

М. А. Акоев, В. А. Маркусова, О. В. Москалева, В. В. Писляков

РУКОВОДСТВО ПО НАУКОМЕТРИИ: ИНДИКАТОРЫ РАЗВИТИЯ НАУКИ И ТЕХНОЛОГИИ

Екатеринбург, ИПЦ УрФУ
2014

УДК 001.1
ББК 72
Р851

Авторы
Марк Анатольевич Акоев Валентина
Александровна Маркусова Ольга
Васильевна Москалева Владимир
Владимирович Писляков

Редактор
Марк Анатольевич Акоев

Дизайн обложки
Сергей Георгиевич Слюсарев

Руководство по наукометрии: индикаторы развития науки и технологии : [монография] /
Р 851 М. А. Акоев, В. А. Маркусова, О. В. Москалева, В. В. Писляков ; [под. ред. М. А. Акоева]. – Екате-
ринбург : Изд-во Урал. ун-та, 2014. – 250 с.

ISBN 978-5-7996-1352-5

DOI 10.15826/B978-5-7996-1352-5.0000

Данное руководство систематизирует знания о наукометрии для читателей, поставивших перед собой цель разобраться в предмете. Материал, изложенный в книге, должен послужить основой для практического применения методов наукометрии. Подчеркивается необходимость грамотной и аккуратной трактовки наукометрических индикаторов при принятии административных решений, распределении грантов, осуществлении кадровой политики.

Руководство предназначено для научных работников, руководителей, сотрудников информационных центров и студентов соответствующих специальностей, а также для всех, кто интересуется вопросами измерения и оценки развития науки и технологии.

УДК 001.1
ББК 72

ISBN 978-5-7996-1352-5

© Thomson Reuters, 2014

СОДЕРЖАНИЕ

CONTENTS

Вступительное слово Юджина Гарфилда	8
От авторов	10
Об авторах	12
В. А. Маркусова. Введение. К 50-летию Science Citation Index: История и развитие наукометрии	14
1. М. А. Акоев. Наука, технология и общество	49
1.1. Процесс оценивания	52
1.2. Целеполагание в научном процессе	56
1.3. Выделяемые ресурсы	61
1.4. Особенности анализа количественных данных	66
1.5. Задержки и артефакты научного и технологического процессов	69
2. В. В. Писляков. Библиометрические индикаторы в ресурсах Thomson Reuters	75
2.1. Библиометрические инструменты. База данных Web of Science Core Collection и аналитические надстройки	76
2.2. Индикаторы влиятельности статей (импакт-индикаторы)	82
2.3. Индикаторы относительной влиятельности статей (относительные, нормализованные импакт-индикаторы)	89
2.4. Индикаторы «экстремальной» цитируемости	97
2.5. «Взвешенные» индикаторы. Собственный фактор (Eigenfactor) и индекс влияния статьи (Article Influence)	100
2.6. Индекс Хирша (h-index) и «хирш-подобные» показатели	102
2.7. Показатели хронологического распределения ссылок	106

Preface to the Russian Readers, by Dr. Eugene Garfield	8
Preface	10
About Authors	12
V. Markusova. Introduction. 50 th anniversary of the Science Citation Index: History and Evolution of Scientometrics	14
1. M. Akoev. Science, Technology and Society	49
1.1. Research Evaluation Process	52
1.2. Goal-setting in Research and Development Management	56
1.3. Allocation of Resources	61
1.4. Quantitative Data Analysis	66
1.5. Research and Development Outputs and Data Lags	69
2. V. Pislyakov. Bibliometric Indicators in the Thomson Reuters Information Resources	75
2.1. Bibliometric Instruments. Web of Science Core Collection and Analytical Tools	76
2.2. Article Impact Indicators	82
2.3. Relative Article Impact Indicators	89
2.4. "Extreme" Citation Indicators	97
2.5. "Weighted" Indicators. Eigenfactor and Article Influence Scores	100
2.6. Hirsch index (h-index) and Hirsch-type indicators	102
2.7. Chronological Distribution of References	106

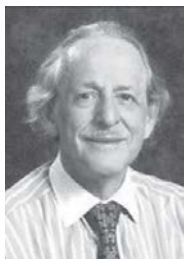
3. О. В. Москалева. Научные публикации как средство коммуникации, анализа и оценки научной деятельности	110	3. O. Moskaleva. Research Publications as a Means of Communication, Analysis and Assessment of Research Activity	110
3.1. История создания и характеристики научных журналов	111	3.1. History and Characteristics of Scholarly Journals	111
3.2. Базы данных публикаций как инструменты поиска и анализа	116	3.2. Publication Databases as Search and Analysis Instruments	116
3.3. Природа цитирования	118	3.3. Nature of Citations	118
3.4. Анализ цитирования	121	3.4. Citation Analysis	121
3.5. Показатели научных журналов	122	3.5. Journal Indicators	122
3.6. Показатели для ученых и организаций	124	3.6. Scientometric Indicators for Researchers and Organizations	124
3.7. Классификаторы, используемые в науке	129	3.7. Scholarly Publications Classification	129
3.8. Важность источника информации о публикациях для оценки научной деятельности	141	3.8. Importance of Publication Data Source for Research Evaluation	141
3.9. Особенности публикаций и цитирования в разных научных областях	144	3.9. Publications and Citations Patterns in Different Scientific Disciplines	144
3.10. Способы оценки публикаций, отличные от анализа цитирования	149	3.10. Publication Activity Evaluation Methods Beyond Citation Analysis	149
3.11. Возможные подходы к сравнительному анализу цитирования публикаций в разных областях знаний	151	3.11. Comparative Citation Analysis in Different Scientific Disciplines	151
3.12. Использование библиометрических данных при построении рейтингов вузов и научных организаций	156	3.12. Use of Bibliometric Data in University and Research Organizations Ratings	156
4. М. А. Акоев. Картирование науки и технологии, прогноз развития	164	4. M. Akoev. Mapping Science and Technology, Forecasting Research and Development	164
4.1. Картирование науки и технологии	164	4.1. Mapping Science and Technology	164
4.2. Прогнозирование и развитие	178	4.2. Forecasting Research and Development	178
Статьи Юджина Гарфилда (перевод)	201	Articles by Eugene Garfield	201
1. Ю. Гарфилд. Их ошибки, наши ошибки и ваши ошибки	201	1. E. Garfield. Errors Theirs, Ours and Yours	201
2. Ю. Гарфилд. Русские идут! Часть 1. Сто самых влиятельных советских ученых 1973–1988. Часть 2. Пятьдесят самых цитируемых советских статей за 1973–1988 гг. в Science Citation Index и взгляд на исследовательские фронты 1988 г.	204	2. E. Garfield. The Russians are Coming. Part 1. The Red-Hot 100 Soviet Scientists, 1973–1988. Part 2. The Top 50 Soviet Papers Most Cited in the 1973–1988 Science Citation Index and a Look at 1988 Research Fronts	204

