будучи предметом количественного анализа в последние годы. В сети цитирования узлами сети служат статьи, а ориентированный конец растягивается от статьи А к статье В, если А цитирует В в своей библиографии. В сети соавторства узлами являются авторы, а неориентированный конец связывает двух авторов, если они написали статью вместе. Оба типа сети могут пролить свет на традиции и модели научного исследования. Сети цитирования, например, могут дать представление о тематических связях между статьями, тогда как сети соавторства укажут на модели сотрудничества, такие как размер сотрудничающих групп или частота повторного сотрудничества.

Однако существует также много интересных вопросов, на которые можно дать ответ только с помощью сочетания данных цитирования и соавторства, вопросы, касающиеся того, как много исследователей цитируют своих коллег по сравнению с другими учеными в своей области, или должен ли ученый, более вероятно, цитиро-

вать других ученых, от которых он предварительно получил ссылку. Сегодня имеется сравнительно мало исследований относительно подобных вопросов [5], и это – те вопросы, на которых фокусируется данная статья.

Мы анализировали большой массив данных, доступ к которому предоставлен Американским физическим обществом (APS), он (массив) содержит библиографические данные и данные цитирования для журнала *Physical Review*\*. *Physical Review* является семейством журналов, управляемым APS и постоянно публикующим в течение более века статьи, охватывающие все аспекты физики. Массив данных, анализируемый нами, распространяется от начала выхода журналов в 1893 г. вплоть до 2009 г. и описывает почти полмиллиона статей, включая их авторство и ссылки между ними. Был опубликован ряд предыдущих анализов этих данных [6-8], но наше исследование принимает несколько отличную от тех исследований точку зрения, концентрируясь на взаимодействии между авторством и ссылками. Помимо прочего мы убеждаемся, например, что ученые цитируют самих себя или статьи соавторов гораздо быстрее после их публикации, чем они цитируют других лиц; что авторы стремятся отблагодарить ссылкой другого

\* Перевод Martin T. et al. Coauthorship and citation in scientific publishing. — 2013. — <http://arxiv.org/pdf/1304.0473v1.pdf>

\* Подробнее о массиве данных можно узнать на сетевом сайте - <https://publish.aps.org/datasets>

автора, особенно предыдущего соавтора; что в противоположность некоторым недавним предположениям наличие общего соавтора, вероятно, не заставит двух авторов сотрудничать в будущем [9-11]; что со временем нет никакого роста в самоцитировании (по крайней мере, в рамках изучаемых нами журналов, это число остается приблизительно постоянным на уровне почти 20% всех ссылок на протяжении более века).

## МАССИВ ДАННЫХ

В своем необработанном виде исследуемый нами массив данных содержит записи для 462 090 статей, опубликованных в различных журналах *Physical Review*, каждая из которых определена уникальной цифровой меткой. Данные для каждой статьи включают название статьи, дату публикации, фамилии и место работы каждого автора, и список цифровых меток предыдущих цитируемых в *Physical Review* статей. Массив данных является необычным по двум причинам: длинный охваченный им временной период, который растягивается на 116 лет с 1893 по 2009 гг., и тот факт, что он (массив) включает данные цитирования, а, значит, позволяет нам сравнить модели соавторства со ссылками, по крайней мере, для той части сети цитирования, которая появляется в *Physical Review*, - ссылки на журналы и из журналов, не входящих в семейство журналов *Physical Review*, большинство которых не включены.

Тем не менее, перед проведением любого анализа необходимо преодолеть некоторые препятствия. Самым первым среди них является тот факт, что имя автора само по себе необязательно определяет его однозначно. Два автора могут иметь одинаковую фамилию, или один и тот же автор может обозначаться по-разному в различных публикациях (например, со средним инициалом или без). В отличие от некоторых журналов, таких как журналы Американского математического общества\*, *Physical Review* не поддерживает единые определители авторов, которые могут быть использованы для точного установления авторства. Поэтому, в качестве первого шага в анализе данных, мы обработали их, используя ряд способов устранения неоднозначности, с тем чтобы сделать вывод об авторской идентичности из имен авторов настолько точно, насколько это возможно. Детали процесса устранения неоднозначности приведены в Приложении.

Кроме того, мы произвели умеренный отбор данных, чтобы удалить резко выделяющиеся значения. Самым важным действием является удаление всех статей с 50 или более авторами, являющихся в основном новыми публикациями в области экспериментальной физики высоких энергий. (Почти все они, около 91%, были опубликованы либо в *Physical Review D*, охватывающим область физики высоких энергий, или в *Physical Review Letters*, либо в *Physical Review C*, который распространяется на ядерную физику.) Как мы сжато показываем, хотя статьи с более 50 авторами являются только небольшой частью целого (около 0,7%), их включение значительно искажает результаты почти 30 последних лет по сравнению с остальным периодом времени. Что касается результатов, которые зависят от присутствия или отсутствия этих статей, то мы для сравнения приводим результаты, как с учетом этих статей, так и без них.

\* Описание системы определения уникального автора, используемой Американским математическим обществом, можно посмотреть на сетевом сайте – <http://www.isdl.org/01-summer/databases.html>.

В табл. 1 приводятся некоторые основные параметры окончательного массива данных.

Таблица 1

### Средние значения некоторой статистики для нашего массива данных с учетом статей, имеющих свыше 50 авторов, и без

	Все статьи	Статьи с 50 авторами или меньше
Всего статей	460889	457516
Всего авторов	235533	226641
Число авторов на статью	5,35	3,34
Число ссылок на статью	10,16	10,16
Число соавторов	59,44	17,24
Число статей на автора	10,47	6,74

## АНАЛИЗ

В следующих нескольких разделах представлено разнообразие видов анализов массива данных по *Physical Review*. Они начинаются с ознакомления с некоторыми основными параметрами авторства и соавторства.

### Модели авторства

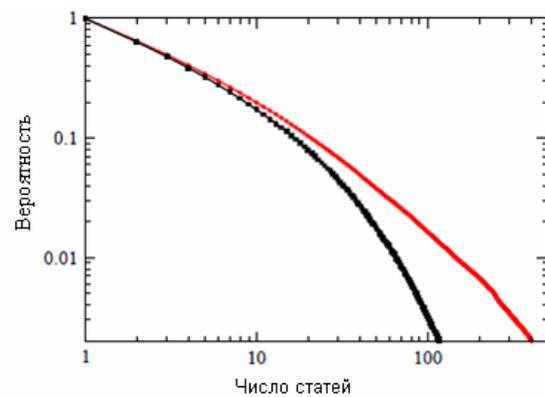


Рис.1 Вероятность того, что автор написал больше определенного числа статей.

Кружочки (верхняя кривая) показывают значения, вычисленные из полного массива данных; квадратик (нижняя кривая) являются значениями, полученными из массива данных после удаления статей с 50 и более авторами. Граф прерывается вблизи отметки 500 статей, так как за этой точкой имеется очень мало данных.

На рис. 1 представлена кумулятивная функция распределения для ряда статей, опубликованных одним автором, усредненная для целого массива данных. То есть на рис. 1 показана доля авторов, опубликовавших  $n$  статей или больше, как функция от  $n$ , которая является неточным измерением научной эффективности. Оси на рис. 1 являются логарифмическими, а приближенная прямолинейная форма распределительной функции означает, что научная эффективность следует, грубо говоря, степенному закону, результату, известному как закон Лотки, впервые наблюдаемому Альфредом Лоткой в 1926 г. [12] и подтвержденному после него множеством других ученых. (Также предполагалось, что это распределение является скорее логарифмически нормаль-

ным, чем степенным законом [13]. Известно, что эмпирически трудно отличить логарифмически нормальное распределение и распределение степенного закона [14].) На рис. 1 даются отдельные кривые с учетом и без учета статей, имеющих 50 или более соавторов. Как показывает рис. 1, различие между этими кривыми находится в основном в хвосте распределения, среди авторов, опубликовавших самое большое число статей, показывая, что значительная доля наиболее продуктивных авторов задействована в крупных соавторствах. Действительно, если составить список из 50 авторов, опубликовавших самое большое число статей, то только один из них останется в этом списке после исключения статей с 50 или более авторами. Это, вероятно, вытекает из сочетания двух эффектов: во-первых, большие группы могут публиковать больше статей просто потому, что они имеют больше людей, доступных для их написания; и, во-вторых, большая и эффективная группа коллег вносит свой вклад в статистику посредством наличия многих несомненно продуктивных авторов - каждый из соавторов получает репутацию высокоэффективного. Именно из-за предубеждений такого рода мы исключаем статьи со многими авторами из некоторых наших подсчетов.

Мы можем в некоторой степени устранить эту проблему путем измерения эффективности в более сложной манере. Вместо подсчета всех статей, внесенных в список автором, мы разделим авторство статьи между сотрудничающими авторами таким образом, чтобы, например, каждому из двух авторов была приписана половина авторства этой статьи. Это значительно сокращает влияние крупных соавторств на статистику, хотя распределение числа авторизованных статей все еще очень искажено из-за некоторых авторов, вносящих в науку больше, чем другие. Общим способом визуализации таких искаженных распределений является использование кривой Лоренца, графа доли статей, созданных самыми продуктивными авторами, в противоположность доли создавших их авторов. Такая кривая изображена для нашего массива данных на рис. 2, и крутой подъем этой кривой в левой части рисунка показывает на сосредоточение научной эффективности среди самых продуктивных ученых. Заметим, например, что продуктивность, кажется, должна приблизительно следовать так называемому правилу 80-20, состоящему в том, что почти 80% результата создается 20% самых продуктивных авторов. Отметим также, что почти нет различий в кривых Лоренца с учетом статей с 50 или более авторами и без него именно потому, что мы разделили авторство, с тем чтобы эффект статей со многими авторами был уменьшен.

Это распределение может быть в дальнейшем определено количественно путем измерения коэффициента Гини, который определяется, как избыточная площадь под кривой Лоренца, сравниваемой со случаем, когда каждый имеет точно такую же продуктивность. В нашем массиве данных коэффициент Гини равен 0,70, что является относительно большой цифрой для действующих коэффициентов, показывая высокое искажение. (Например, коэффициенты Гини для имущественного неравенства, являющегося контекстом, в котором такие коэффициенты, вероятно, хорошо известны, редко поднимаются выше 0,60.)

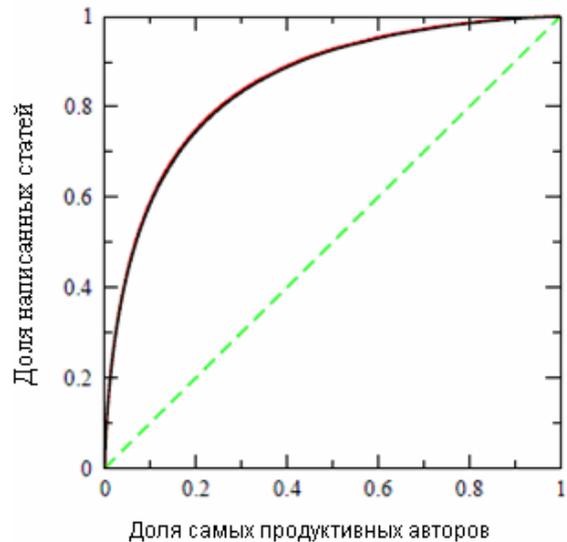


Рис. 2. Доля статей, написанных самыми продуктивными авторами (с учетом значимости много-авторских статей, разделенной между соавторами, как описано в тексте). Одна кривая представляет значения, вычисленные из общего массива данных; другая кривая представляет значения после удаления статей с 50 или более авторами. Отметим, что две кривые являются почти неразличимыми. Пунктирная линия показывает форму кривой, которую бы она имела, если бы все авторы опубликовали одно и то же число статей.

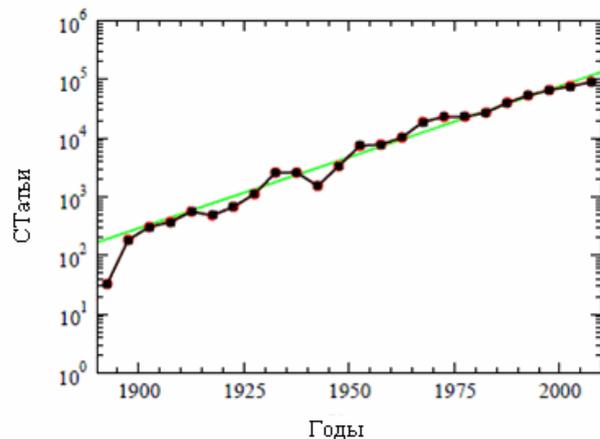


Рис. 3. Число статей, опубликованных в каждом пятилетнем периоде. Кружочки показывают числа, подсчитанные из полного массива данных; квадратикки означают числа, подсчитанные из массива данных после удаления статей с 50 или более авторами. Отметим, что два значения почти неразличимы. Прямая линия наилучшим образом соответствует экспоненте.

Этот массив данных также позволяет нам измерить продуктивность области физики в целом за определенное время, что не может быть сделано с помощью многих других массивов данных. На рис. 3 показано общее число статей, опубликованных в *Physical Review* за пятилетние периоды, начиная с 1893 г. Учитывая важное предостережение, что эти результаты предназначены только для одного массива журналов, и, более того, тая роль внутри области эволюционирует на протяжении его истории — от локального, быстро достигшего успеха журнала, до одного из ведущих изданий по физике на планете, мы видим, что имеется устойчивый рост объема опубликованной работы, который, кажется, должен приблизительно следовать экспоненциальному за-

кону (прямая линия на полулогарифмической шкале рисунка). Интересной особенностью является то, что за снижением кривой в 1940-х гг., которое совпадает со Второй мировой войной, следует ее (кривой) восстановление в 1950-х гг., возможно приписываемое отчасти растущему финансированию науки в послевоенный период. Однако общий результат этих отклонений должен только вернуть кривую назад, на тот же самый путь экспоненциального роста после войны, каким он уже был до нее. В своих ранних исследованиях вековых тенденций в сфере научных результатов Дерек де Солла Прайс [15,16] отмечал подобный экспоненциальный рост, прерванный войной, и измерял удвоенное время процесса роста, колеблющееся от 10 до 15 лет. Самое лучшее экспоненциальное соответствие нашим данным дает совместимая цифра в 11,8 лет.

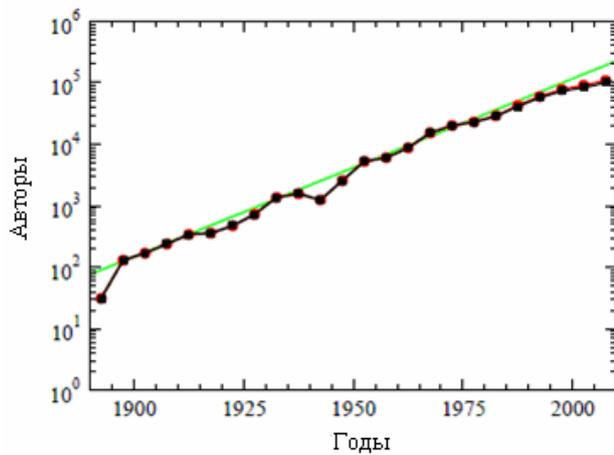


Рис. 4. Число уникальных авторов, опубликовавших статью в каждый пятилетний период. Кружочки показывают числа, подсчитанные из целого массива данных, тогда как обозначенные квадратиками числа подсчитаны из массива данных после удаления статей с 50 или более авторами. Отметим, что два значения почти неразличимы. Прямая линия является наилучшим соответствием экспоненте.