3. Ю. Гарфилд. Эволюция	
Science Citation Index	227
Список сокращений	239
Предметный указатель	241
Именной указатель	246
Реферат	248

3. E. Garfield. The Evolution	
of the Science Citation Index	227
Abbreviations	239
Index	241
Persons	246
Abstract	248



ВСТУПИТЕЛЬНОЕ СЛОВО ЮДЖИНА ГАРФИЛДА



Ю. Гарфилд

PREFACE TO THE RUSSIAN READERS, BY DR. EUGENE GARFIELD

DOI 10.15826/B978-5-7996-1352-5.0001

This monograph is a comprehensive review of various methods and techniques used now in monitoring and assessing progress in scholarly research and technology. For me, it is fascinating to watch the development of this kind of research, from scientometrics to webometrics and altmetrics. These new metrics were started at the Institute for Scientific information (ISI). ISI was launched in 1954 in Philadelphia. In my many visits to the Soviet Union, and then to Russia, I was pleasantly surprised by the interest shown by Soviet and Russian scientists in ISI products including Current Contents, Science Citation Index, Journal Citation Reports and others. This interest contrasted with the poor availability of these products – except for Current Contents. In those days the Soviet Union had only ten subscriptions to the Science Citation Index. In 1992 ISI was acquired by Thomson Reuters. Now the Thomson Reuters Web of Science – the successor of the Science Citation Index – is widely available and easily accessible on the Web to thousands scientists in Russia and the neighbouring countries. Wider use of modern information

resources and tools can help to unlock further the great potential of the Russian science. Thus, this volume, partly focusing on the possibilities provided by Web of Science, is quite timely and welcome.

Scientometric indexes are widely used in evaluation and promotions of scholars, grant awarding, and research personnel recruiting. However, I would like to stress that these quantitative data should never be used as a single criterion of the quality of institutional research and/or individual researchers. These indexes can help in developing background for wise and expert judgment. Uninformed and uncritical use of the quantitative data can be quite detrimental to the progress of research and individual careers.

I hope that this book will facilitate the difficult task of reasonable and thoughtful evaluation of the impact and productivity of scientists and scholars.

A handwritten signature in dark ink, reading "Eugene Garfield".

Eugene Garfield
Philadelphia, PA. USA
November 13, 2014

Данная монография – исчерпывающий обзор ряда современных методов и техник мониторинга и оценки прогресса научных исследований и технологий. Я увлеченно наблюдаю за процессом развития такого вида исследований, от наукометрии до вебометрии и альтметрии. Многие из этих исследовательских направлений появились на свет и развивались в стенах Института научной информации. Институт научной информации был создан в 1954 г. в Филадельфии. Во время моих многочисленных визитов в Советский Союз, а потом и в Россию я был приятно удивлен интересом, проявленным советскими и российскими учеными к разработкам Института научной информации, например, среди прочих, к Current Contents, Science Citation Index и Journal Citation Reports. Этот значительный интерес шел вразрез с крайне ограниченной доступностью этих информационных ресурсов, за исключением Current Contents. В те времена Советский Союз имел только 10 подписок на Science Citation Index. В 1992 г. Институт научной информации стал составной частью компании Thomson Reuters. Сегодня Thomson Reuters Web of Science, наследница Science Citation Index, доступна тысячам ученым в России и соседних странах. Более широкое использование современных информационных ресурсов и инструментов

поможет раскрыть огромный потенциал российской науки, поэтому данную книгу, отчасти посвященную и возможностям Web of Science, можно считать актуальной и долгожданной.

Наукометрические показатели широко используются при оценке и продвижении научных сотрудников, выделении грантов и приеме на работу научно-исследовательского персонала. Однако мне хотелось бы подчеркнуть, что эти количественные данные не должны использоваться как единственный критерий оценки качества исследований научного учреждения или индивидуальных исследователей. Эти показатели могут способствовать формированию взвешенного экспертного суждения. Использование количественных данных без должного предварительного обучения обращению с ними и критического взгляда может воспрепятствовать прогрессу исследований и научных сотрудников.

Я надеюсь, что эта книга упростит сложную задачу по тщательной и осмысленной оценке влияния и продуктивности ученых и научных коллективов.



Юджин Гарфилд
Филадельфия
13.11.2014 г.

ОТ АВТОРОВ

PREFACE

DOI 10.15826/B978-5-7996-1352-5.0002

Настоящее руководство по наукометрии является ответом на все возрастающий интерес к наукометрическим методам анализа работы научных и образовательных организаций России и стран СНГ.

Многие университеты и исследовательские институты применяют для поиска научной информации базы цитирования. Ряд государств СНГ приобрел национальные лицензии на доступ к этим базам, ставя перед собой задачу стимулирования и управления развитием научных исследований. Базы цитирования, являясь наукометрическими инструментами, позволяют рассчитывать показатели научной результативности, поэтому доступ к базам зачастую создает у администраторов «от науки» ошибочное впечатление того, что они могут с помощью автоматически рассчитываемых показателей легко оценить отдельных ученых и научные коллективы.

Цель книги — дать читателю представление об основных методах измерения науки и технологии. Полученные знания, как мы надеемся, помогут найти ответы на ключевые вопросы наукометрии: что мерить, как мерить и зачем ме-

рить. Авторы не ставят целью дать полное и систематическое видение наукометрии по примеру первых опубликованных на русском языке книг по данному предмету (см., например, классические работы Г. М. Доброва, В. В. Налимова, З. М. Мульченко, Ю. В. Грановского, А. И. Яблонского и др., начиная с 1960-х гг.). Информация в книге должна послужить для читателя основой грамотного применения на практике методов наукометрии. Приведенные в книге примеры иллюстрируют связь практических аспектов использования наукометрии с теоретическими исследованиями в области измерения науки и технологий. Основное внимание здесь уделено оценке результативности и принятию решений о поддержке существующих или о развитии новых научных групп. Во всех главах руководства подчеркивается необходимость грамотной и аккуратной трактовки наукометрических индикаторов при принятии административных решений, распределении грантов, осуществлении кадровой политики и т. д.

Во вводной главе представлена история уникальной идеи Юджина Гарфилда по использованию на-

учных ссылок как средства научного поиска, его борьба за воплощение этой идеи в жизнь в форме универсального инструмента для научного поиска — Science Citation Index.

В первой главе нами рассмотрены возможности и ограничения в применении наукометрии при принятии решений о выделении ресурсов для поддержки научной работы и подчеркнута важность привлечения экспертов в предметной области к проведению оценок.

Во второй главе представлены ключевые библиометрические показатели и примеры расчетов с использованием конкретных инструментов. При выборе показателей мы стремились сформировать у читателя понимание того, как научные процессы отражаются в ключевых наукометрических индикаторах. Расчет выбранных показателей может быть выполнен с использованием любого инструмента, содержащего достаточно полную и достоверную информацию о научных публикациях и их цитировании.

В третьей главе представлена связь наукометрических показателей с природой научной коммуникации. Описаны альтернативные цитированию способы оценки публикаций и рассмотрено использование библиометрических показателей при построении рейтингов университетов и научных организаций.

В четвертой главе дана характеристика методов визуализации наукометрической информации, рассмотрены вопросы ограничений в прогнозировании научно-технического развития и аспекты повышения качества существующих научных направлений и развития новых направлений в организации.

Завершают книгу три статьи Юджина Гарфилда, дополняющие главы книги и демонстрирующие его интерес к науке СССР и России.

Авторы выражают свою благодарность всем сотрудникам компании Thomson Reuters, участвовавшим в реализации проекта по изданию данной книги. Отдельная благодарность Валентину Григорьевичу Богорову за идею создания этой книги, организацию и сопровождение проекта ее издания, Олегу Геннадьевичу Уткину, без участия и поддержки которого эта книга бы не состоялась; а также Ясиру Тухами (Yassir Touhami) и Филиппу Пурнеллу (Philipp Purnell) за активное участие в обсуждении концепции книги и поддержку проекта.

Авторы также признательны организаторам конференции Science Online, благодаря которой они познакомились друг с другом; директору Научной библиотеки им. М. Горького СПбГУ Марине Эдуардовне Карповой и главному библиотекарю Научной библиотеки им. М. Горького СПбГУ Екатерине Михайловне Полниковой, планомерно продвигающим идею наукометрии; профессору Александру Ивановичу Пудовкину за ценные замечания и помощь в переводе; Мехер Мистри (Meher Mistry), жене доктора Ю. Гарфилда, и Ж. А. Налимовой-Дрогалиной за помощь в подборе фотоматериалов для иллюстраций; профессору П. Вoutersу (P. Wouters) за его книгу *The Citation Culture*, которая была неоценимым источником информации при работе над историей отечественного науковедения и наукометрии; Юлии Владимировне Запарий за активное обсуждение текста книги и вопросы; Валентине Викторовне Барминой за помощь в работе над текстом и уточнение цитат и фактов; Е. А. Ляшенко, А. А. Сваловой, Н. А. Мазову и В. Н. Гурееву за замечания и корректуру текстов и сотрудникам издательства УрФУ, обеспечившим оперативное издание книги.

Все тексты данной книги отражают только личное мнение их авторов.

ОБ АВТОРАХ

ABOUT AUTHORS

Марк Анатольевич Акоев

Заместитель директора Центра мониторинга науки и образования Уральского федерального университета имени первого Президента России Б. Н. Ельцина; член ACM, IEEE Computer Society, System Dynamic Society и EuroCRIS.
m.a.akoev@urfu.ru

Mark Akoev

Deputy Director of Monitoring Centre of Science and Education, Ural Federal University.
Member of the ACM, IEEE Computer Society, System Dynamic Society and EuroCRIS.
m.a.akoev@urfu.ru

Валентина Александровна Маркусова

Доктор педагогических наук, заведующая отделением научно-информационного обслуживания РАН и регионов России Всероссийского института научной и технической информации РАН (ВИНИТИ РАН). Специалист в области наукометрии, автор более 150 статей и глав в монографиях; более 30 лет ведет преподавательскую деятельность в России и в качестве visiting professor читала курс лекций по наукометрии в колледжах по библиотечным и информационным наукам в Catholic University of America, University of Tennessee, Alabama University; эксперт Российского гуманитарного научного фонда; член редколлегии научных журналов: «Научно-техническая информация», «European Science Editing», Collnet Journal of Scientometrics and Information Management.
markusova@viniti.ru

Valentina Markusova, Ph.D. (Dr. Sc.)

Head of the Information Department, All Russian Institute for Scientific and Technical Information (VINITI), the Russian Academy of Sciences.
Author of more than 150 publications, including articles in Russian and international scholarly journals, and books chapters. More than 35 years of experience in teaching at Russian and American universities. Editorial board member: Scientific and Technical Information, European Science Editing, Collnet Journal of Scientometrics and Information Management. Member of ISSI and COLLNET.
markusova@viniti.ru

Ольга Васильевна Москалева

Советник директора Научной библиотеки им. М. Горького Санкт-Петербургского государственного университета, кандидат биологических наук, член совета Научной электронной библиотеки РФФИ, член редакционного совета журнала «Научная периодика. Проблемы и решения».
o.moskaleva@spbu.ru

Olga Moskaleva, Ph.D.

Advisor to the Director of Scientific Library, Saint Petersburg State University.
Member of the Board of the Scientific Electronic Library of the Russian Foundation for Fundamental Research, Editorial board member, Scientific Communications. Problems and Solutions.
o.moskaleva@spbu.ru

Владимир Владимирович Писляков

Заместитель директора библиотеки Национального исследовательского университета «Высшая школа экономики», кандидат физико-математических наук. Автор более 30 научных работ, среди которых статьи в журналах Journal of the Association for Information Science and Technology, Scientometrics, «Научно-техническая информация», «Управление большими системами» и др.; рецензент журналов Journal of the Association for Information Science and Technology, Science and Public Policy, «Вопросы образования», «Информационно-управляющие системы» и др.
pislyakov@hse.ru

Vladimir Pislyakov, Ph.D.

Assistant Library Director, National Research University Higher School of Economics.
Author of more than 30 publications, including articles in the Journal of the Association for Information Science and Technology, Scientometrics, Nauchno-technicheskaya Informatsia (Scientific and Technical Information, in Russian) and other journals. Reviewer of the Journal of the Association for Information Science and Technology, Science and Public Policy, Voprosy Obrazovania (Problems of Education, in Russian) and other journals.
pislyakov@hse.ru

ВВЕДЕНИЕ. К 50-ЛЕТИЮ SCIENCE CITATION INDEX:

ИСТОРИЯ И РАЗВИТИЕ НАУКОМЕТРИИ



В. А. Маркусова

Доктор педагогических наук, заведующая отделением научно-информационного обслуживания РАН и регионов России ВИНИТИ РАН.