На рис. 4 приводится соответствующий граф числа уникальных авторов в массиве данных в каждый пятилетний период как функции времени. Подобно числу опубликованных статей, число авторов, по-видимому, растет экспоненциально и с примерно одинаковым (но слегка меньшим) временем удвоения в 10,4 лет. Таким образом, несмотря на отмеченный в целом рост продуктивности в области, оказывается, что каждый отдельный ученый создал приблизительно постоянное или даже слегка снижающееся число статей в год за все время.

Естественным дополнением к измерению числа статей на автора является измерение числа авторов на статью, т.е. размер сотрудничающих групп. На рис. 5 представлено среднее число авторов на статью в нашем массиве данных как функция времени, и существует определенная усиливающаяся тенденция на протяжении большей части охватываемого временного периода, со средним размером сотрудничающей группы, немного выросшей за прошедший век от 1 до почти 4 сегодня. Подобный эффект отмечался в предыдущих исследованиях, например Гроссмана и Йона [3] для случая математического сотрудничества. В наших подсчетах мы снова вычислили отдельные кривые с учетом статей, имеющих 50 или более авторов или без, и сравнение между ними выявляет удивительный эффект: в то время как почти совсем не существует отличия между кривыми

до 1975 г., имеется большой и быстро растущий пробел между ними в последующие годы. Без этих статей рост размера группы был медленным и устойчивым в течение десятилетий; с этими статьями он значительно отстает от исторической тенденции после 1970-х гг., показывая большую и растущую роль в области физики (или, по крайней мере, в публикациях по физике) для больших сотрудничеств.

Альтернативная точка зрения на ту же тенденцию показана на рис.6, который приводит число уникальных соавторов, которых имеет автор в течение пятилетнего отрезка времени. Подсчитывается каждый соавтор во временном периоде, даже если он также был подсчитан в предыдущем временном отрезке (но предыдущие соавторы не подсчитываются, если только они не являются также соавторами в новом временном отрезке). Как показывает рис. 6, это число также значительно возросло за последнее столетие, с чуть выше 1 до более 10 сегодня (и даже более 60, если включить сотрудничества с 50 или более членами). Так как мы имеем только данные из *Physical Review*, то вероятно, что мы опускаем некоторых сотрудников, поэтому эти числа на практике находятся только в нижних границах действительных чисел.

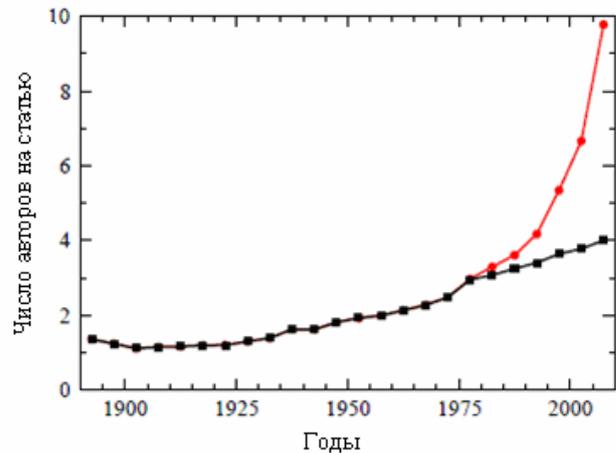


Рис. 5. Среднее число авторов на статью за пятилетний период. Кружочки показывают полный массив данных; квадратик представляет массив данных после удаления статей с 50 или более авторами.

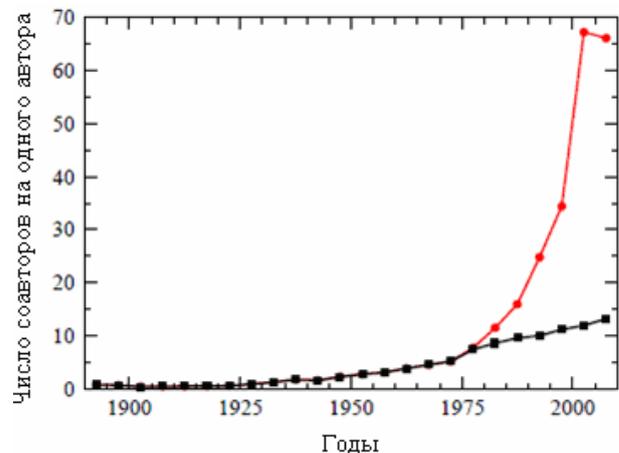


Рис. 6. Среднее число уникальных соавторов автора, усредненное в пятилетний период. Кружочки показывают полный массив данных; квадратик представляет массив данных после удаления статей с 50 или более авторами.

## Модели цитирования

Давайте теперь добавим долю ссылок массива данных к нашему анализу и исследуем модели цитирования в журнале *Physical Review* во времени, а также взаимодействия между ссылками и соавторством.

На рис. 7 дается среднее число ссылок в статье и на статью за временной период, охватывающий массив данных *Physical Review*. Кривая с квадратиками, число ссылок, которое создает статья, показывает устойчивый рост со временем - авторы обычно цитировали меньше статей, но в последние десятилетия они цитируют все больше. Одним возможным объяснением этого явления является рост объема доступной для цитирования литературы, хотя также предполагается, что авторы были под большим давлением в последние десятилетия, например, со стороны редакторов журналов или рецензентов, чтобы добавлять больше ссылок на статьи [17].

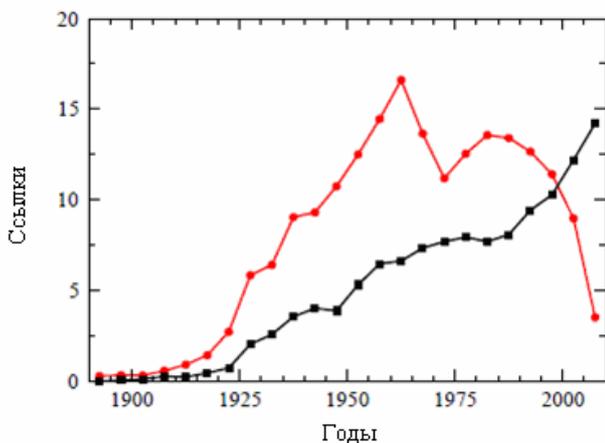


Рис. 7. Среднее число сделанных (квадратики) и полученных (кружочки) ссылок на статью по пятилетним периодам.

Кривая с кружочками на рис. 7 представляет среднее число ссылок, полученных статьей, которое показывает более изменчивое поведение, вырастая до пика дважды, перед тем как снизиться в последнее время. Здесь работает ряд эффектов. Во-первых, если (как мы вскоре увидим) большая часть ссылок приходится на статьи недавнего прошлого, тогда устойчивый рост ссылок в статьях должен привести к росту ссылок на статьи, опубликованные совсем недавно. Поведение такого рода наблюдалось в предыдущих исследованиях, таких как всестороннее исследование Валласа и др., использующее данные из Web of Science [18]. Однако этот рост числа полученных ссылок не может продолжаться до самого предела массива данных, так как большинство недавних статей является слишком новыми, чтобы получить значительное число ссылок и, значит, мы ожидаем падения в самом правом конце кривой, как показано на рис. 7.

Однако, имеется также заметное падение в этой кривой около 1970 г., происхождение которого представляется менее ясным. (Его не видно, например, в работе Валласа и др.) Исследуя данные за этот период подробно, мы полагаем, что падение числа ссылок на статью обусловлено в основном ростом числа статей, опубликованных в *Physical Review* (которое значительно расширилось в течение этого периода), хотя число ссылок, полученных этими статьями, в совокупности остаются приблизительно постоянными. Рост опубликованных статей может быть отчасти ответом на общее расширение исследований по физике в США в течение 1960-х гг., последовавшее за учреждением фонда National Science Foundation, но эти данные показывают, что боль-

ший объем исследований не привел, по крайней мере на начальной стадии, к большему числу полученных ссылок, и значит их соотношение делает наглядным на рис. 7 это падение кривой. Тем не менее, восходящее направление кривой восстанавливает ее положение почти с 1970 г., предполагая, что в конечном счете был рост не только числа опубликованных статей, но и числа статей, являющихся достаточно влиятельными для их последующего цитирования.

Интересно сравнить данные для полученных ссылок с прогнозами теоретических моделей для процесса цитирования. Возможно, самым известным классом моделей являются предпочтительные прикладные модели [19] и особенно модель Прайса 1976 г. [20], простая модель, в которой скорость, с которой статья получает ссылки, предположительно, должна изменяться линейно относительно уже имеющегося числа. В своем самом простом применении эта модель делает прогнозы, которые сильно отличаются от наблюдений, изображенных на рис. 7. Эта модель предсказывает, что наибольшее число ссылок должно относиться к самым старым статьям, а наименьшее - к самым новым, поэтому кривая с кружочками на рисунке монотонно снижается. Имеется ряд возможных объяснений такого несоответствия. Популярная теория состоит в том, что статьи «стареют» со временем, становясь менее цитируемыми по мере старения [21,22], это, вероятно, происходит потому, что их область сдвигается к другим аспектам, так как они вытесняются более продвинутой или точной работой, или потому, что их результаты настолько хорошо известны, что авторы больше не чувствуют необходимости в их цитировании. Если бы это было так, то большинство ссылок пришлось бы на недавние статьи, и кривая полученных ссылок, главным образом, отображала бы кривую данных ссылок, хотя с промежутком времени, длина которого определялась бы скоростью, с которой статьи стареют. Альтернативная теория, для которой имеется некоторое эмпирическое свидетельство, заключается в том, что предпочтительные прикладные модели представляют модели ссылок вполне достойно в рамках отдельных подобластей [23], но не тогда, когда применяются к литературе в целом. Центральным параметром в предпочтительных прикладных моделях является дата становления подобласти, а так как различные подобласти имеют различные даты становления, то эта модель, как ожидается, могла бы работать в рамках подобластей, а не целого массива данных.

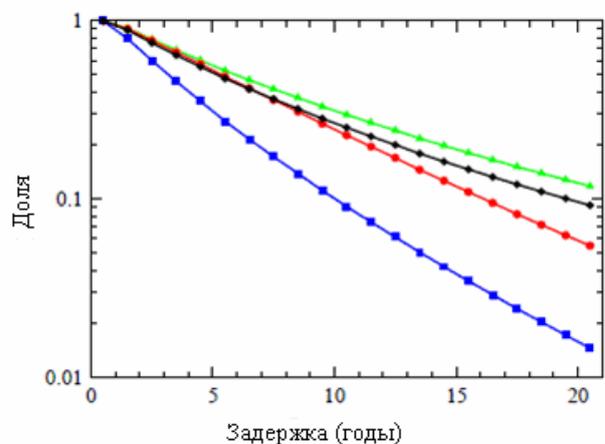


Рис. 8. Доля ссылок, сделанных в период, больший определенного числа лет после публикации.

Кривая с ромбиками включает все ссылки, кривая с квадратиками представляет самоцитирование, кривая с кружочками отражает ссылки соавторов и кривая с треугольниками показывает отдельные ссылки.

На рис. 8 анализируется старение статей в рамках массива данных *Physical Review* путем изображения доли ссылок, которые относятся к статьям определенного времени в прошлом. Давайте сосредоточимся в данный момент на кривой с ромбиками, которая включает все ссылки в целом массиве данных. Этот рисунок показывает, что действительно появляется сильный эффект старения, со скоростью цитирования, уменьшающейся приблизительно экспоненциально со временем (которой должна быть прямая линия на полулогарифмической шкале рисунка). Это наблюдение согласуется с предыдущими исследованиями старения [21], которые также обнаружили экспоненциальный спад. Альтернативной интерпретацией данных, однако, является то, что совсем не происходит старения, и что спад в ссылках является чисто механическим эффектом, который вытекает из спада литературы - в небольшой, молодой области имеется только несколько статей для цитирования, и значит каждая из них получает множество ссылок; в области постарше имеется много статей, и поэтому отдельные темпы цитирования ссылок падают. В той степени, в какой проводилось тестирование, последняя теория кажется хорошо согласуется с доступными данными цитирования, а также с предсказанием предпочтительных прикладных моделей [24], поэтому в настоящее время свидетельство в пользу (или против) старения в нашем массиве данных является неубедительным.

#### Взаимодействия между ссылками и соавторством

Возможно самым интересным аспектом данных *Physical Review*, тем не менее, является интервал, который он нам дает по взаимодействию между ссылками и соавторством. Одним способом доказать это взаимодействие является разделить ссылки в соответствии с общими ролями, взятыми на себя авторами цитирующих и цитируемых статей, и затем сравнить окончательные модели ссылок. В настоящей работе мы делим ссылки на три класса в соответствии с Валласом и др. [25]: самоцитирования, где цитирующие и цитируемые статьи делили, по крайней мере, одного соавтора; ссылки на соавторов, где, по крайней мере, один автор цитирующей статьи ранее сотрудничал с одним автором цитируемой статьи (но нет общих авторов между статьями, поэтому самоцитирования и ссылки на соавторов являются разобщенными); отдельные ссылки, которые включают все ссылки, отличные от самоцитирований и ссылок на соавторов. (Другие авторы, исследовавшие цитирование и сотрудничество, пошли дальше и рассмотрели также ссылки между соавторами соавторов [25], но это оказалось технически невыполнимым в настоящем случае из-за размера массива данных *Physical Review*.) Мы обращаем внимание на то, что рассматриваем в качестве соавторов только индивидуумов, если они предварительно были соавторами, когда встретилась ссылка. Соавторство, которое рождается после ссылки, не учитывается. Также наши данные ограничены *Physical Review*, поэтому число ссылок на соавторов будет в реальности выше, чем здесь, как из-за того, что некоторые ссылки отсутствуют в наших данных, так и из-за того, что некоторые соавторства присутствуют.

На рис. 9 показана доля ссылок, подпадающих под каждый из трех классов, как функция года издания цитирующей статьи. Грубо говоря, три кривые кажутся плоскими со временем. Имеется умеренный рост в доле соавторских ссылок (самая нижняя кривая с кружочками на рисунке), но это может объясняться ростом числа доступных для цитирования соавторов, показанных на

рис. 6, который характеризуется одинаковой протяженностью. Что касается других ссылок, то правило пальца, кажется, должно быть таким, что константа в 20% или около того является самоцитированиями, 75 или 80% являются отдельными ссылками, а небольшая оставшаяся доля относится к соавторам.

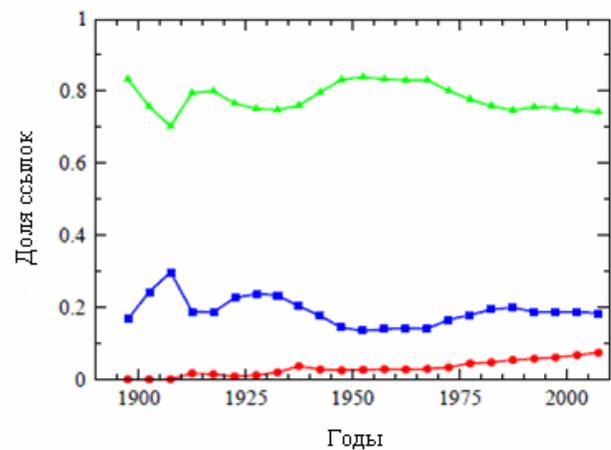


Рис. 9. Доля сделанных ссылок по типам в пятилетние периоды. Не было сделано ссылок в период 1890-1894 гг. Кривая с квадратиками представляет самоцитирования, кривая с кружочками обозначает ссылки на соавторов, а кривая с треугольниками показывает отдельные ссылки.

Распределение времени между датами публикации новой статьи и статей, которые она цитирует, представлено для трех классов ссылок на рис. 8 в виде кривой с квадратиками, кривой с кружочками и кривой с треугольниками. Здесь мы отмечаем значительное отличие между классами. В частности, самоцитирования падают быстрее, чем ссылки на соавторов или отдельные ссылки. Это означает, что большая доля самоцитирований встречается быстрее после публикации по сравнению со ссылками в других классах. Это не является неожиданным, учитывая, что ученый предположительно узнает о своем собственном исследовании быстрее и более подробно, чем о других исследованиях. Заметим также, что соавторские ссылки являются чуть более ранними, чем отдельные ссылки, что снова кажется разумным. Тем не менее, нужно быть осторожным в интерпретации этих результатов. Альтернативным объяснением тех же наблюдений является то, что статья может еще долго цитироваться другими учеными после ухода или оставления автором области, что может сделать среднее замедление для ссылок, осуществленных другими, более длительным, чем для самоцитирований. Нельзя сказать только из-за задержки самой статистики, какое объяснение является лучшим.

В табл. 2 суммируется средняя задержка ссылок для трех классов. Мы изучаем отличия между классами ссылок в следующем разделе.

Таблица 2

#### Среднее время задержки между датой публикации статьи и датами статей, ее цитирующих

Тип цитирования	Средняя задержка (годы)
Самоцитирования	4,12
Ссылки соавторов	6,92
Отдельные ссылки	9,02
Всего ссылок	7,89

Таблица 3

Доля (%) статей, которые делают или получают, по крайней мере, одну ссылку определенного типа

Тип цитирования	Сделано (%)	Получено (%)
Самоцитирования	68,9	60,3
Ссылки соавторов	42,0	31,3
Оба типа	35,6	26,3
Один из двух типов	75,0	64,2
Определенно оба типа возможны	76,4	66,4

Рассмотрим табл. 3, представляющую доли статей (%), которые делают или получают, по крайней мере, одно самоцитирование или одну ссылку соавторства при условии возможности такой ссылки. Около 70% статей цитируют, по крайней мере, одну статью одного и того же автора (или одного из тех же самых авторов, если их несколько), а 60% из них получают такую ссылку. Эти цифры могут на первый взгляд показаться большими, и, учитывая использование подсчета ссылок как меру влияния, вызвать беспокойство относительно того, что авторы могут раздуть свои подсчеты самоцитированием [26, 27]. Но, принимая во внимание тот факт, что число ссылок на статью и доля, в которую входят самоцитирования, являются значительного размера, эти большие числа не выглядят неожиданными. Рис. 9 показывает, что общее самоцитирование остается постоянным и умеренным, около 20%, и что нет недавно имевшего место изрядного излишка в самоцитировании.

Более интересный вопрос состоит в том, существует ли среди исследователей тенденция взаимного обмена ссылками в целях благодарности. Если автор  $A$  цитирует статью автора  $B$ , платит ли той же монетой автор  $B$ , позднее цитируя автора  $A$ ? Чтобы изучить этот вопрос, мы измеряем долю ссылок одного автора на другого автора (исключая ссылки его собственных статей), которые появляются в одной или более поздних публикациях. Мы подсчитываем отдельные цифры для пар авторов, до этого написавших в соавторстве статью и не написавших, и приходим к выводу, что только 13, 5% ссылок между не соавторами оплачиваются взаимностью цитирования по возможности, тогда как производящее впечатление 43,8% ссылок между соавторами оплачиваются взаимностью. (Следует иметь в виду, что ни один автор не может совместить цитирующую и цитируемую статью ссылкой так, чтобы она считалась ссылкой на соавтора, а не самоцитированием.) Обе эти цифры являются очень большими по сравнению с ожидаемой взаимностью, если ссылки были сделаны унифицированно случайно, но это необязательно подразумевает в качестве благодарности взаимный обмен ссылками. Ссылка предположительно является показанием того, что две статьи попадают в одинаковую предметную область, и, таким образом, присутствие ссылки значительно увеличивает шансы того, что авторы работают в одной и той же области, а это в свою очередь увеличивает правдоподобие ссылки вообще и значит шансы взаимной ссылки. В случае предыдущих соавторов шансы работы в одной и той же области являются, вероятно, более высокими. К сожалению, в настоящий момент мы не имеем ни одной модели процесса цитирования, настолько подробной, чтобы сделать количественное предсказание размера этого эффекта, на фоне которого мы могли бы сравнить наши измерения, чтобы проверить их на значимость.

## Транзитивность

Транзитивность в контексте сетей относится к наблюдению того, что «друг моего друга является также моим другом» [28]. В контексте соавторства, например, она наблюдается так, что если  $A$  выступает соавтором в статье с  $B$ , а  $B$  – с  $C$ , тогда  $A$  и  $C$  наиболее вероятно также должны являться соавторами статьи. Можно определить так называемый коэффициент кластеризации, который количественно определяет этот эффект, измеряя среднюю вероятность того, что друг вашего друга является также вашим другом [29], и такие коэффициенты измеряются во многих сетях [10, 30-32]. Как правило, можно прийти к выводу, что эти значения значительно выше ожидаемых, если сетевые связи были сделаны исключительно наугад, и наша сеть соавторства не является исключением. Что касается исследуемого здесь массива данных, мы находим коэффициент кластеризации 0, 212, который сравним с другими цифрами, сообщаемыми для сетей соавторства [4].

Однако в этом случае характер массива данных позволяет нам идти дальше. Обычное объяснение высокой транзитивности в сетях полагается на механизм тройного смыкания, согласно которому два автора, имеющие общего соавтора, наиболее вероятно должны будут сотрудничать в будущем возможно потому, что они встречаются в одних и тех же кругах, посещают одни и те же конференции, работают в одной и той же организации или представлены друг другу их общим знакомым [9-11]. Характер настоящего массива с временным разрешением данных позволяет нам непосредственно протестировать это предположение. Мы можем подсчитать, какую долю времени индивидуумы, имеющие общего соавтора, но которые до этого не сотрудничали вместе, тратят на совместное написание статьи. Когда мы делаем это измерение для данных *Physical Review*, мы приходим к выводу, что доля пар таких авторов составит только 0,0345 - гораздо меньшая доля, чем приведенный ранее коэффициент кластеризации целой сети. Причиной этой маленькой цифры является то, что большая доля транзитивности, просматриваемая в сетях сотрудничества, происходит из статей с тремя и более авторами, которые автоматически способствуют более замкнутым триадам узлов в сетях соавторства. Однако такие триады исключены из нашего подсчета вероятности более позднего сотрудничества. Большое различие между двумя вероятностями, которые мы подсчитываем, подразумевает, что только небольшая доля сетевой транзитивности происходит из истинных процессов сближения триад.

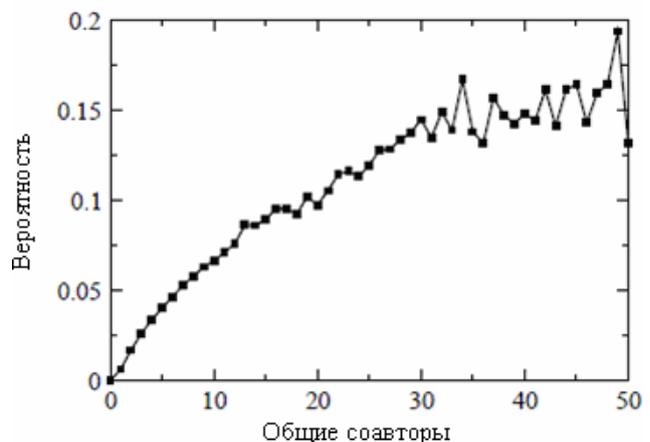


Рис. 10. Вероятность будущего соавторства с другим автором как функция числа общих соавторов. Число общих соавторов подсчитано ко времени первого соавторства или дате любой последней опубликованной статьи соавтора, когда она появляется впервые.

Тем не менее, процесс сближения триад, по-видимому, должен быть представлен в нашем массиве данных. На рис. 10 показана вероятность будущего соавторства между двумя индивидуумами как функция числа их общих соавторов, и мы видим, что эта вероятность быстро растет, этот вывод не противоречит предыдущим результатам [9, 33].

## ЗАКЛЮЧЕНИЕ

В этой статье мы анализировали большой массив данных из семейства журналов *Physical Review* с точки зрения сети. Чтобы не фокусироваться только на сетях цитирования или сетях соавторства, как было в большей части предыдущих исследований, мы объединили обе сети, что позволяет нам исследовать вопросы относительно способов, когда люди - а не только статьи - цитируют друг друга, и степени, с которой ученые сотрудничают с теми, кого они цитируют, или цитируют тех, с кем они сотрудничают. Временная протяженность массива данных является необычно большой, охватывающей публикации, вышедшие более чем за вековой период, что позволяет нам исследовать долговременные изменения в сотрудничестве и моделях цитирования, которые не доступны меньшим массивам данных.

Нашим основным наблюдением является то, что *Physical Review*, по-видимому, должен расти экспоненциально, с удвоением чуть меньше 12 лет и ростом числа ссылок на статью в рамках журналов. Доля самоцитирования и ссылок среди соавторов является более или менее постоянной со временем, и авторы стремятся быстрее цитировать свои собственные статьи после публикации, чем их соавторы, которые, в свою очередь, цитируют быстрее, чем не соавторы. Мы наблюдаем сильную тенденцию к перекрестным ссылкам, ученые, цитирующие другого автора, часто получающего ссылку в свою очередь позднее, в особенности с высокими долями ссылок между соавторами. В противоположность некоторым предыдущим требованиям [9-11], тем не менее, имеется только небольшой эффект триадной близости в моделях соавторства; два ученых, имеющие общего соавтора, но никогда не сотрудничавшие вместе, будут менее вероятно сотрудничать в будущем - около 3,5%. Это число, тем не менее, гораздо выше, чем вероятность для двух наугад выбранных ученых, и более того оно (число) быстро возрастает с ростом числа общих соавторов.

Имеется много других вопросов, которые могут быть исследованы с помощью этого массива данных, необычно длинная временная протяженность и сочетание данных публикаций и ссылок открывает множество возможностей. Например, мы знаем, какие статьи публикуются и в каких различных журналах *Physical Review*, и, значит, мы имеем грубое измерение тем статей, которое позволит нам ответить на вопросы о том, как модели соавторства и цитирования варьируются между областями внутри физики. Мы могли также изучить географические вариации, воспользовавшись данными организационной принадлежности авторов [34]. Наш анализ длительных исторических направлений мог быть также расширен; для ученого, занимающегося историей физики в США, без сомнения, имеется много интересных характерных черт исторических событий, скрытых внутри этих данных. Этот массив данных также предлагает возможность проследить карьерный рост отдельных ученых, возможно за более длинные периоды времени, или найти исследование по отдельной теме. И наконец, любой из наших анализов может быть расширен массивами данных, охватывающими другие журна-

лы или области, отличные от физики, если и когда такие данные станут доступными. Все это будет превосходными темами для дальнейшего исследования.

*Благодарность.* Эта работа была частично профинансирована Национальном научном фондом (США) по гранту DMS-1107796 (BB and MEJN) и организацией NSF IGERT Fellowship (TM).

## ПРИЛОЖЕНИЕ: ОБРАБОТКА ДАННЫХ

Как упоминается в основном тексте, мы выполнили некоторую предварительную обработку сырых данных *Physical Review*, чтобы устранить неясности в авторских именах и удалить экстремальные значения. Это приложение описывает предпринятые шаги.

### *Устранение неясности авторских имен*

Данные были представлены в двух блоках: 1) список статей с соответствующей информацией, такой как авторы, места работы авторов, журнал и год публикации; 2) список ссылок, использующий уникальные определители статей, которые соответствуют записям в первом блоке. Тем не менее, нет уникальных определителей для авторов, являющихся совместимыми между статьями, затрудняя точно выраженную авторскую идентификацию. Не все авторы используют одинаковую форму своего имени в каждой публикации, и имеется много примеров различных ученых с одним и тем же именем. Поэтому, перед тем как использовать массив данных, мы предприняли попытку связать имена авторов с определенными людьми. Как и в предыдущей работе по устранению неясностей в именах авторов, наш процесс начинается с предположения, что каждое имя в каждой статье представляет отдельного индивидуума, затем вводятся критерии сходства авторов и предполагается, что авторы, которые являются достаточно похожими по этим критериям, должны быть одной и той же личностью. После завершения этого процесса устранения неясностей мы проверили вручную подмножество этих результатов, чтобы оценить долю ошибок для этого процесса, и пришли к выводу, что он выполняется хорошо. Подробности описываются ниже.

Наш подход полагается не только на сами имена авторов, чтобы определить сходство, но и на модели сотрудничества и организационную принадлежность, так как авторы с похожими именами, имеющие много одинаковых соавторов или работающие в одной и той же организации, наиболее вероятно будут одними и теми же лицами. Однако информация о месте работы, как и само имя автора, может быть неточной и противоречивой, поэтому нашим первым шагом является объединение места работы, которые считаются достаточно похожими. Мы измеряем сходство, используя вариант редакторского расстояния, применяемого к установлению текстовых строк организаций, внедренного с использованием библиотеки Python *difflib library*.

Если места работы были однажды обработаны таким образом, то при обработке имен авторов мы придерживаемся следующего:

1. Объединяем всех авторов с идентичными именами, которые относятся к одной и той же организации. Публиковаться под идентичными именами оказывается нетипичным для двух физиков из одного и того же университета, поэтому это кажется надежным шагом.

2. Находим пары авторов с похожими, но не идентичными именами. Нашим критерием сходства на этом этапе является то, что у авторов должны быть идентичными фамилии и совместимые первые/средние имена (т.е. идентичными, если написаны полностью,

или если сопоставимы инициалы, где они используются). Также авторы не должны публиковаться вместе в одной статье (что исключает, например, членов семьи с одинаковыми именами, которые публикуются вместе). Для всех пар с одинаковыми именами мы затем подсчитываем меру дальнейшего сходства, основанную на том, где они вместе работают, как много соавторов они имеют, идентичны ли их полные имена и публикуются ли они в одном и том же журнале. Авторы с достаточно высоким сходством объединяются, в первую очередь самые похожие пары.

Мы проверяли точность этого процесса путем составления наугад двух списков из его результатов, первый, содержащий 79 примеров, в котором авторы с похожими именами объединялись в одного автора, и второй, включающий 111 примеров, в котором этого не было. Затем мы вручную произвели «слепой» поиск – без знания выбора, сделанного алгоритмом, – для публично доступной интерактивной информации об именах, вызывающих сомнение, чтобы определить, действительно ли они представляют одних и тех же авторов или различных. Мы обнаружили ложную положительную долю в 3% (т.е. 3% пар неправильно оценены как одно и то же лицо, когда в реальности они являются разными лицами) и ошибочную отрицательную долю в 12%.

Мы также проверяли этот эффект на наших результатах процесса устранения неясностей путем подсчета статистических данных, приведенных в этой статье, как для противоречивых данных, так и необработанного массива данных до устранения противоречий, в котором мы наивно предполагали, что каждая строка уникального (единственного) автора представляет одного автора, а каждая пара авторов с одинаковой строкой является одним и тем же лицом. Мы находили реальные отличия между этими двумя результатами в ряде наиболее важных статистических данных, таких как общее число определенных авторов: это число составляло 328 938 в необработанном массиве данных, но снизилось до 235 533 после устранения неясностей. С другой стороны, еще одна статистика дала очень небольшие изменения, указывая на то, что они не являются особо чувствительными к деталям авторской идентификации. Например, коэффициент кластеризации изменяется от 0,222 в необработанном массиве данных до 0,212 в точном массиве данных.