INTRODUCTION. 50TH ANNIVERSARY OF THE SCIENCE CITATION INDEX: HISTORY AND EVOLUTION OF SCIENTOMETRICS

DOI 10.15826/B978-5-7996-1352-5.0003

This chapter is a short history of the breakthrough idea proposed by Eugene Garfield to use citations analysis for scholarly information search. It describes how he turned this idea into reality, through creation of Science Citation Index and other information resources and tools. The chapter further discusses the usage of citation analysis for the study of cognitive structure of science, collaboration networks, informal and formal scholarly communications, and research evaluation. The chapter covers the development of scientometrics as a scholarly discipline, with a special focus on development of scientometrics in Russia. The role of the Internet in development of modern scientometric resources is considered, particularly of the Web of Science platform. Finally, the growing demand for scientometric tools as means of research assessment for government bodies, universities, and research institutions is discussed.

Во введении рассмотрена история уникальной идеи Ю. Гарфилда по использованию научных ссылок как средства информационного поиска и взаимосвязи новых научных достижений со старыми и борьба за воплощение этой идеи в жизнь и создание универсального инструмента для поиска научной литературы Science Citation Index (SCI). Обсуждается история методов анализа цитирования и их применение для изучения развития научных идей и направлений, изучения социальной структуры науки и сетей научного соавторства. Описывается рождение и становление наукометрии, развитие ее методологии и создание новых информационных продуктов на основе новейших достижений информационных технологий на примере платформы Web of Science. Отмечается непрерывно возрастающая потребность государственных органов всех стран в использовании библиометрических индикаторов как средства оценки эффективности научной деятельности университетов и научных организаций.

Я думаю, что Вы делаете историю, Джин.

*Дж. Ледерберг,
президент Рокфеллеровского университета
New York, 1978–1990 гг.
лауреат Нобелевской премии
по физиологии и медицине (1958 г.)
(письмо Ю. Гарфилду от 24.01.1962 г.)*

*Юджин Гарфилд внес выдающийся вклад
в построение упорядоченных картин в миро-
вой науке. Его средством являются не сами
исходные данные, а восприятия ученых, ко-
торые исследуют неизвестное в природе.*

*Барух Блумберг,
лауреат Нобелевской премии по физио-
логии и медицине (1976 г.)*

В 2014 г. исполнилось 50 лет со дня выпуска первого в мире издания Science Citation Index — SCI (Указателя цитированной литературы), созданного доктором Ю. Гарфилдом (Eugene Garfield), регулярный выпуск которого начался в 1964 г. в Institute for Scientific Information — ISI (Институте научной информации США). На основе анализа библиографических массивов этого издания появилась новая научная дисциплина — наукометрия (или, как теперь чаще говорят, библиометрия). Сама история создания этого указателя и его повсеместное использование как инструмента изучения отдельных направлений науки, сетей научных коммуникаций, средства оценки результативности научных исследований заслуживают внимания.

В середине 40-х г. двадцатого столетия возникла проблема информационного кризиса, связанного с ростом объемов научных исследований и коллективов, а также с возникновением мультидисциплинарных проектов. Чтобы успешно выполнить свою главную социальную роль (которая заключает-

ся в производстве нового знания), ученый непременно должен быть информирован о том, что было известно до него. Выдающийся американский ученый и администратор науки д-р Ванневар Буш (Vannevar Bush) в своей книге Science — The Endless Frontier («Предел науки — бесконечность») призвал ученых сделать доступным обширный тезаурус знаний (багаж) и таким образом расширить физические и умственные способности человека. Извлекая уроки из опыта Второй мировой войны и руководства шестью тысячами ведущих американских ученых в области использования научных достижений в военных целях, он увидел насущную необходимость удовлетворения их информационных и коммуникационных потребностей. По его мнению, одним из важных уроков Второй мировой войны было осознание того, что научный прогресс необходим для блага страны и наука должна быть особой заботой правительства [Bush, 1945]. Его усилия привели позже к созданию в 1950 г. фонда National Science Foundation — NSF (Национального научного

фонда), агентства, которое сыграло важную роль в развитии системы научно-технической информации (НТИ) в США.

Об информационном кризисе была написана масса статей, и эта проблема вызывала горячие дискуссии в научном сообществе, а также в администрации и среди политиков США. Казалось, что мир тонет в потоке информации. Традиционные предметные указатели и реферативные службы, построенные по отдельным направлениям знания, не решали проблемы поиска научной литературы по таким проблемам, как атомный проект, исследование космоса или интенсивно развивающиеся масштабные исследования по генетике. Научное сообщество нуждалось в принципиально новом подходе для решения этих задач с использованием зарождающейся вычислительной техники.

В это время молодой химик со степенью бакалавра Ю. Гарфилд поступил на работу в химическую фирму. В лаборатории раскрылись его способности не в экспериментальной химии (был ряд неудачных экспериментов), а в синтезе научной информации. В фирме он узнал о существовании старейшего реферативного журнала *Chemical Abstracts* и создал обширный указатель химических препаратов, имеющихся в лаборатории. Изучая научную литературу, он обратил внимание на обзорные статьи, каждый параграф которых подтверждался соответствующей ссылкой. Этот факт привел его к пониманию ассоциативной связи между новой публикацией и работой предшественников. Примерно в это же время он попал на годовичное собрание Американского химического общества (*American Chemical Society* — CAS) и увидел возможности сочетания профессии химика с информационными технологиями. Вскоре по рекомен-

дации директора отдела химической литературы CAS Дж. Перри (J. Perry) он получил должность ассистента в проекте *Welsh Machine Project* (далее — *Welsh*) по машинному индексированию в Медицинской библиотеке Университета Дж. Гопкинса (*John Hopkins University*). Проект возник как один из способов борьбы с «информационным кризисом» в медицине и необходимостью разработки нового способа ее обработки с помощью машинных методов [Михайлов и др., 1965].

В 1953 г. по результатам проекта *Welsh* был проведен первый симпозиум по машинному индексированию *The First Symposium on Machine Methods in Scientific Documentation*, который привлек огромное внимание масс-медиа благодаря фразе его организатора профессора Л. Рида (L. Reed — вице-президент Университета Дж. Гопкинса) «человек тонет в потоке информации» [Hyslop, 1953]. Вследствие этого внимания прессы в адрес проекта пришло много писем, среди которых было письмо от У. Адейра (W. Adair), ушедшего в отставку вице-президента юридической компании «Шепард». Эта компания издавала с 1873 г. *Shepard's Citations* («Указатель ссылок Шепарда»). У. Адейр предложил рассмотреть возможность использования метода цитирования, применявшегося в указателе *Shepard's Citations*, для индексирования медицинской литературы. По воспоминаниям Ю. Гарфилда, он буквально воскликнул «эврика!», поскольку понял, что принцип цитирования может быть использован для индексирования не только обзорных статей по медицине, содержащих много библиографических ссылок, но и всей медицинской литературы. Он стал переписываться с У. Адейром и работать над статьей об указателях цитирования для научной литературы.