Подбор данных

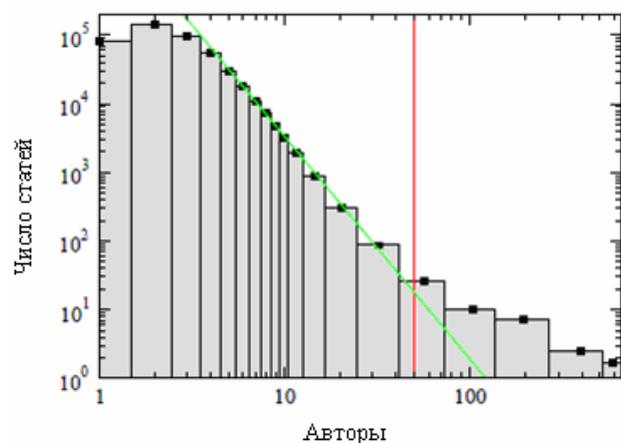


Рис. 11. Гистограмма статей с определенным числом авторов. Вертикальная линия падает на 50 авторов и грубо соответствует той точке, где распределение отклоняется от формы степенного закона, показанной соответствием. Данные для десяти и более авторов собраны логарифмически для минимизации статистических колебаний.

В дополнение к устранению неясностей в отношении авторов мы подбираем данные в соответствии с несколькими простыми правилами. Ряд статей в массиве данных не указывает авторов, в основном это редакционные статьи и другие, относящиеся к материально-техническому обеспечению производства, статьи без научного содержания. Они удаляются полностью. Как упоминается в тексте, мы также определяем все статьи с 50 или более соавторами, и многие из наших подсчетов выполнены в двух версиях, с учетом и без учета этих статей. Выбор 50 авторов в качестве объекта устранения был сделан путем тщательного изучения распределения числа авторов, показанного на рис. 11. Как показывает рисунок, ряд статей с определенным числом соавторов кажется, грубо говоря, следует степенному закону (в соответствии с некоторыми предыдущими [35], но не с другими [36] исследованиями), но имеется заметное отклонение от формы степенного закона для самых крупных чисел соавторов, около или выше 50, показывающее потенциально различные статистические законы в этом режиме и, возможно, различные соответствующие процессы сотрудничества.

Мы также удалили из данных небольшое число ссылок. В нескольких случаях статья приводится как цитирующая саму себя, что мы посчитали ошибкой. В ряде других случаев статьи цитируют другие статьи, которые были опубликованы в более позднее время, что искажает причинную связь. Они также считались ошибочными и удалялись. Наконец, данные показывают, что некоторые статьи цитировали какие-нибудь одни и те же статьи несколько раз в одной библиографии; такие множественные ссылки мы считали за одну ссылку.

## Литература

1. Price D. J. d. S. Networks of scientific papers// Science. — 1965. — Vol. 149. — P. 510–515.
2. Grossman J. W., Ion P. D. F. On a portion of the well-known collaboration graph// Congressus Numerantium. — 1995. — Vol. 108. — P. 129–131.
3. Grossman J.W. The evolution of the mathematical research collaboration graph// Congressus Numerantium. — 2002. — Vol.158. — P. 201–212.
4. Newman M.E.J. The structure of scientific collaboration networks// Proc. Nat. Acad. Sci. USA. — 2001. — Vol. 98. — P. 404–409.
5. Milojević S. How are academic age, productivity and collaboration related to citing behavior of researchers?// PLoS ONE. — 2012.— Vol.7, e49176.
6. Redner S. Citation statistics from 110 years of Physical Review// Physics Today. — 2005. — Vol. 58. — P. 49–54.
7. Chen P., Redner S. Community structure of the Physical Review citation network// Journal of Informetrics. — 2010. — Vol. 4. — P. 278–290.
8. Gualdi S., Medo M., Zhang Y.-C. Influence, originality and similarity in directed acyclic graphs// Europhysics Letters. — 2011. — Vol. 96. — No.18004.
9. Newman M. E. J. Clustering and preferential attachment in growing networks// Physical Review E. — 2011. — Vol. 64, 025102.
10. Davidsen J., Ebel H., Bornholdt S. Emergence of a small world from local interactions: Modeling acquaintance networks// Physical Review Letters. — 2002. — Vol. 88, 128701.

11. *Holme P., Kim B. J.* Growing scale-free networks with tunable clustering// *Physical Review E*. — 2002. — Vol.65, 026107.
12. *Lotka A.J.* The frequency distribution of scientific productivity// *Journal of the Washington Academy of Sciences*. — 1926. — Vol. 16. — P. 317-324.
13. *Shockley W.* On the statistics of individual variations of productivity in research laboratories// *Proceedings of the IRE*. — 1957. — Vol. 45. — P. 279-290.
14. *Clauset A., Shalizi C.R., Newman M.E.J.* Power-law distributions in empirical data// *SIAM Review*. — 2009. — Vol. 51. — P. 661-703.
15. *Price D. J. d. S.* *Science since Babylon*. — Yale University Press: New Haven, CT, 1961.
16. *Price D. J. d. S.* *Little Science, Big Science*. — New York, NY: Columbia University Press, 1963.
17. *Wilbite A.W., Fong E.A.* Coercive citation in academic publishing// *Science*. — 2012. — Vol. 335. — P. 542-543.
18. *Wallace M.L., Larivière V., Gingras Y.* Modeling a century of citation distributions// *Journal of Informetrics*. — 2009. — Vol. 3. — P. 296-303.
19. *Barabási A.-L., Albert R.* Emergence of scaling in random networks// *Science*. — 1999. — Vol. 286. — P. 509-512.
20. *Price D. J. d. S.* A general theory of bibliometric and other cumulative advantage processes// *Journal of the American Society for Information Science*. — 1976. — Vol. 27. — P. 292-306.
21. *Zhu H., Wang X., Zhu J.-Y.* Effect of aging on network structure// *Phys. Rev. E*. — 2003. — Vol. 68, 056121.
22. *Sanyal S.* Effect of citation patterns on network structure// 52nd Annual March Conference of the American Physical Society, Denver, CO, USA. — 2007.
23. *Newman M.E.J.* The first-mover advantage in scientific publication// *Europhys. Lett*. — 2009. — Vol. 86, 68001.
24. *Karrer B., Newman M.E. J.* Random graph models for directed acyclic networks// *Physical Review E*. — 2009. — Vol. 80, 046110.
25. *Wallace M.L., Larivière V., Gingras Y.* A small world of citations? The influence of collaboration networks on citation practices// *PLoS ONE*. — 2012. — Vol. 7, e33339.
26. *Hirsch J.E.* An index to quantify an individuals scientific research output// *Proceedings of the National Academy of Science*. — 2005. — Vol.102. — P. 16569-16572.
27. *Bartneck C., Kokkelmans S.* Detecting h-index manipulation through self-citation analysis// *Scientometrics*. — 2011. — Vol. 87. — P. 85-98.
28. *Wasserman S., Faust K.* *Social Network Analysis*. — Cambridge University Press, Cambridge, 1994.
29. *Watts D.J., Strogatz S.H.* Collective dynamics of 'small-world' networks// *Nature*. — 1998. — Vol. 393. — P. 440-442.
30. *Banks D.L., Carley K.M.* Models for network evolution// *Journal of Mathematical Sociology*. — 1996. — Vol. 21. — P. 173-196.
31. *Jin E.M., Girvan M., Newman M.E.J.* Structure of growing social networks// *Physical Review E*. — 2001. — Vol. 64, 046132.
32. *Newman M.E.J., Park J.* Why social networks are different from other types of networks// *Physical Review E*. — 2003. — Vol. 68, 036122.
33. *Bložnelis M., Kurauskas V.* Clustering function: A measure of social influence. — Preprint arXiv:1107.1155.— 2012.
34. *Pan R.K., Kaski K., Fortunato S.* World citation and collaboration networks: Uncovering the role of geography in science. — Preprint arXiv:1209.0781. — 2012.
35. *Newman M.E.J.* Scientific collaboration networks: I. Network construction and fundamental results// *Physical Review E*. — 2001. — Vol. 64, 016131.
36. *Hsu J.W., Huang D.W.* Distribution for the number of coauthors // *Physical Review E*. — 2009. — Vol. 80, 057101.

# Влияние публикаций в трудах на измерение ядра соавторов\*

**Януш МИСКЕВИЧ**  
**(Janusz MIŚKIEWICZ)**

Институт теоретической физики,  
Вроцлавский университет, Отделение  
физики и биофизики, Вроцлавский  
университет окружающей среды и  
наук о жизни, г. Вроцлав, Польша

*Соавторы (СА) «ведущего исследователя» (ВИ) могут получить ранг (r) в соответствии с их «важностью» при выпуске совместных с ВИ публикаций. Общепринято, без каких-либо доказательств, что публикации в рецензируемых журналах и, например, трудах конференций не имеют одинакового «значения» в послужном резюме (Curriculum Vitae – CV). То же самое относится к статьям в энциклопедиях и книжных главах. Здесь исследуется, сохраняется ли еще взаимосвязь между числом публикаций отдельного ученого (J) с его соавторами, ранжируемыми в соответствии с их уменьшающейся важностью, т.е.  $J \propto 1/r^a$ , как обнаружил Ауслус [1], все еще имеет силу, если общий список публикаций включает такие особые типы публикаций. Некоторые авторы с различными статусами, но в основном занятые в области статистической механики, изучаются здесь, чтобы получить ответы на эти вопросы. Показатель a становится слабо зависимым от ученого, только если максимальное значение J и r является большим и тогда равным  $\sim + 1$ . Значение ядра  $m_A$ , т.е. основное число соавторов, только в трудах составляет около половины всего значения, т.е. когда подсчитываются все публикации. Вклады в числовые значения как энциклопедий, так и книжных глав являются незначительными. Роль временного пробела на  $m_A$  также исследуется в двух случаях относительно рассмотрения карьерной активности. Можно полагать, что эти наблюдения служат в качестве противоположной точки зрения на то, как определить качество индивидуальной (публикационной) карьеры, что недавно проделано Петерсеном и др. [2-4], выделяя здесь скорее размер сотрудничества и развития, а не подсчет ссылок, более того, особо подчеркивая тип публикации. Через разные  $m_A$  можно выделить различные модели поведения научной публикации с соавторами.*

## ВВЕДЕНИЕ

Для подтверждения и/или продвижения работы молодого ученого он часто представляет свое исследование на научных встречах и публикует результаты исследования в трудах. Также некоторым образом общепринято без доказательств, что статьи в трудах содержат больше соавторов, чем в рецензируемых журналах. Здесь не ставится вопрос, следует ли рассматривать публикации в трудах и публикации в рецензируемых журналах равнозначными для измерения важности некоторого научного сообщения. В пользу опубликования материалов в трудах высту-

пает то, что это, кажется, единственный способ более быстро оправдать время, затраченное членом команды или лабораторией, так как труды конференций считаются менее строгими или занимают меньше времени рецензентов, чем хорошо устоявшиеся рецензируемые журналы. Обязательно, чтобы лидер команды или координатор лаборатории ассоциировался с такими публикациями, подтверждая свой статус главного исследователя (ГИ). Безусловно, число публикаций ГИ в таком случае очень увеличивается. Основной вопрос, который, кажется, должен возникнуть в отношении таких аспектов научной жизни, касается числа релевантных соавторов для множества публикации ГИ. Можно заинтересоваться количественной ролью ГИ или команды.

Определение «ядра соавторов», данное Ауслусом, и его последующее измерение [1] способствуют рассмотрению таких вопросов конструктивным способом через

\* Перевод Miśkiewicz J. Effects of publications in proceedings on the measure of the core size of coauthors. —2013.—  
<http://arxiv.org/pdf/1306.2604.pdf>

взаимосвязь между числом совместных публикаций ( $J$ ) с соавторами, ранжируемых в соответствии с рангом важности ( $r$ ). Этот подход сильно отличается от соображений времени карьерной жизни, роста и/или упадка, представленных Петерсеном и др. [2] и позднее расширенных в [3, 4]. Они уделяют внимание подсчету ссылок в рамках идеи, касающейся индекса Хирша [5,6], учитывающая некоторую нормализацию, основанную на размере группы и времени ее развития, однако в основном проверяя популярность статьи. Подход Ауслуса ставит акцент на роли *людей* в команде по производству научной продукции, что скорее касается ГИ, а не ссылок.

Фактически, ГИ, стоящего за публикацией или в самой публикации, иногда трудно определить. Известно, что понятие ГИ возникает из административных задач. Иногда соавтор, совсем необязательно получающий первое место в списке соавторов публикации, играет важную роль в научных исследованиях. Однако без этого критерия нельзя утверждать, что такой соавтор выполнил «больше работы», чем ГИ, или наоборот. Это является одной из основных причин, почему словосочетание ГИ будет впоследствии отклонено и заменено на ВИ, подразумевающее «ведущего исследователя», например, как лицо, принятое сообществом в качестве хорошо известного лидера для некоторого исследования.

Как правило, положение в списке соавторов публикации намекает на некоторую ответственность, но не всегда. Ожидаемое выделение положения ученого в списке публикации является деликатным вопросом при определении доли сотрудничества. Тем более из-за того, что считается, что вводящее в заблуждение соавторство и публикационный паразитизм существуют, как подчеркивает Квок [7], т.е. имеет место «эффект белого быка». Действительно, иногда большое количество так называемых статей из трудов или приглашенных лекций имеют много соавторов, как правило, для того, чтобы учитывать различные вклады в сотрудничество по рассматриваемому предмету и/или создавать видимость размера команды среди, вероятно, многих других причин.

С другой стороны, следует помнить, что индекс  $m_A$  [1] измеряет ядро соавторов в научной команде, как будто сконцентрированной на ученом, ВИ, которым может быть любой. Индекс  $m_A$  выводится из графа числа совместных публикаций ( $J$ ) этого ВИ с соавторами (СА), ранжируемыми в соответствии с их рангом важности ( $r$ );  $r=1$  является самым продуктивным соавтором с ВИ.

Ауслус [1] обнаружил простой степенной закон, связывающий  $J$  и  $r$

$$J \propto 1/r^\alpha. \quad (1)$$

Показатель  $\alpha$  степенного закона не равняется точно +1. Он, как правило, зависит от исследуемого диапазона данных, это хорошо известно [8]. Можно также предположить, что нарушения и отклонения от такого простого аналитического закона, уравнение (1), могут быть вызваны: обесцениванием публикаций, подсчетом трудов, раздуванием соавторства по какой-либо причине [9], даже полагая, что все подсчеты сделаны на надежной основе данных. Тем не менее, исходя из этой точки зрения можно получить индекс  $m_A$ , устанавливая диапазон ядра соавторов через условие, подобное тому, которое определяет  $b$ -индекс ученого [5,6], т.е.

$$m_A \equiv r, \text{ такому, что } r \leq J. \quad (2)$$

Наиболее вероятно, что закон Ауслуса, уравнение (1), кажется лучшим для больших команд и/или для авторов, имеющих много публикаций и много соавторов. В самом деле, как уже указывалось в [1], когда автор имеет мало публикаций или мало соавторов, этот закон

может быть статистически недостаточным описанием (редких) эмпирических данных. С другой стороны, отклонения при наличии большого множества публикаций и большого множества соавторов могут быть вызваны несколькими причинами.

Так называемые «внутренние причины» могут возникать из большой продуктивности группы, основанной на высоком привлечении молодых исследователей и студентов-старшекурсников, с  $r \gg 1$ , но имеющей несколько  $J$  с ВИ. С другим диапазоном  $r$ , т.е. для малых  $r$ , многие вклады в сотрудничество могут происходить от постоянных партнеров, которые обеспечивают представительность команды, посещая конференции и летние школы. Среди «внешних причин» можно просто упомянуть условия финансирования или введения ограничений по субсидированию, запрашиваемого скорее с тем, чтобы показать большой размер команды ВИ. Действительно, государственные и частные агентства по финансированию исследований предъявляют права на поиск и продвижение таких коллективов. Снова, оценка качества таких соавторов далека от ясности, не говоря об измерении их внутреннего влияния на команду и ВИ.

Поэтому интересно рассмотреть применимость уравнения (1). Это делается здесь путем разбиения полного списка публикаций на отдельные типы публикаций как в рецензируемых журналах, трудах, вступительных главах книг или энциклопедий, ... Здесь изучаются отдельные случаи ВИ. Как пример, два хорошо известных в статистической физике ВИ: Н.Е. Stanley (HES), самый продуктивный из таких авторов (с самым высоким, известным для физиков  $b$ -индексом,  $b > 115$ , хотя это irrelevantно для настоящей цели); и D. Stauffer (DS), признанный лидер в теоретической и цифровой статистической физике. Более того, M. Ausloos (MA), который придумал индекс  $m_A$  [1], включен, благодаря его большому списку публикаций с соавторами, в тот же качественный диапазон, что и DS, но со статьями гораздо меньше цитируемыми, чем статьи HES и DS.

Для сравнения, включены списки публикаций 6 ученых, работавших или все еще работающих в области статистической механики, хотя во Вроцлаве не требуется быть ГИ в англо-саксонском или даже американском статусе\*. Поскольку список публикаций для некоторых из них является «довольно коротким» по сравнению с HES, DS и MA, то подчеркивается, что эти данные даны в основном для охвата большого диапазона  $J$  и  $r$ . Однако их анализ подтверждает заявление Ауслуса, что существует что-то еще, кроме  $b$ -индекса, для измерения качества или полезности научного автора, т.е. его способности создания научной команды с релевантными соавторами.

Методология относительно отделения и подсчета различных типов публикаций кратко объясняется в разделе «Методология». Данные анализа характеристик соавторства приводятся в разделе «Анализ и обсуждение масштаба данных», делающим различие между всеми публикациями, подраздел «Численный анализ NJP», и публикациями в трудах - подраздел «Численный анализ NJPr». В разделе «Обсуждение» некоторое обсуждение статистических (механика) аспектов этих иллюстративных примеров представлено параллельно с общими предположениями относительно «подядер» соавторов для ВИ. Кроме того, изучается, следует ли рассматривать временную зависимость настоящих данных. Следуя двум временным интервалам, рассматриваемым HES для своего списка публика-

\*Автор настоящей статьи не претендует на то, чтобы называться ВИ.

ций, те же временные интервалы использовались для «сравнения» ядра соавторов HES и DS как функция временного интервала в разделе «Система времени». Раздел «Заключение» представляет собой предложение относительно дальнейшего исследования и применения.

## МЕТОДОЛОГИЯ

Как упоминалось во «Введении», чтобы определить качество закона Ауслуса и подтвердить его предел обоснованности в выше определенном контексте, были отобраны некоторые списки публикаций. С одной стороны, список публикаций с упомянутыми соавторами HES, DS и MA доступен через сетевые сайты, а список вроцлавских ученых – через личные контакты. С другой стороны, первые три являются достаточно большими, чтобы позволить ожидать приемлемую и значительную статистическую работу. Поскольку мой список (JM) и список вроцлавских ученых являются короткими, они определенно могут служить в качестве меры непостоянства результатов.

В каждом списке многие главы в книгах и энциклопедиях или статьи для последующей научной презентации на различных научных встречах, могут «легко» отличаться от статей в рецензированных журналах, в случае сомнения следует провести поиск, например, в «необычных» журналах, выясняя, «мотивировала» ли конференция публикацию через некоторые труды. Фактически отдельные труды могут появляться даже в выпусках научных журналов или как целые выпуски, которые совсем необязательно указываются издателем как «специальные номера». Иногда встречается некоторая неясность в отношении того, относится ли публикация к научному докладу, представленному на конференции, или является тщательно подготовленной статьей. При сомнении публикация включалась в список «трудов». Отметим, что книги, написанные, переведенные, или труды, отредактированные рассматриваемыми учеными, не подсчитывались здесь как релевантные данные.

Таким образом, статьи для энциклопедий и глав книг впервые классифицировались как особые группы (e) и (bc) (см. табл. 1)\*. Оказывается, что такие вклады в сотрудничество являются немногочисленными. Кроме того, проверка уравнения (1) на таких публикациях, как, например, HES (см. рис. 1) и MA (см. рис. 2), показывает, что числовой эффект NJPbc является незначительным вкладом в NJPr, что касается NJPbr, то возникает вопрос, используется или нет полный диапазон  $r$  или центральная область  $J(r)$ , близкая к  $m_{,b}$ , для соответствия степенному закону. Наблюдаем, что соответствующий показатель  $a$  приближается, но выше +1 в случае HES и приближается к +0,86 в случае MA. Более того, точность соответствия скорее является приемлемой в случае полного режима соответствия, т.е.  $R^2 \sim 0,873$  и  $\sim 0,981$  для случаев HES и MA, соответственно, выгодно показывающей легкое отклонение от идеального степенного закона в случае HES. Однако это соответствие является значительно улучшенным, т.е.  $R^2 \sim 0,98$  в случае HES, если используется только центральная область ( $r \leq 30$ ) (на рис. не показано). Это показано для MA на рис. 3:  $a$  поднимается к 0,88 с  $R^2$  некоторым образом снижающимся до  $\sim 0,96$ . Отклонения можно считать слабыми, но в обоих случаях  $\alpha \neq +1$ . Легко понять, что эффект масштаба, число точек данных, подразумевается в таких отклонениях. Случай других авторов не показан из-за

отсутствия места, но если бы и был показан, то не принес бы никакой отличной качественной информации.

Следовательно, было решено сгруппировать все публикации только в два множества, в общем сливая (bc) и (e) в (обобщенный) список «трудов» (p), т.е. NJPbc  $\rightarrow$  NJPr, в противоположность NJPj для списка рецензируемых журналов (j). Тем не менее, было доказано, что все они являются качественными научными статьями, иногда цитируемыми в работах, которые получают право регулярно приниматься в качестве «статей». Однако, например, после завершения анализа в разделе «Численный анализ NJPr» было обнаружено, что статья [10], которая, по-видимому, на первый взгляд должна быть обычной статьей в хорошо известном журнале, но (после ее прочтения) оказывается докладом на конференции и была неправильно помещена. Отметим, что при попытке устранить неясности предполагалось, что можно сделать различие между статьями, опубликованными либо в книге с твердой, либо с мягкой обложкой. Но журналы также обладают мягкой обложкой\*. Возможно, могут быть найдены другие подобные случаи. Нужно признаться, что нельзя прочитать все статьи, чтобы узнать, существует ли такое неправильное расположение. Тем не менее, большая часть анализируемых данных берется из глубоко проверенных баз данных. Отметим, что акцент был сделан на том, чтобы не подсчитать дважды идентичные научные публикации. В качестве образца, HES имеет много репродуцированных публикаций, например, в учебном пособии или в другом журнале. Более того, *Erratum* не считалась статьей, а «*Reply to*» или «*Comment on*» считались настоящими статьями. В целом, ошибка в каждом подсчете оценивалась почти в 2,5 %.

Главная проблема связана с опечатками в именах соавторов: например, Buldryev и Giovambattista правильно пишутся Buldyrev и Giovambattista, соответственно. Главная проблема также касалась польских, немецких, китайских и корейских имен. Имена и средние имена, последние иногда опускались, проверялись: например, F. W. Star и F. Star являются один и тем же человеком; аналогично A. Petersen и A.M. Petersen. Такая (утомительная) проверка вручную также позволила мне отличить имена-омонимы, такие как Ch. Laurent и Ph. Laurent. HES также упоминает, что одна из его наиболее часто цитируемых статей относится к некому HFS! Все такие «дефекты» и опечатки были апостериори исправлены перед подсчетом авторов вручную. Хотя группирование статей может быть спорным, как и выбор авторов, эти списки могут уверенно рассматриваться как надежные для дальнейшего анализа в рамках настоящего подхода.

Статистические характеристики (средина, медиана, стандартное отклонение, искажение, экспесс, ...) различных распределений  $J(r)$  не представлены, но они доступны по запросу.

Для последующего текста и таблиц используются следующие понятия:

**NJpmfCA:** число совместных публикаций с наиболее частым (mf) соавтором;

**NJPICA:** число соавторов, написавших только 1 публикацию с ВИ (с любым числом соавторов в этой публикации);

**TNCA:** общее число соавторств, любая частота соавторов и число публикаций; т.е. это интеграл гистограммы  $J$  в противоположность  $r$ ;

**NDCA:** число различных соавторов.

\*Все таблицы и рисунки даны в конце статьи.

\*Например, в случае MA совместная публикация, т.е. работа [12], рассматривается как глава книги.

## АНАЛИЗ И ОБСУЖДЕНИЕ МАССИВА ДАННЫХ

Список публикаций HES насчитывает свыше 1100 «работ», разбитых на подгруппы, как показано в табл. 1. Его рабочее резюме и список подобранных публикаций, взятые с сайта [polymer.bu.edu/hes/vitahes-messina.pdf](http://polymer.bu.edu/hes/vitahes-messina.pdf), приводят, помимо прочего, список 15 книжных глав и 5 энциклопедических статей, 619 научных статей за период 1966-1999 гг. плюс свыше 490 статей в журналах за период 2000 - 2012 гг. [Список приведен в порядке ранжирования путем подсчета ссылок.] Отметим, что его резюме (CV) упоминает 136 научных сотрудников и приглашенных ученых на конец 2012 г. DS имеет свыше 600 публикаций, ряд из них на немецком языке, иногда в виде комментариев. Некоторые из его статей появляются в журналах сообществ или в педагогических журналах и редко в рецензируемых журналах. Поступая таким образом, он имеет большее (чем многие другие ученые в этой области) число статей с одним автором (sA), см. табл. 1. Что касается HES и DS, в отдельных случаях вклады одного автора или совместных публикаций возникают из приглашенных лекций, опубликованных по случаю юбилеев. Такие работы подсчитываются в трудах (p). MA имеет около 600 публикаций, смешанные статьи в рецензируемых журналах, трудах (p), вклады в энциклопедии (e) и главы книг (bc). В случаях DS и MA происходит некоторая неясность, так как некоторые работы относятся к статьям, опубликованным в лекциях летних школ. Они были включены в труды (p). Два польских ученых (JMK и AP) имеют похожий карьерный путь, как HES, DS и MA. Один (JMK) частично строит карьеру в США. Четверо других являются более молодыми и, по существу, делают карьеру в Польше. Однако DG начал свою карьеру в физике высоких энергий; он имеет статью с 93 соавторами в трудах. С точки зрения половой принадлежности, все за исключением KSW, являются мужчинами. С таким выбором авторов данные относительно карьеры показывают на различных этапах возрастания размера списка публикаций, см. табл. 1. Такой выбор позволяет охватить широкий диапазон «типов» соавторства.

Для сравнения, характеристики одного автора и совместных публикаций 10 ученых, рассмотренных таким образом, приводятся в табл. 1. Отметим, что в табл. 1 статьи одного автора (sA) отличаются от совместных публикаций (JP) либо в рецензируемых журналах (j), либо в других категориях (p, bc, e). Примечательно, что число совместных публикаций является значительно большим, чем число публикаций одного автора; последнее поднимается только от 5 до 10%. Тот же самый порядок величины, т.е. около 40%, сохраняется в соотношении  $N_{JPp}/N_{JP}$  для HES и MA. Однако это соотношение близко к 15% для DS, AP, KSW и JM. Наоборот, DG, RW и JKM имеют гораздо более высокое соотношение (~0,7), чем все другие. Отметим, хотя это находится вне рамок настоящего обсуждения, что большая часть  $N_{JP}$  или  $N_{sAPp}$  не часто цитируется, как показывается в любом исследовании на основе *h*-индекса! Так как требуется некоторое время, прежде чем публикация будет процитирована в отдельной другой работе, следует учитывать, что некоторая значительная величина ошибки существует в таком измерении ссылок, соответственно в *h*-индексе, - в противоположность настоящему рассмотрению относительно соавторов, которые «располагаются в иерархии» публикации навсегда.

Грубая оценка показывает, что около 20% JP не цитируются. Интересное число порядка возрастания Парето!

Данные, сконцентрированные на соавторах, суммируются в табл. 2-4, группируя статьи в «трудах» (p) и статьи журнального типа (j), а некоторые данные отображены на нескольких рисунках. Чтобы проверить уравнение (1) для каждого соответствия использовался общий диапазон  $r$ , т.е.  $\in [1, r_M]$ . Однако, значение  $R^2$  не всегда было  $\geq 0,9$  (как уже показано на рис. 1). Поэтому соответствие степенному закону, подобное уравнению (1), часто предпринималось в ограниченном диапазоне, около области интереса, т.е. около  $m_A$  такого, что  $r \in [1, r_{cf}]$  и  $J \in [1, r_{cf}]$  с  $r_{cf} \sim 3 m_A$ , см. подрисуночные подписи к каждому рисунку для получения большей информации об отдельном значении  $r_{cf}$  которое используется. Это придает особое значение наиболее частым соавторам, т.е. низким режимам  $r$ , но накладывает некоторое ограничение на  $N_{JPmfCA}$ , т.е. самые высокие  $J$  случаи. Используются два типа размера осей, т.е.  $\ln\text{-}\ln$  или  $\log\text{-}\log$ , но соответствия не всегда показаны дважды. Диапазон соответствия также приводится в каждой подрисуночной подписи. Для лучшего представления показаны только нижний диапазон  $r (< 3m_A)$  и симметрично диапазон данных  $J \leq 3m_A$ . Исследуемые авторы были сгруппированы на иллюстрационных графиках так, чтобы данные были настолько легко читаемыми, насколько это возможно, даже если было неизбежно, что они часто перекрывались.

### Численный анализ $N_{JP}$

Не все данные и соответствия могут быть показаны, но релевантные результаты суммируются в табл. 2-4. Несколько (3) случаев данных общего числа совместных публикаций ( $TN_{JP}$ ), ранжируемых в соответствии с важностью соавторов, представлены и численно соответствуют рис. 4-5. Наблюдаем, что  $a$  ( $TN_{JP}$ ) приближается к  $+1$  для HES и MA (рис.4). Значения  $R^2 \sim 0,9$  и  $a$ , таким образом, являются надежными независимо от диапазона соответствия, т.е. на целый диапазон или близкое  $m_A$ , безусловно для MA, но отметим, что  $a$  для HES опускается от 1,14 до 0,74, если этот диапазон сокращается. Наблюдаем также, что общее  $a$  ( $TN_{JP}$ ) опускается ниже 0,8 для DS и JMK и около 0,7 для KSW и RW.

Находится значение ядра  $m_A$ , например, как указано для HES и MA стрелкой на рис. 5 и для DS, JMK и AP – на рис. 7,8,9. Значения  $TN_{JP} m_A$  для всех авторов также даны в табл. 2-4. Они колеблются от 2 до 26.

Наблюдаем одну крайность, т.е. JM, который имеет небольшое число частых соавторов (CA) вместе с малым числом точек на графике, CA с  $r \geq m_A$ , все типа 1CA, очень противоречат «соответствию», фактически снижая  $a$  до низкого значения, т.е. 0,63, по сравнению с  $a$   $N_{JP}$ . Это в некотором роде похоже на случай DS, для которого значение  $m_A$  НЕ растет от своего значения  $N_{JP}$  к  $TN_{JP}$ . Наконец случай DG представляет также интерес с точки зрения  $TN_{JP}$ , так как он имеет самое низкое  $a$ , ослабленное поведением  $N_{JPp}$ , как подразумевалось здесь выше.

Такое наблюдение в дальнейшем показывает интерес к исследованию эффекта числа соавторов на отдельные типы публикаций, такие как «труды», что и делается в следующем разделе.