Необходимо отметить, что, участвуя в осуществлении проекта Welsh, Ю. Гарфилд разработал новые методы подготовки предметных указателей, а также информационного поиска на счетно-перфорационных машинах, применение которых в то время было одним из главных направлений механизации информационных процессов [Гиляревский и др., 1995]. Когда завершились работы по проекту Welsh, Ю. Гарфилд поступил в аспирантуру Columbia University (Нью-Йорк) по специальности «Библиотечноеведение». В это время он поддерживал переписку с У. Адейром и начал работать над подробной статьей об указателях цитирования для научной литературы. Эта статья была написана им в 1954 г. И пока она ожидала опубликования в журнале Science, Ю. Гарфилд, который к тому времени стал помощником редактора журнала American Documentation, предложил У. Адейру написать для этого журнала статью, в которой был бы кратко описан механизм функционирования указателя Shepard's Citations. Такая статья У. Адейра появилась в журнале American Documentation в июне 1955 г. [Adair, 1955], а статья Ю. Гарфилда Citation Indexing for Science была опубликована в журнале Science в июле 1955 г. [Garfield, 1955]. Идеи, изложенные в этой статье, в частности соображения о возможности создания мультидисциплинарного указателя цитирования, самым непосредственным образом касались решения проблемы поиска научной литературы. То есть появлялась реальная возможность реализовать мечту о «мировом мозге», которая со времен Г. Уэллса занимала многих выдающихся ученых [Wouters, 1999].

Работая над проектом Welsh, Ю. Гарфилд знакомится с дайджестом Current List of Medical Literature и решает создать соб-

ственный информационный продукт. В 1956 г. он организовал фирму Eugene Garfield Associates, Information Engineers и начал выпускать первое в мире недельное сигнальное издание библиографической информации Current Contents, содержащее оглавления только что вышедших научных журналов в области фармацевтики и биомедицины, а также пермутационный¹ и авторский указатели. Вскоре издание Current Contents охватило самые разные области знания от физики до психологии и литературы. Еженедельные выпуски Current Contents стали необходимым источником информации для мирового научного сообщества [Маркусова, Черный, 1985]. Издание оказалось чрезвычайно успешным, и в 1960 г. Ю. Гарфилд переименовал свою компанию в Institute for Scientific Information. О выборе этого названия Ю. Гарфилд писал в 1978 г.: «Несомненно, мы были воодушевлены (inspired) созданием в СССР Всесоюзного института научной и технической информации. Кроме того, для многих людей слово “институт” дополнительно означает бесприбыльную организацию. В те дни многие из моих коллег и заказчиков испытывали антипатию к коммерческим организациям в области информации. По-видимому, многие из них все еще придерживаются этого мнения» [Garfield, 1980]².

Популярность издания Current Contents в научном сообществе была настолько значительной, что его юбилею была посвящена

¹Пермутационный предметный указатель (от англ. слова «permuterm» — перестановка) — это был новый вид указателя, подготовка осуществлялась с конца 40-х гг. XX в. на больших ЭВМ. Указатель состоял из слов (их называли «key words» — ключевыми словами) заглавий публикаций, рядом с которыми помещалась фамилия автора статьи источника.

²С 1952 по 1955 г. ВИНТИ назывался Институт научной информации АН СССР.

статья в журнале Science [Broad, 1978]. Интересный факт, характеризующий популярность Current Contents среди ученых, приведен в этой статье. Автор — биохимик из Кембриджского университета (Англия), — опубликовал в журнале Analytical Biochemistry статью, описывающую теорию модификации белка. Когда к этому ученому стали поступать запросы на оттиски статьи, то из-за различия в написании адреса в журнальной статье и в авторском указателе Current Contents он легко смог определить, кто узнал о его работе из журнала Analytical Biochemistry, а кто — из Current Contents. Оказалось, что 2100 запросов поступило от ученых, просматривающих Current Contents, и всего 350 запросов от ученых, читающих Analytical Biochemistry. Следует отметить, что появление Интернета и его широкое использование редакциями научных журналов для размещения оглавлений новых номеров сказалось на значимости этого информационного инструмента.

Задуманная и воплощенная Юджином в Current Contents идея верификации фамилий авторов и организаций получила дальнейшее развитие при создании в Web of Science (WoS) специальных инструментов для верификации: авторского профиля исследователя ResearcherID и объединенного профиля организации Organization Enhanced. На сайте ResearcherID³, доступном любому пользователю Интернета, исследователь может создать список своих собственных публикаций, включая публикации, индексированные в Web of Science. Это позволяет затем проводить поиск в WoS по номеру профиля ResearcherID и находить все индексированные публикации автора, даже при наличии ошибок и разных вариан-

тов написания его фамилии. В объединенный профиль Organization Enhanced объединяются различные варианты названий организаций (вузов, научных институтов и т. д.), указанных в публикациях в качестве аффилиаций (мест работы) авторов. Объединенные профили Organization Enhanced значительно облегчают сбор и анализ статистики по публикационной активности и цитируемости в целом по организациям.

О популярности и влиянии Current Contents свидетельствует тот факт, что редакторы научных журналов, настаивающие на включении издаваемых ими журналов в Current Contents, утверждают, что охват или не охват выпускаемых ими изданий означает жизнь или смерть научного журнала. По мнению этих редакторов, ученые не хотят представлять статьи для публикации в журнал, если он не отражается в Current Contents. На обложках многих зарубежных журналов написано: «отражается в Current Contents, Chemical Abstracts, Science Citation Index». Именно финансовый успех этого издания позволил Ю. Гарфилду приступить к выпуску Science Citation Index (Указателя цитированной литературы), который принес ему и ISI мировую известность.

Необходимо отметить, что успех серий Current Contents был связан также с тщательным отбором наиболее престижных журналов, соответствующих тематике направлений отдельных серий этого издания. Этот принцип отбора основан на использовании «закона рассеяния» научных публикаций в мировой периодике, сформулированного в 30-х гг. прошлого века английским библиографом и статистиком С. Брэдфордом (S. Bradford) [Михайлов и др., 1965]. Как писал Брэдфорд, «...статья, интересующая специалиста, может появиться не только в периодике, специ-

³<http://researcherid.com>

ализирующей на его предмете исследования, но также, время от времени, в других периодических изданиях, число которых растёт относительно этих областей таким образом, что уменьшается материал по его предмету исследования, и количество публикаций в каждом из этих периодических изданий снижается» [Garfield, 1971]. Согласно формулировке Брэдфорда, совокупность публикаций по какой-либо отрасли можно разделить на три концентрических круга, первый из которых содержит небольшое число ядерных (ключевых) журналов по этой области, вторая группа содержит значительно большее количество журналов, смежных с данной отраслью, и, наконец, третья группа содержит огромное количество журналов, в которых появление публикаций по этой области трудно предполагать. Соотношение количества журналов в этих трех группах выглядит как $1 : n : n^2$, где n — постоянная, зависящая от тематики области [Михайлов и др., 1965]. Именно эта закономерность является одним из основных принципов для комплектования фондов библиотек и деятельности всех информационных реферативных служб, в том числе создания SCI и других информационных продуктов ISI. Закон Брэдфорда объясняет, почему мультидисциплинарный указатель SCI в целом значительно более эффективен, чем любые другие дисциплинарно ориентированные службы. Сам Ю. Гарфилд называет закон Брэдфорда законом «концентрации» информации [Garfield, 1971].

Создание SCI было чрезвычайно затратным, и на протяжении нескольких лет Ю. Гарфилд безуспешно пытался получить грант от NSF — государственной организации США. Одной из проблем являлся тот факт, что NSF не был склонен поддерживать малоизвестную коммерческую компа-

нию, которой являлась фирма Ю. Гарфилда. В какой-то мере первая финансовая поддержка осуществлению издания SCI была связана и с выдающимися успехами СССР в исследовании космоса. Известно, что запуск советского спутника 4 октября 1957 г. имел огромный резонанс в мире и застал врасплох президента Д. Эйзенхауэра и правительство США. Этот факт был сигналом для общества о значимости научно-технической информации. Администрация Белого дома и научное сообщество недоумевали, почему они не знали из научной литературы о научном и техническом потенциале СССР. Сенатор Х. Хамфри (Hubert Humphrey), впоследствии вице-президент США, объявил, что «наступил новый век науки — век информатики», и потребовал объяснений, «что исполнительная власть собирается сделать для управления потоками информации, генерируемой за счет миллиардов долларов, выделяемых правительством на научные исследования»⁴. Президент Д. Эйзенхауэр назначил специальную комиссию для оценки состояния обеспечения ученых научно-технической информацией. В докладе комиссии уделялось особое внимание организации переводов советской научной литературы и рассматривалась возможность создания государственного центра по обработке и распространению информации — аналога ВИНИТИ (!!!). Аналог ВИНИТИ решили не создавать, но для осуществления научных исследований в документалистике, как тогда называли информатику, по рекомендации комиссии в NSF был создан отдел НТИ [Wouters, 1999]. Этот отдел сыграл важную роль в судьбе SCI.

В истории создания SCI большую роль сыграла поддержка двух выдающихся

⁴Материал взят с сайта NFAIS (URL: <http://www.nfaais.org/>).

американских ученых. В январе 1957 г. Ю. Гарфилд получил весьма ободряющее письмо от известного генетика профессора Гордона Аллена (Gordon Allen), занимавшего ответственный пост в National Institutes of Health — NIH⁵ (Национальном институте здравоохранения). Г. Аллен писал, что после ознакомления со статьей Ю. Гарфилда в Science он с нетерпением ждал новостей по созданию SCI. По своей инициативе он обращался в American Society on Human Genetics — ASHG (Американское общество по генетике человека) и призывал к взятию инициативы по участию в таком проекте. Однако заинтересованности не было проявлено, поскольку отдел информации ASHG занимался созданием классификационной системы по генетике человека. Г. Аллен полагал, что группа заинтересованных специалистов смогла бы привести к прогрессу в этом вопросе. Началась активная переписка с Алленом, которому Ю. Гарфилд рассказал о своих неудачных попытках получения гранта [Wouters, 1999].

В 1957 г. в Комитете конгресса США состоялись специальные слушания представителей NSF о мерах борьбы с информационным кризисом. В результате этого NSF опубликовал заявление о желании провести исследование, основанное на использовании метода цитирования по аналогии с Shepard's Citations. Как отмечалось в заявлении, «...этот метод, заслуживший признание в области юриспруденции, ранее никогда не был использован в естественных науках» [Wouters, 1999]. Ю. Гарфилд узнал

об этом позднее и направил заявку на грант в NSF. В этот раз заявка была направлена совместно с Г. Бедфордом (G. Bedford), занимавшим тогда должность профессора-ассистента в Университете Пенсильвании (University of Pennsylvania). Проект должен был быть двухлетним исследованием и завершиться созданием указателя. Несмотря на поддержку А. Гордона, проект был отклонен, но NSF выразил интерес к методу индексирования ссылок. Г. Аллен, служивший экспертом в NSF, предупреждал президента ASHG о необходимости создания заинтересованной и компетентной группы для рассмотрения заявки на проект [Wouters, 1999].

Летом 1958 г. Ю. Гарфилд получил письмо от профессора Дж. Ледерберга (J. Lederberg), лауреата Нобелевской премии по медицине в 1958 г., выразившего энтузиазм в поддержке проекта Ю. Гарфилда. Он был шокирован, узнав о неудачных попытках Юджина получить грант и поведении бюрократов NSF. Одновременно он предупредил Ю. Гарфилда о необходимости продолжать дискуссии с NSF. Узнав от Ю. Гарфилда о заинтересованности профессора Г. Аллена в создании SCI, Ледерберг предложил объединить усилия и начать совместную борьбу за поддержку исследования и пропаганду этой идеи среди генетиков и специалистов по молекулярной биологии. В марте 1960 г. Ю. Гарфилд направил переработанную заявку в NSF и попросил Дж. Ледерберга поддерживать проект.

Дж. Ледерберг обратился в NSF со следующим письмом: «...я убежден, что мое письмо мало что добавит к утверждению, что научное сообщество столкнулось с огромными проблемами, пытаясь ознакомиться с текущей научной литературой. Нет универсального решения этой проблемы; хотя отдель-

⁵National Institutes of Health — государственная научная организация США с годовым бюджетом в 2012 г. 33,6 млрд долл., имеющая под своей эгидой 40 научных институтов и финансирующая на конкурсной основе все исследования по биомедицине и медицине в стране. Для сравнения, годовой бюджет National Science Foundation — 7,2 млрд долл.

ные реферативные службы выполняют полезные функции. С моей точки зрения, и это после размышлений, SCI имел бы неоценимую ценность в улучшении эффективности научного исследования, поскольку оно зависит от поиска литературы. Я глубоко заинтересован и с огромным энтузиазмом отношусь к успеху д-ра Гарфилда и счастлив иметь возможность поддержать эту попытку, служа советником в его консультативном комитете» [Lederberg, 2000].