### Численный анализ $N_{JPp}$

Отдельные графики поведения  $N_{JPp}$  уже были даны на рис. 1-3. Другие представлены на рис. 6-9. В табл. 2-4 можно найти множество релевантных численных результатов. Интересно, что соответствия  $N_{JPp}$  степенно-

му закону являются лучшими ( $R^2 \sim 0,92$ ) почти для всех авторов (HES и DG!), чем для TNJP (см. рис.6-9).

Давайте понаблюдаем, удовлетворяет условиям уравнения или нет гиперболический закон, уравнение (1), во-первых, в случае HES, рис. 1, для NJPr. Соответствия для NJPr являются вполне хорошими ( $R^2 \geq 0,87$ ) с  $a \approx 1,05$  вне зависимости от рассматриваемого диапазона. Более того, значение  $a$  может быть больше, как для KSW, т.е.  $\approx 1,255$  и ниже, как для DG, т.е.  $\approx 0,239$ .

Проверка была проведена на нескольких случаях, касающихся диапазона соответствия. Пример центрального диапазона соответствия показан на рис. 3 для МА, рис. 5 – для HES и МА, и рис. 9 – для DS, JKM и AP. Численное различие  $a$  между соответствиями с различными диапазонами не является большим. Значение  $m_A$  для NJPr МА считается равным 10.

Тем не менее, некоторое отклонение от точного гиперболического соответствия просматривается около  $m_A$ , указанного стрелкой. Однако, значения  $m_A$  всегда хорошо определяются.

Во всех случаях, можно наблюдать, что значения  $m_A$  являются довольно небольшими:  $m_A = 15$  и  $10$  для HES и МА, соответственно, но опускаются ниже 4 для всех других. Наоборот DS имеет очень низкое значение NJPr  $m_A (= 3)$  и заметно низкое значение  $a (\sim 0,5)$ . Фактически, похожие низкие значения  $m_A$  и  $a$  также найдены для польских ученых. Все значения NJPr  $m_A$  опускаются гораздо ниже значений ядра соавторов NJP  $m_A$ , т.е.  $m_A = 25$  и  $20$  для HES и МА, соответственно, как упоминается в табл. 2.

#### Обсуждение

Аномальное поведение, т.е. большие отклонения от ожидаемого закона, уравнение (1), может, с одной стороны, относиться к самим данным и приближению закона простого шкалирования. С другой стороны, оно может отражать некоторую более глубокую информацию о роли ВИ, лидерстве, управлении, финансировании, научных темах, ... и организации команды, истории, противоречиях [2,3,4].

Во-первых, в самом деле, можно ожидать наблюдения, как во многих графиках Ципфа, некоторой неоднородности при низком  $r$  и большом  $r$ . Так называемый «эффект королевы» [1], показанный своего рода горизонтальным окаймлением в данных, см. рис. 4 или рис. 7, может быть представлен гиперболическим законом, подобным закону Брэдфорда-Ципфа-Манделброта,

$$J = \frac{J^*}{(v+r)^\zeta}, \quad (3)$$

предположительно с  $\zeta \approx +1$  [13]. В этом диапазоне низкого  $r$  известно, что так называемый «эффект короля» [14], т.е. устойчивый рост в значениях низкого  $r$  существует, если такой низко ранжируемый элемент имеет больше преимущества. В примерах с HES также встречается эффект короля и королевы в нескольких случаях. Эффект короля вызван соавторами Хавлином, Буддыревым, Амаралом, Гопикришнаном и Плери, для  $r =$  от 1 до 5, эти соавторы с NJPr  $\geq 40$ . Порядок соавторов отличается от порядка в NJPr и TNJP (не показано). Подобным образом для МА эффектом короля является Клутс, который явно имеет NJPr  $\gg 40$  (см. табл. 2). Для МА порядок ранжирования соавторов также отличается от порядка в NJPr и TNJP (не показано).

Некоторый рост, наблюдаемый в случае HES, около  $r > 10$  может относиться к тому факту, что HES был очень активен во многих различных областях, для кото-

рых соавторы являются похожими, но с различным акцентом в их деятельности, больше благодаря распределению исследования во времени. Эти случаи чаще связаны с медициной. Понятно, что эффект команды, с королями и королевами, должен ожидать в таких областях. Отмеченные эффекты области обязательно замаскированы в общем графике, но допускают некоторое отличие между статьями в «трудах» и рецензируемых журналах.

Наконец, тот факт, что в «трудах» ( $a$  также в журнале) значения  $m_A$  опускаются гораздо ниже общего значения ядра соавторов  $m_A$ , может быть объяснен только размером числа соавторов. Тем не менее, очень похожие короли и королевы встречаются в различных типах публикаций. Указанное выше наблюдение может предполагать, что  $m_A$  зависит от тематики. Нормализация, приводящая к частотному графику, может быть полезной работой, чтобы улучшить  $m_A$  как стандартную меру. Однако другое объяснение может также относиться к типу самой публикации. В самом деле, можно наблюдать, что приблизительно одно и то же значение обнаруживается в соотношении NJPr/NJP для TNCA как для HES, так и для МА, т.е.  $\sim 0,47$ . Тем не менее, это соотношение и для NJPrCA и для NDCA является весьма различным для HES и МА. Значения для случая HES составляют почти половину значений для случая МА. Так как NJPr является (весьма) грубо одинаковым для HES и МА (см. табл. 2), это обязательно показывает вполне различный акцент и успех в публикации научных статей HES и МА с соавторами в рецензируемых журналах.

С точки зрения значения  $m_A$  для NJPr, а также его низкого значения  $a$ , можно прийти к выводу, что DS является особым (основным) ученым в этом отношении.

#### СИСТЕМА ВРЕМЕНИ

Так как HES разбивает свой список публикаций в рабочем резюме на два периода [1966-1999] и [2000-...], то интересно обнаружить, имеется ли некоторого рода временная зависимость выше указанных данных. С одной стороны, научные темы изменяются; с другой стороны, соавторы также могут меняться. Более того, научный результат исследовательских команд может меняться благодаря использованию многих электронных средств и из-за финансовых ограничений. Таким образом, в табл. 3 характеристики JP HES и DS сообщаются для обоих временных режимов, хотя рис. 10 и 11 показывают данные в случае HES, а рис. 12 – в случае DS. Интересно, что значения NJP, NPmfCA, NDCA, NCA,  $a$ ,  $R^2$  и  $m_A$  не кажутся зависимыми от временного режима, таким образом, от акцента на одну или другую тему или на соавторов. Это очень удивительно.

Просматривается, что NsA отличаются в различных режимах: это может быть понято как следующее: в начале карьеры наиболее вероятно, что либо (молодой) ученый принадлежит к большой группе и, таким образом, однажды получает много соавторов, либо он ищет признания, соответственно публикуясь в статьях одного автора, - когда это делать разрешается.

#### ЗАКЛЮЧЕНИЕ

Научная продукция и другие факты, такие как ссылки и (совместные) публикации, используются сегодня для количественного определения научных достижений [2 - 6,15] ученого или целой команды. Однако, каждый признает, что статистические методы, основанные на простых арифметических подсчетах, только отчасти являются адекватными, поскольку любые количественные смещения пропускают релевантные качественные харак-

теристики [17], а подсчет чувствителен к объему данных, хотя взаимодействия и контекстуальные отклонения некоторым образом спрятаны\*. Обрушивается шквал статей, чтобы «улучшить» эти «измерения», что походит на неистощающийся список! [19 – 23].

В некотором роде это публичное знание, что статьи в трудах содержат больше соавторов, чем статьи в рецензируемых журналах. Это видно здесь в соотношении NDCA/TNCA либо для NJPj, либо для NJPp (см. табл. 2 - 4). Официально соавторство должно подразумевать ответственность за содержание статьи. Кроме того, незаслуженное авторство, авторство, приписываемое из-за авторитета или престижа, в качестве благодарности, кажется более позволительным в статьях трудов, являющихся результатом презентаций на конференции. Вместе с увеличивающимся числом соавторов растет доля незаслуженного (ложного) авторства [25,26]. Наоборот, «обязательство» в некотором смысле, страх обидеть кого-либо, потребность в подтверждении некоторого сотрудничества, давление со стороны другого соавтора или определенное требование, обеспечивающее взаимную выгоду, [24] являются признанными фактами.

Определение, данное Ауслусом относительно ядра соавторов, и последующее измерение [1] дают возможность рассматривать такие вопросы конструктивным способом, через взаимосвязь между числом ( $J$ ) (совместных) публикаций с соавторами, ранжируемыми в соответствии с рангом ( $r$ ) важности. Этот подход очень отличается от индекса Хирша. Первый подчеркивает роль человека. Второй контролирует популярность статьи.

Проверка данных в [1], т.е.  $J \propto 1/r^a$  с  $a \approx 1$ , была проведена и обсуждалась в данной статье выше; рассматривались два продуктивных высокоуважаемых автора, часто цитируемые, и некоторые другие авторы, работающие в области статистической физики. Исследуется роль коллег, фактически соавторов, с учетом предположения, что публикации в рецензируемых журналах и в так называемых трудах могут иметь в какой-то степени различное влияние на ядро соавторов ВИ. Это доказано здесь для исследуемых выше случаев. Наблюдается, что значения соответствия  $R^2$  обычно весьма хорошие, особенно учитывая, что простой степенной закон соответствия является пробным. Значения  $m_A$  так называемых подмножеств соавторов опускаются ниже значения целого ядра  $m_A$  соавторов и отличаются, находятся ли публикации в рецензируемых журналах или в так называемых трудах. Практически, значения  $a$  и  $m_A$  являются зависимыми как от размера, так и от типа публикации. Таким образом, практические рассмотрения поведения ВИ и научной команды можно наблюдать через показатель  $a$  TNJP и значение ядра соавторов  $m_A$ .

Отсюда следует, что еще многое может быть сделано после указанного выше, в частности, проверка: отклонений от регулярного гиперболического закона; эффектов короля и королевы; эффекта длинного хвоста NDCA; влияния NJP1CA; эффекта числа соавторов на статью [27,28]; глобализации измерений при рассмотрении роли основного автора и ранжировании его коллег по команде. Проблема некоторого моделирования остается для дальнейшей работы. Можно представить расширения модели Эгте «успех порождает успех» [29], приводящей к необязательному снижению распределений.

Технически можно таким образом измерить релевантную силу научной группы, сконцентрированной вокруг некоторого лидера, например, через различные со-

\*Работа [18] по критической массе и зависимости качества исследования от размера группы.

отношения, которые обсуждались в основном тексте статьи. Тем не менее, невидимый колледж [30,31] становится видимым и легко определяемым, включая центры таких процессов. Можно предположить, что есть возможность ввести политику отбора и вознаграждения в финансирование команды (или ВИ) с помощью измерения ядра соавторов  $m_A$ , предложенного Ауслусом [1]. Демонстрация дальнейшего практического применения, например, для финансирования команды (или ВИ) или для критериев карьерного роста, находится за рамками этой статьи.

*Благодарность.* Автор выражает благодарность М. Ауслусу за обсуждение по теме [1], первоначальные комментарии, сделанные до представления рукописи и доступ к его списку публикаций, преобразованному в соответствии с настоящим исследованием.

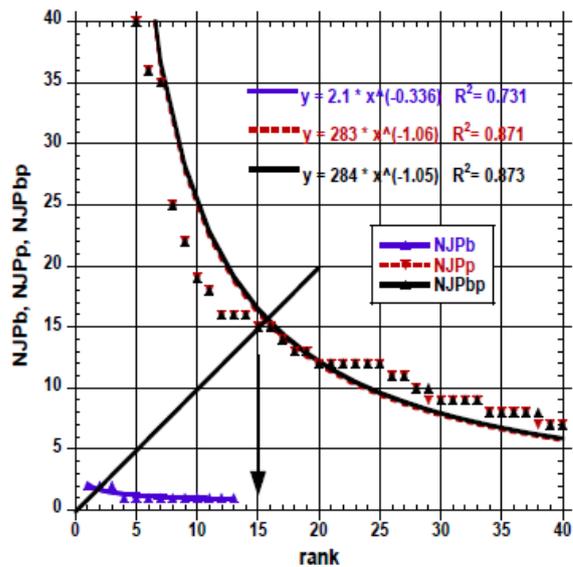


Рис. 1. Число совместных публикаций (JP) HES с соавторами, ранжируемых по снижению важности в книгах и энциклопедиях (NJPbe), в трудах (NJPp), а также «окончательная сумма» сочетания (NJPbp) в близости измерения ядра соавторов [1], показанная по диагонали; соответствующее значение  $m_A (=15)$  указано стрелкой; каждое наилучшее соответствие дано для диапазона *целого r*

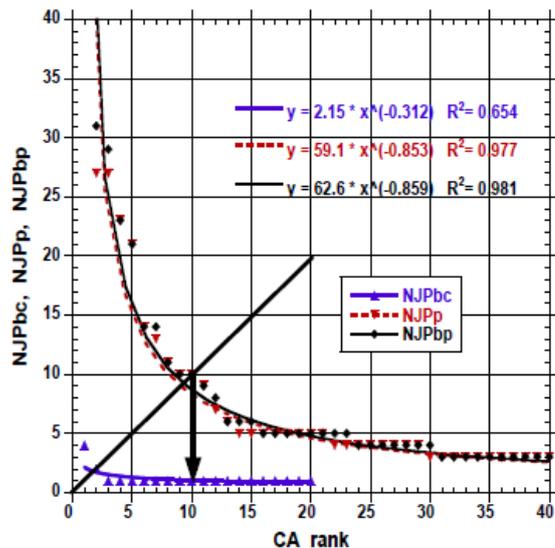


Рис. 2. Число совместных публикаций (JP) MA с соавторами, ранжируемых по снижению важности в книгах и энциклопедиях (NJPbe), в трудах (NJPp), а также «окончательная сумма» сочетания (NJPbp) в близости измерения ядра соавторов [1], показанная по диагонали; соответствующее значение  $m_A (=10)$  указано стрелкой; каждое наилучшее соответствие дано для диапазона *целого r* ( $r_M$ , NJPmfCA)

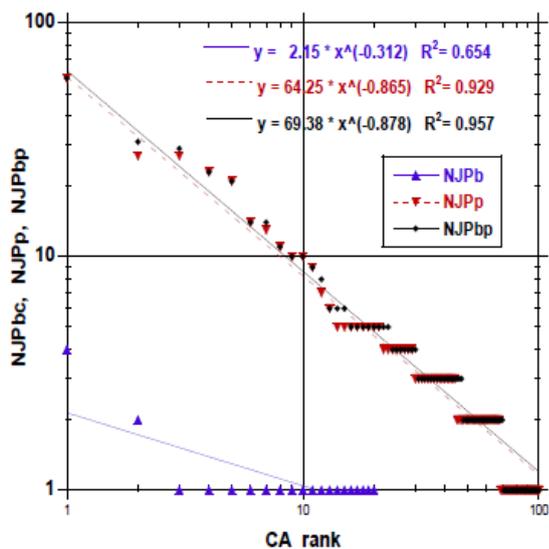


Рис. 3. Показ log-log числа совместных публикаций (NJP) МА с соавторами, ранжируемых по уменьшению важности в книгах и энциклопедиях (NJPbe), в трудах (NJPp), а также их «окончательная сумма» сочетания (NJPbp); каждое наилучшее соответствие сделано на данных диапазона центрального  $r$  ( $r \leq 40, J \leq 40$ )