Пока в NSF размышляли о вынесении решения по проекту Ю. Гарфилда, он направил заявку в NIH. Начались его совместные консультации с обеими организациями по поддержке проекта. NIH был заинтересован в создании Genetic Citation Index — GCI (Индекса цитируемости по генетике). В NSF, обеспокоенные схожестью двух заявок, предложили составить единый бюджет, объединивший задачи по выполнению проекта. В канун католического Рождества Ю. Гарфилд получил ответ. В письме Дж. Ледербергу он писал: «Санта Клаус был очень мил к нам. Мы услышали, что NIH утвердил половину бюджета, который NSF просил меня представить в сумме 100 тыс. долл. в течение трех лет» [Wouters, 1999]. В 1961 г. Ю. Гарфилд получил грант в 300 тыс. долл. для создания GCI. В этот же год Ю. Гарфилд завершил образование по структурной лингвистике, получив степень доктора философии в Пенсильванском университете. Его диссертация называлась «Алгоритм для перевода химических названий в молекулярную формулу» [Гиляревский и др., 1995].

Прежде чем создать Genetic Citation Index, Ю. Гарфилду было необходимо создать Указатель цитированной литературы, охватывающий все области знания. Для создания SCI Ю. Гарфилд использовал массив из 613

ведущих научных журналов, обрабатываемых в его компании для подготовки различных серий издания Current Contents. Все журналы, используемые для подготовки SCI, называются журналами-источниками (source journals). Статьи, опубликованные в этих журналах, называются статьями-источниками (source article). Авторы этих статей-источников называются цитирующими авторами (citing authors). Ссылки, содержащиеся в этих статьях, называются цитируемыми статьями (cited articles), а их авторов — цитируемыми авторами (cited authors).

Genetic Citation Index был создан на массиве статей за 1961 г., опубликованных в 613 научных журналах-источниках, и 1,4 млн содержащихся в них статей (цитируемых статей). Первое печатное издание SCI появилось в 1963 г., и с 1964 г. оно стало выходить ежеквартально, а последний том был кумулятивный. SCI состоял из четырех частей, построенных по алфавитному признаку: Source Index (указатель источников), Citation Index (указатель ссылок), Permuterm Subject Index (пермутационный указатель ключевых слов из заглавий статей-источников) и Corporated Index (указатель организаций, в которых работали авторы статей-источников).

Идею использования ссылок как поискового инструмента проверил и профессор Г. Аллен в 1961 г. Он построил схему использования и взаимовлияния полтора десятков работ по вопросу выявления нуклеиновых кислот, опубликованных в 1941–1960 гг. Было установлено, что для получения сведений обо всех этих работах нужно было обратиться к библиографии по меньшей мере шести статей, опубликованных в период с 1956 по 1960 г. Отдельные статьи, которые почти не цитировались, вообще трудно выявить таким путем. Однако

сведения почти обо всех работах можно легко получить, если узнать, где содержатся ссылки на три ранние работы. Это было важное подтверждение идеи работоспособности метода цитирования, выполненное известным специалистом-генетиком [Михайлов и др., 1965].

В основу создания Current Contents был положен постулат о том, что научная литература является формой существования науки, а научная статья является общепринятой нормой для распространения и оценки результатов исследования, а также решающим (критическим) фактором для продвижения по служебной лестнице в большинстве научных дисциплин.

Нам всем хорошо знакома американская поговорка «публикуйся или погибни» (“publish or perish”). Эти слова означают постоянное давление на исследователей для опубликования своих результатов и создания имени в науке. Но это не новый феномен. Эта крылатая фраза, сказанная в 1950 г. американским генетиком К. Атвудом (K. C. Atwood) из Columbia University (Нью-Йорк) и не опубликованная в то время, проникла в сердце каждого исследователя. Согласно легенде, месяц спустя она вернулась назад к К. Атвуду, когда его посетил профессор из другого университета. Несмотря на давнюю историю, эта фраза будет звучать в холлах университетов и в аудиториях столько, сколько будет существовать соревнование между исследователями за получение финансовых ассигнований. Десятки тысяч статей, заметок, писем в редакцию и обзоров, которые публикуются ежедневно в научных журналах, и миллионы ссылок между этими статьями обеспечивают путь проникновения в коммуникацию знания, способствуют процессам его распространения в науке

и получению эмпирических данных о значимости исследования и научной активности «единицы» производства знания. По образному выражению Д. Прайса, научная статья стала в XX в. «атомом научной связи» [Price, 1961].

В наши дни, как и 100 лет назад, статья обычно содержит ссылки на работы предшественников [Price, 1963]. Отсутствие ссылок в статье начинающего автора рассматривается как один из признаков низкой квалификации и затрудняет публикацию. Индексирование документа посредством использования содержащихся в нем библиографических ссылок основано на простой мысли, что указанные в этих ссылках другие, ранее опубликованные, документы близки по тематике данному документу. Но эту идею было очень непросто воплотить в инструмент практического использования. Что же такое «ссылки» и что подразумевается, когда один автор цитирует другого?

Теоретической основой указателей цитирования является тот факт, что ссылки представляют собой символы научных концепций. Являясь формой научной коммуникации, ссылки помогают проследить развитие вклада автора публикации, придают достоверность работе исследователя. Социологи говорят, что ссылки — это своеобразная валюта, которой ученые «оплачивают» вклад предшественников. Ученые обычно отдают должное исследованиям, которые оказали влияние на них, цитируя релевантные источники (статьи, книги, заметки, обзоры и т. д.) в списке использованной литературы или сносок в этих публикациях. Таким образом, публикация — это своеобразная дорога, на которой оставлены следы деятельности самого исследователя, его научных связей и использования

внешних научных источников. Цитировать кого-то — это признать воздействие предшественника на вашу работу, своеобразный способ вернуть интеллектуальный долг человеку, которого вы цитируете.

Причины цитирования публикаций:

- признание вклада;
- определение/описание методов, оборудования;
- корректирование собственной работы или работы других;
- критика предыдущих работ;
- указание дополнительной литературы по данной теме;
- отрицание предыдущих работ;
- подтверждение данных;
- оказание уважения.

Говоря о цитируемости, нужно различать два понятия: «citation» и «references». Эти два английских слова являются синонимами, но мы используем слово «citation», когда говорим о цитируемости статьи, а слово «references» — когда говорим о списке литературы, приведенном в статье.

При традиционном поиске ученый, просматривая литературу, находит ссылки на работы предшественников и изучает их. Противоположный путь предлагает метод цитирования. Исследователю нужно знать фамилию его предшественника и, обратившись к указателю библиографических ссылок, он может найти все современные работы, цитирующие работы этого предшественника. Поскольку за каждой ссылкой скрыт ее автор, она оказалась знаком не только научного результата, в поисках которого ученый обращается к тексту, но и конкретного лица, с которым автор текста пожелал соотносить свой результат. Там, где есть люди, неизбежно вступают в действие психологические факторы. Новые информационные технологии, изобретенные для решения

информационных задач, которые прежде решались библиографическими средствами, позволили новыми глазами взглянуть на социальную жизнь науки, вторгнуться в глубины психологии ее людей. Мотивация цитирования стала предметом изучения психологов и социологов науки. Основные положения и принципы теории учета, анализа и использования цитирования были позднее изложены Ю. Гарфилдом в его монографии [Garfield, 1979].