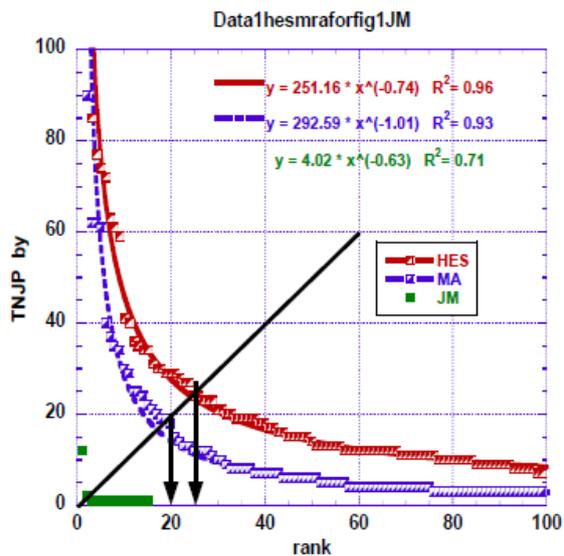


Рис. 5. Общее число совместных публикаций (TNJP) HES, МА и JM с соавторами (СА), ранжируемых по уменьшению важности вблизи так называемого ядра соавторов, измеренного Луслуом [1], показанного по диагонали; значения TNJP  $m_A$  для HES и МА показаны стрелками; наилучшие соответствия даны только для *центральной* области, т.е.  $r$  и  $J \leq 100$

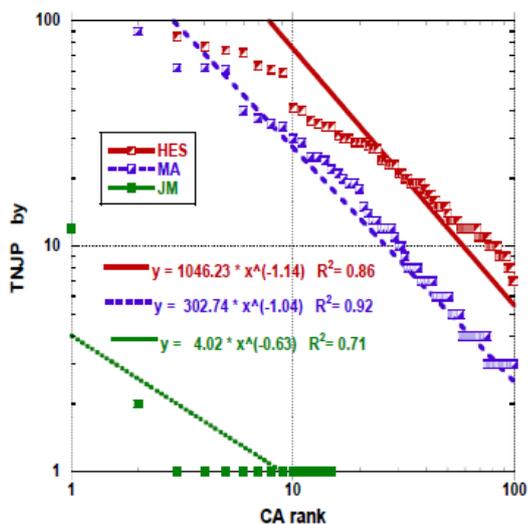


Рис. 4. Показ шкалы log-log общего числа совместных публикаций (TNJP) HES, МА и JM с соавторами (СА), ранжируемых по уменьшению важности вблизи так называемого ядра соавторов, измеренного Луслуом [1]; наилучшие соответствия даны для *целого* диапазона ( $r, J$ )

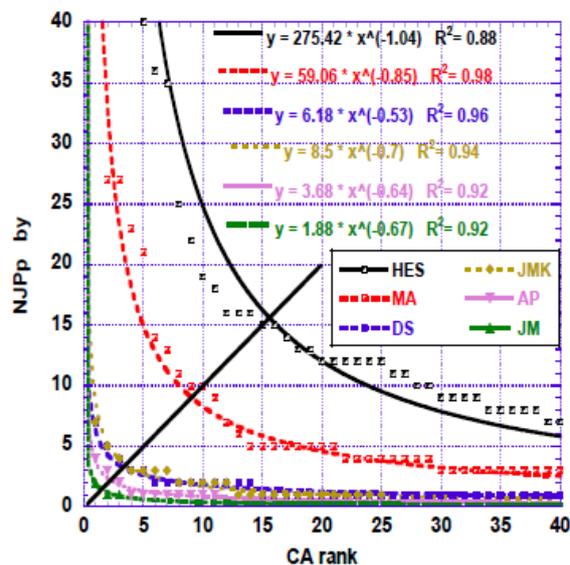


Рис. 6. Показ числа совместных публикаций в трудах (NJPp), сделанных шестью ВИ, т.е. HES, DS, МА, JMK, AP и JM с соавторами, ранжируемых по уменьшению важности, близко к  $m_A$ ; его значение, предполагаемое по диагонали, приводится в табл. 2-4; каждое наилучшее соответствие дано для *целого*  $r$  и  $J$  диапазона ( $r_M, NJPpmfCA$ )

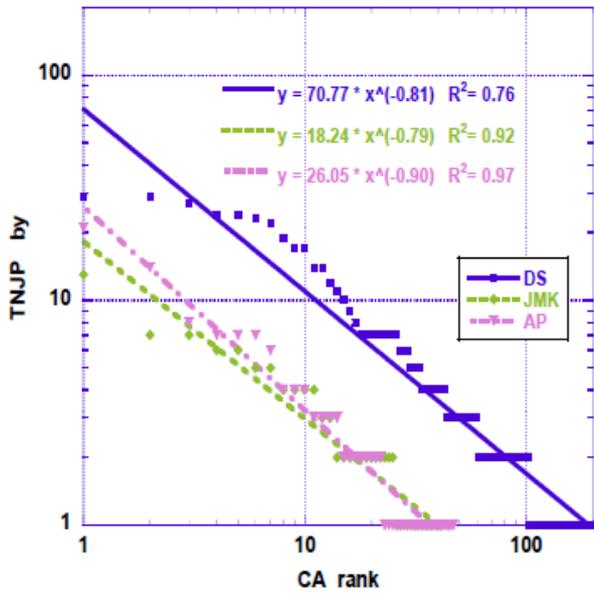


Рис. 7. Показ шкалы log-log общего числа совместных публикаций (TNJP) DS, JMK и AP с соавторами (CA), ранжируемых по уменьшению важности, наилучшие соответствия степенному закону даны для диапазона *целого* ( $r, J$ )

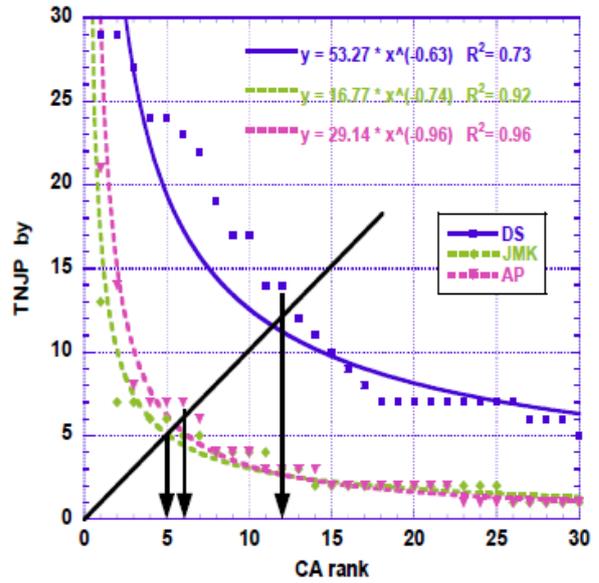


Рис. 9. Общее число совместных публикаций (TNJP) DS, JMK и AP с соавторами (CA), ранжируемых по уменьшению важности; каждое наилучшее соответствие степенному закону дано для диапазона ( $r, J$ ), [30, 30] т.е. близко к индексу  $m_A$ , т.е. так называемое ядро соавторов, измеренное Ауслусом[1], показанное стрелкой для каждого ВИ

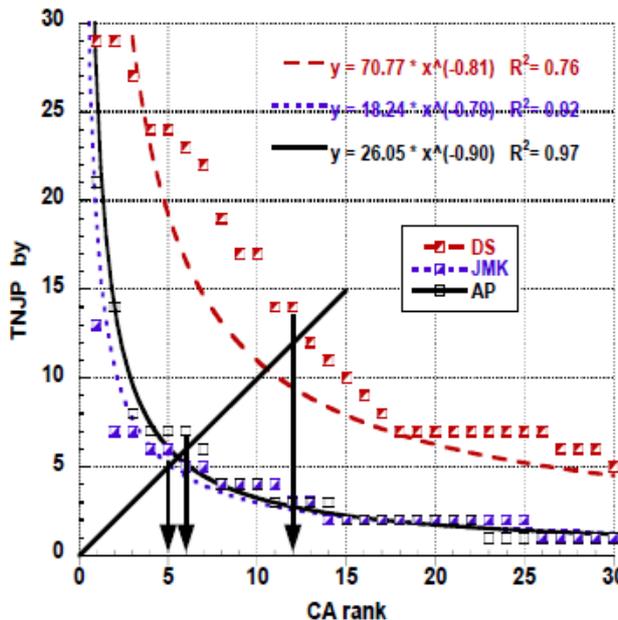


Рис. 8. Общее число совместных публикаций (TNJP) DS, JMK и AP с соавторами (CA), ранжируемых по уменьшению важности, каждое наилучшее соответствие степенному закону дано для диапазона *целого* ( $r, J$ ); так называемое ядро соавторов, измеренное Ауслусом [1], показано стрелкой для каждого ВИ

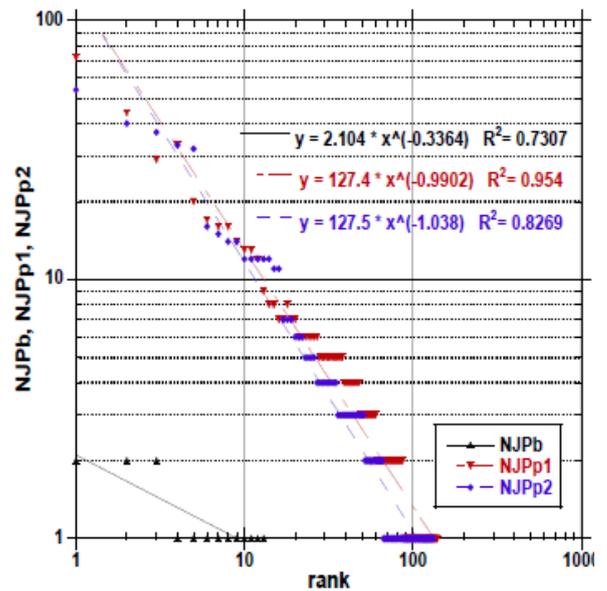


Рис. 10 Показ шкалы log-log числа совместных публикаций (NJP) HES с соавторами (CA), ранжируемых по уменьшению важности, в книгах и энциклопедиях (NJPb) и «трудах» (NJPp) в течение различных временных пределов ( $p1$  и  $p2$ ), см. табл. 5; каждое наилучшее соответствие дано для целых диапазонов

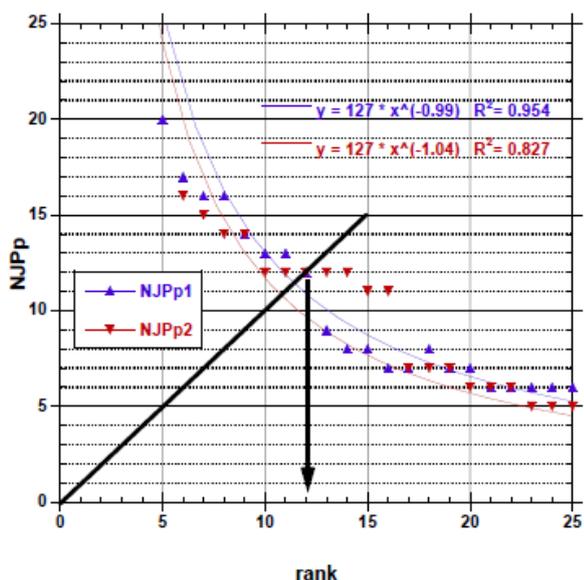


Рис. 11. Общее число совместных публикаций (TNJP) HES с соавторами (CA), ранжируемых по уменьшению важности в «трудах» в течение различных временных пределов (p1 и p2), см. табл. 5, вблизи измерения ядра соавторов [1], показанного по диагонали; значение  $m_A (= 12)$  показано стрелкой; каждое наилучшее соответствие дано для целых диапазонов

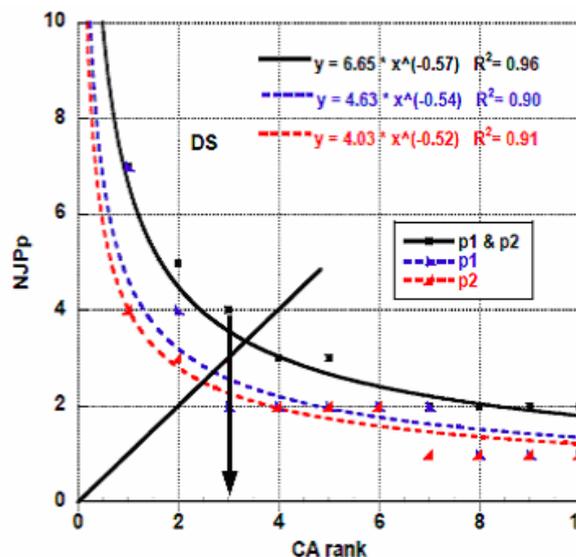


Рис. 12. Общее число совместных публикаций (TNJP) DS с соавторами (CA), ранжируемых по уменьшению важности в «трудах» в течение различных временных пределов (p1 и p2), см. табл. 5, вблизи измерения ядра соавторов [1], показанного по диагонали; значение  $m_A (= 3)$  показано стрелкой; каждое наилучшее соответствие дано для целых диапазонов

Таблица 1

	суммарные данные								
	NP	NJPj	NsAPj	NJPp	NsAPp	NJPbc	NsAPbc	NJPe	NsAPe
HES	1148	791	15	288	34	9	6	2	3
DS	612	374	115	46	77	-	13	-	4
MA	599	359	29	148	31	15	5	1	1
RW	98	32	7	21	5	11	1	0	0
JMK	60	28	8	19	5	-	-	-	-
AP	111	79	22	9	1	-	-	-	-
DG	45	12	14	9	10	0	0	0	0
KSW	43	27	7	4	5	0	0	0	0
JM	27	15	8	2	1	1	1	0	0

Примечание: NP – общее число (N) публикаций (P); NJPj – число совместных публикаций (JP) в журналах (j); NJPp – число совместных публикаций в трудах (p); NJPbc – число совместных публикаций в главах книг (bc); NJPe – число совместных публикаций в энциклопедиях (e); sA – один автор. Данные приводятся на дек. 2012 г.

Таблица 2

	HES			DS			MA		
	NJPj	NJPp	TNJP	NJPj	NJPp	TNJP	NJPj	NJPp	TNJP
Самая старая публикация	1965	1969	1965	1968	1973	1968	1971	1971	1971
Самая новая публикация	2012	2012	2012	2012	2011	2012	2012	2011	2012
NJPMfCA	195	104	299	27	7	29	97	58	155
NJP1CA	275	102	269	172	32	178	152	103	172
TNCA	2639	1250	3889	691	72	763	1055	502	1557
NDCA	568	242	592	268	46	280	273	168	319
$\alpha$	0,999	1,045	1,135	0,688	0,535	0,796	1,03	0,860	1,04
R <sup>2</sup>	0,914	0,873	0,86	0,752	0,965	0,722	0,91	0,981	0,92
$m_A$	20	15	26	12	3	12	15	10	20

Примечание: Суммарные характеристики данных для числа совместных публикаций (NJPP) HES, DS и MA в соответствии с нотациями, данными в тексте статьи, т.е. различная журналы (j), «труды» (p) и общую (T) сумму (TNJP). Ядерные характеристики закона соавторов,  $\alpha$  и  $m_A$  [1] также выведены из соответствий, обсуждаемых в тексте и подрисуночных подписях.

Таблица 3

	RW			JMK			AP		
	NJPj	NJPp	TNJP	NJPj	NJPp	TNJP	NJPj	NJPp	TNJP
Самая старая публикация	1997	1995	1995	1969	1971	1969	1969	1980	1969
Самая новая публикация	2012	2010	2012	1999	1999	1999	2012	2010	2012
NJPmfCA	5	10	15	6	7	13	17	4	21
NJP1CA	13	11	24	24	13	16	25	8	25
TNCA	62	62	124	60	51	111	118	17	135
NDCA	34	23	46	35	25	41	45	11	47
$\alpha$	0,561	0,767	0,743	0,618	0,702	0,787	0,89	0,64	0,94
R <sup>2</sup>	0,867	0,901	0,918	0,887	0,943	0,915	0,945	0,92	0,96
m <sub>A</sub>	4	4	6	4	3	5	5	2	6

Примечание: Суммарные характеристики данных для совместных публикаций RW, JMK и AP в соответствии с нотациями, приведенными в тексте статьи, т.е. различная журналы (j), «грудь» (p) и общую (T) сумму (TNJP). Ядерные характеристики закона соавторов,  $\alpha$  и  $m_A$  [1] также выведены из соответствий, обсуждаемых в тексте и подрисуночных подписях.

Таблица 4

	DG			KSW			JM		
	NJPj	NJPp	TNJP	NJPj	NJPp	TNJP	NJPj	NJPp	TNJP
Самая старая публикация	2000	2000	2000	1996	1999	1999	2001	2006	2001
Самая новая публикация	2012	2000	2012	2012	2006	2012	2012	2006	2012
NJPmfCA	5	5	10	6	3	9	10	2	12
NJP1CA	4	3	7	10	0	10	10	0	10
TNCA	14	104	118	44	5	49	23	2	27
NDCA	7	99	104	21	3	21	12	1	14
$\alpha$	0,755	0,239	0,547	0,594	1,255	0,715	0,67	0,67	0,63
R <sup>2</sup>	0,933	0,321	0,615	0,865	0,954	0,961	0,76	0,92	0,71
m <sub>A</sub>	2	2	2	3	1	3	2	1	2

Примечание: Суммарные характеристики данных для совместных публикаций DG, KSW и JM в соответствии с нотациями, приведенными в тексте статьи, т.е. различная журналы (j), «грудь» (p) и общую (T) сумму (TNJP). Ядерные характеристики закона соавторов,  $\alpha$  и  $m_A$  [1] также выведены из соответствий, обсуждаемых в тексте и подрисуночных подписях.

Таблица 5

	HES			DS		
	временной интервал			временной интервал		
	[1966-1999]	[2000-2012]	[1966-2012]	[1966-1999]	[2000-2012]	[1966-2012]
	NJPp1	NJPp2	NJPp	NJPp1	NJPp2	NJPp
NJP	150	138	288	41	31	72
NJPmfCA	72	54	104	7	4	7
NJP1CA	53	66	102	20	15	32
NDCA	140	132	242	27	21	46
NsA	28	8	36	46	32	78
$\alpha$	0,99	1,04	1,045	0,54	0,52	0,565
R <sup>2</sup>	0,954	0,827	0,875	0,896	0,912	0,965
m <sub>A</sub>	12	12	15	2	2	3

Примечание: Система времени. Суммарные характеристики данных для публикаций HES и DS на дек. 2012 г. – в «трудах»: число совместных публикаций (NJP) с одним соавтором (CA); NPmfCA – число публикаций с наиболее частыми (mf) соавторами; NP1CA – число публикаций только с одним соавтором (CA); NDCA – число различных соавторов; показатель степенного закона  $\alpha$  в уравнении (1) и соответствие R<sup>2</sup>; мера m<sub>A</sub> ядра CA [1] также дана.

## Литература

1. *Ausloos M.* A scientometrics law about co-authors and their ranking. The co-author core// *Scientometrics*. — 2012 (в печати).—arXiv : 1207.1614.
2. *Petersen A.M., Wang F., Stanley H.E.* Methods for measuring the citations and productivity of scientists across time and discipline// *Phys. Rev. E*. — 2010. — Vol.81, 036114.
3. *Petersen A.M., Riccaboni M., Stanley H.E., Pammolli F.* Persistence and uncertainty in the academic career// *Proc Natl Acad Sci USA*. — 2012. — Vol. 109. — P. 5213-5218.
4. *Petersen A.M., Jung W.-S., Yang J.-S., Stanley H.E.* Quantitative and empirical demonstration of the Matthew effect in a study of career longevity// *Proc Natl Acad Sci USA*.— 2011.— Vol. 108. — P.18-23.
5. *Hirsch J.E.* An index to quantify an individual's scientific research output// *Proc Natl Acad Sci USA*. — 2005. — Vol.102. — P.16569-16572.
6. *Hirsch J.E.* An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship// *Scientometrics*. — 2010. — Vol. 85. — P. 741-754.
7. *Kvok L.S.* The White Bull effect: Abusive coauthorship and publication parasitism// *Journal of Medical Ethics*. — 2005. — Vol.31. — P. 554-556.
8. *Clauzet A., Shalizi A.C.C., Newman M.E.J.* Power-law distributions in empirical data// *SIAM Rev.* — 2009. — Vol. 51. — P. 661-703.
9. *Kretschmer H., Rousseau R.* Author inflation leads to a breakdown of Lotkas Law// *Journal of the American Society for Information Science and Technology*. — 2001. — Vol. 52. — P. 610-614.
10. *Stanley H.E., Buldyrev S.V., Goldberger A.L., Havlin S., Peng C.-K., Simons M.* Scaling features of noncoding DNA// *Physica A*. — 1999. — Vol. 273. — P.1-18.
11. *Urbanc B., Borreguero J.L., Cruz L., Stanley H.E.* Amyloid  $\beta$ -Protein Aggregation: Ab initio discrete molecular dynamics approach to protein folding and aggregation, invited review [Amyloid, Prions, and other Protein Aggregates, Pt. B]// *Methods in Enzymology*. — 2006. — 412, Chapter 19.—P. 314-338.
12. *Ausloos M., Lecomte G.V., Harrison J., Krusius M.* Low temperature physics at the millenium, in physics 2000 as it enters a millenium/ P. Black, G. Drake, L. Jossem (Eds.)// (IUPAP-36. — 1999). — P. 27-33.
13. *Fairthorne R. A.* Empirical hyperbolic distributions (Bradford-Zipf-Mandelbrot) for bibliometric description and prediction// *Journal of Documentation*. — 1969. — Vol. 25. — P. 319-343.
14. *Laberrère J., Sornette D.* Stretched exponential distributions in nature and economy: Fat tails with characteristic scales// *Eur. Phys. J. B 2*. — 1998. — P. 525-539.
15. *Fernández-Cano A., Torralbo M., Vallejo M.* Reconsidering Price's model of scientific growth: An overview// *Scientometrics*. — 2004. — Vol. 61. — P. 301-321.
16. *Kelly C.D., Jennions M.D.* The h index and career assessment by numbers// *Trends in Ecology and Evolution*. — 2006. — Vol. 21. — P. 167-170.
17. *Ormerod J.R.* Is content analysis either practical or desirable for research evaluation?// *Omega*. — 2000. — Vol. 28. — P. 241-245.
18. *Kenna R., Berche B.* Critical mass and the dependency of research quality on group size// *Scientometrics*. — 2011. — Vol.86. — P. 527-540.
19. *Durieux V., Gevenois P.A.* Bibliometric indicators: Quality measurements of scientific publication// *Radiology*. —2010. — Vol. 255. — P. 342-351
20. *Schreiber M.* Twenty Hirsch index variants and other indicators giving more or less preference to highly cited papers// *Annalen der Physik (Berlin)*. — 2010. — Vol. 522.— P. 536-554.
21. *Schreiber M., Malesios C.C., Psarakis S.* Exploratory factor analysis for the Hirsch index, 17 h-type variants, and some traditional bibliometric indicators// *Journal of Informetrics*. — 2012. — Vol. 6. — P. 347-358.
22. *Zhang R.* An index to link scientific productivity with visibility. — 2009. —arxiv.org/pdf/0912.3573.
23. *Bornmann L., Mutz R., Daniel H.* Are there better indices for evaluation purposes than the h-index? A comparison of nine different variants of the h-index using data from biomedicine// *Journal of the American Society for Information Science and Technology*. — 2008. —P. 830- 837.
24. *Martin T., Ball B., Karrer B., Newman M.E.J.* Coauthorship and citation in scientific publishing.
25. *Slone R. M.* Coauthors contributions to major papers published in the AJR: Frequency of undeserved coauthorship// *American Journal of Radiology*. — 1996. — Vol. 167. — P. 571-579.
26. *Vuckovic-Dekic Lj.* Multiauthorship - what is it, and does it matter?// *Archive of Oncology*. —2000. — Vol. 8. — P. 139.
27. *Newman M. E. J.* Scientific collaboration networks: I. Network construction and fundamental results// *Physical Review E*. — 2001. — Vol. 64, 016131.
28. *Hsu J. W., Huang D.W.* Distribution for the number of coauthors// *Physical Review E*. — 2009. — Vol. 80, 057101.
29. *Egge L., Rousseau R.* Introduction to Informetrics. Quantitative methods in library, documentation and information science. — Elsevier, Amsterdam, 1990.
30. *Kretschmer H.* Coauthorship networks of invisible colleges and institutional communities// *Scientometrics*. — 1994. — Vol. 30. — P. 363-369.
31. *Zuccala A.* Modeling the invisible college// *Journal of the American Society for Information Science and Technology*. — 2006. — Vol. 57. — P.152-168.

## Аркадий Иванович Черный (1929 –2013)

27 августа 2013 года на 85-м году жизни после тяжелой продолжительной болезни скончался один из основоположников отечественной информатики, старейший сотрудник ВИНТИ РАН, доктор технических наук, профессор, Заслуженный деятель науки Российской Федерации, Аркадий Иванович Черный.

Человек редкого постоянства и целеустремленности, он никогда не менял места работы, направления научных интересов, жизненных убеждений. В 1952 г. он с отличием окончил Московский институт востоковедения и с 14 октября был принят младшим научным сотрудником в только что созданный Институт научной информации Академии наук СССР. За более чем полувековую деятельность в ВИНТИ Аркадий Иванович последовательно занимал должности заведующего научно-методическим отделом, заместителя директора по научной работе (1968—1978), заведующего сектором общесистемных исследований и разработки автоматизированной справочно-информационной системы (1983—1986), заведующего отделом научной информации по информатике, был главным редактором РЖ «Информатика», заместителем главного редактора журналов «Научно-техническая информация» и «Международный форум по информации».

В 1958 г. А.И. Черный участвовал в Международной конференции по информационной науке в Вашингтоне, в 1961 г. как стипендиат ЮНЕСКО изучал опыт научно-информационной работы в Великобритании, Нидерландах и Франции. Он обладал обширными знаниями об организации мировой информационной деятельности и щедро делился ими с коллегами по профессии. В 1999 г. защитил докторскую диссертацию на тему: «Автоматизированная система подготовки баз данных и информационных изданий по естественным и техническим наукам: принципы построения, технология, перспективы».

А.И. Черный является автором (наряду с А.И. Михайловым и Р.С. Гиляревским) известных в стране и в мире монографий «Основы научной информации» (1965), «Основы информатики» (1968), «Введение в теорию информационного поиска» (1974), «Научные коммуникации и информатика» (1976) и нескольких сотен других научных публикаций. В течение десятилетий он ежедневно следил за наиболее значимыми работами по информатике в мировом потоке научной литературы, формулируя направления развития этой научной дисциплины. В этом отношении он — уникальный ученый. Это отразилось в написанной им совместно Ю.М. Арским и Р.С. Гиляревским монографии «Инфосфера» (1996), опубликованной небольшим тиражом и малоизвестной научному сообществу.

А. И. Черный — создатель Автоматизированной справочно-информационной системы по науке и технике (АССИСТЕНТ), которая стала основой всех достижений ВИНТИ в области автоматизации и одной из самых прогрессивных систем в стране. Он создал и неизменно поддерживал на высоком научном уровне реферативный журнал «Информатика», который является солидным информационным обеспечением научных исследований по информатике и библиотековедению – в стране и на постсоветском пространстве. Большая научная заслуга Аркадия Ивановича состоит и в том, что информатика как научная дисциплина не растворяется в модных направлениях вычислительной техники, программирования, искусственного интеллекта, науковедения, семиотики, информационного менеджмента. Он был ревностным патриотом своего Института и написал книгу по истории ВИНТИ: «Всероссийский институт научной и технической информации: 50 лет служения науке» (2005).

Мы, его коллеги, горды и счастливы тем, что каждый день работали вместе с человеком редкой эрудиции, доброжелательным, отзывчивым и внимательным товарищем, и тем неформально высоким положением, которое по праву занимал Аркадий Иванович в нашей профессии и в науке. Светлая память о Аркадии Ивановиче навсегда сохранится в наших сердцах.

Приглашаем российских и зарубежных авторов к сотрудничеству  
в журнале «Международный форум по информации».  
Оригинальные статьи и другие материалы (рецензии, письма)  
можно присылать на русском или английском языке  
по почтовому адресу, указанному в «Памятке для авторов»  
или по электронной почте: [mfi@viniti.ru](mailto:mfi@viniti.ru).

Ответственный за выпуск *Л. В. Кобзева*

Компьютерная верстка *М. А. Филимонова*

ИД № 04689 от 28.04.2001 г.

Подписано в печать 13.09.2013 г.

Бумага “Хегох”. Формат 60x841/8. Гарн. литер. Печать цифровая

Усл. печ. л 5,50 Уч.-изд. л. 5,02 Тираж 55 экз.

Адрес редакции: 125190, Россия, г. Москва, ул. Усиевича, д. 20

Тел. (499) 155-44-95

# РОССИЙСКАЯ АКАДЕМИЯ НАУК

## Федеральное государственное бюджетное учреждение науки ВСЕРОССИЙСКИЙ ИНСТИТУТ НАУЧНОЙ И ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ РОССИЙСКОЙ АКАДЕМИИ НАУК

**предлагает научным работникам, аспирантам и другим специалистам в области естественных, точных и технических наук, желающим быстро и эффективно опубликовать результаты своей научной и научно-производственной деятельности, использовать способ публикации своих работ через *систему депонирования*.**

**«Депонирование (передача на хранение) – особый метод публикации научных работ (отдельных статей, обзоров, монографий, сборников научных трудов, материалов научных мероприятий – конференций, симпозиумов, съездов, семинаров) узкоспециального профиля, разрешенных в установленном порядке к открытому опубликованию, которые нецелесообразно издавать полиграфическим способом печати, а также работ широкого профиля, срочная информация о которых необходима для утверждения их приоритета. Депонирование предусматривает прием, учет, регистрацию, хранение научных работ и обязательное размещение информации о них в специальных информационных изданиях».**

Подготовка и передача на депонирование научных работ происходит в соответствии с «Инструкцией о порядке депонирования научных работ по естественным, техническим, социальным и гуманитарным наукам» (М., 2003).

Результатом депонирования является публикация информации о депонированных научных работах в информационных изданиях ВИНТИ РАН – Реферативном журнале и аннотированном библиографическом указателе «Депонированные научные работы».

В соответствии с “Положением о порядке присуждения ученых степеней”, утвержденным Постановлением Правительства Российской Федерации от 30.01.2002 № 74 (в ред. Постановлений Правительства РФ от 20.04.2006 № 227, от 02.06.2008 № 424, от 20.06.2011 № 475), научные работы, депонированные в организациях государственной системы научно-технической информации, признаны публикациями, учитываемыми при защите кандидатских и докторских диссертаций.

Подать научную работу на депонирование можно обратившись в Отдел депонирования ВИНТИ РАН по адресу:

**125190, Москва, ул. Усиевича, 20.**

**ВИНТИ РАН, Отдел депонирования научных работ.**

**Тел.: 8 (499) 155-43-28, Факс: 8 (499) 943-00-60.**

**e-mail: [dep@viniti.ru](mailto:dep@viniti.ru)**

С инструкцией о порядке депонирования можно ознакомиться на сайте ВИНТИ РАН:  
**<http://www.viniti.ru>**