После выпуска тестового варианта Genetic Index NSF отказался поддержать идею создания мультидисциплинарного SCI. Позднее Ю. Гарфилд писал: «В 1964 г. я принял наиболее важное финансовое решение в истории ISI — начать выпуск индекса. понадобилось пять лет, чтобы превратить это издание в издание, приносящее прибыль. Однако в 1964 г. мы не были готовы рисковать и пропагандировать SCI как инструмент для оценки научной производительности. Но сейчас мы готовы открыть ящик Пандоры, и пусть научное сообщество решает, насколько надежен SCI как социометрический инструмент» [Garfield, 1969].

Большую роль в пропаганде SCI играли статьи Ю. Гарфилда в Current Contents, публиковавшиеся с 1962 г. Сначала они появлялись эпизодически, однако с 1972 по 1992 г. включительно стали еженедельными. Эти статьи привлекали неизменное внимание широких кругов ученых и специалистов, были посвящены той или иной актуальной проблеме науки, научной коммуникации или информатики. Они составили 15 томов сборников Essays of an Information Scientist (1977–1994 гг.), образующих своего рода энциклопедию информатики, в которой можно найти сведения по всем важным вопросам наукометрии, науковедения и социологии науки.

В начальный период своего развития наукометрия рассматривалась как часть науковедения. Импульсом к развитию науковедения явились работы английского философа Дж. Бернала (J. D. Bernal), прежде всего опубликованная им в 1939 г. книга «Социальная функция науки» (The Social Function of Science), ставшая классикой [Garfield, 1983].

В этой монографии излагался совершенно новый подход к роли науки в жизни общества и его процветании. Как отмечал биограф Дж. Бернала Эндрю Браун (Andrew Brown), «Социальная функция науки» была попыткой Бернала обратить внимание общества на тот факт, что «...наука должна быть не охраняемой территорией интеллектуалов, а являться неперменной составляющей улучшения жизни человечества» [The social function..., 1953]. Это был новаторский трактат как в изучении возможностей науки и техники, так и в создании государственной политики, отвечающей за использование достижений науки и техники. Впоследствии Ю. Гарфилд вспоминал, какое огромное впечатление произвела на него эта книга. В 1954 г. вышла четырехтомная монография Дж. Бернала Science in History, анализирующая взаимоотношения науки и общества. Ю. Гарфилд обратился к профессору Берналу в 1962 г., и тот с воодушевлением отнесся к его идее создания SCI. Правда, как вспоминал Ю. Гарфилд, «Бернал не верил, что идею создания SCI будет легко осуществить» [Garfield, 1983]. Бернал писал: «Я думаю, что SCI внесет что-то новое и ценное в годичную премию за обзоры, присуждаемую наукой по документалистике» [Garfield, 1983]. С 1964 г. профессор Бернал вошел в редакционный совет ISI по созданию SCI. По мнению Ю. Гарфилда, профессора Бернала можно считать "grandfather" of science of science («дедушкой» науковедения).

Невозможно переоценить вклад, внесенный в развитие науковедения и наукометрии выдающимся американским историком науки, физиком по образованию, профессором Yale University (штат Коннектикут) Дерекком де Солла Прайсом (Derek de Solla Price). Д. Прайс был советником по науке президентов США Дж. Кеннеди и Л. Джонсона. Его книги «Наука после Вавилона» (Science after Babilone) [Price, 1961] и «Малая наука, большая наука» (Little Science, Big Science) [Price, 1963] стали мировыми бестселлерами, были переведены и изданы в СССР. В книге «Малая наука, большая наука» Д. Прайс изложил свой подход к исследованию развития науки количественными методами и обосновал идею экспоненциального роста научной литературы. Д. Прайс был горячим сторонником роли государственных инвестиций в науку, поскольку он полагал, что «наиболее характерной чертой развития общества является настоятельная потребность в науке» [Price, 1963]. Блестящий специалист и оратор, он активно включился в исследования по анализу массивов SCI и указал на огромные потенциальные возможности нового инструмента для изучения социологии и истории науки. Он одним из первых выполнил исследование на массиве сетей библиографических ссылок. В 1965 г. Д. Прайс в статье Network of Scientific Papers, опубликованной в журнале Science, отмечал, что изучение связей между документами по их цитированию может позволить описывать в географических терминах структуру науки, «в которой разделы науки воспринимаются как карты территории с национальными границами, горами и пастбищами, свободными дорогами и непроходимыми джунглями» [Price, 1965]. По мнению одного из наиболее ав-

торитетных специалистов по наукометрии профессора Б. Гриффита (Belver C. Griffith, Drexel University), «это блестящее исследование Д. Прайса [Price, 1965] позволило впервые построить сети цитирования, которые послужили прообразом карт науки» [Small, Griffith, 1974]. Д. Прайс первым не только построил сети цитирования, но и разработал математическую теорию роста этих сетей, сформулировал закон Прайса о старении научной литературы. В память о выдающемся вкладе профессора Д. Прайса в наукометрию и социологию науки Ю. Гарфилд учредил в 1984 г. золотую медаль имени Д. Прайса, присуждаемую раз в два года Международной организацией по наукометрии и информетрии (International Society on Scientometrics and Informetrics — ISSI). Активная переписка и поддержка Д. Прайса, философа и политика профессора Дж. Бернала и выдающегося социолога Р. Мертон укрепили веру Ю. Гарфилда в то, что новый информационный инструмент может быть использован для эмпирического изучения истории и социологии науки.

Регулярный выпуск SCI с 1964 г. вызвал большой интерес в научном сообществе, и ряд выдающихся исследователей стали инициаторами и разработчиками методов наукометрии на основе анализа массивов SCI. Ученые осознали, что эти массивы являются ценным источником изучения отдельных направлений науки, сетей научных коммуникаций и средством оценки результативности научных исследований. Первое кумулятивное издание SCI появилось в ВИНИТИ АН СССР уже в 1965 г., и с тех пор началось приобретение этого издания ведущими библиотеками и АН СССР. Детальное описание SCI и принципы, заложенные при его создании, были

освещены в книге «Основы информатики», опубликованной в 1965 г. [Михайлов и др., 1965]. Эта монография и последующая книга «Научные коммуникации и информатика» [Михайлов и др., 1976] были переведены на английский, немецкий, французский и другие языки мира и изданы в США, Европе и Японии. Можно констатировать, что обе монографии оказали колоссальное влияние на огромное поколение библиотечных и информационных специалистов в СССР, странах Восточной Европы, Индии, Бразилии и других странах мира.

Следует подчеркнуть, что интенсивное сотрудничество Ю. Гарфилда и Дж. Ледерберга стало частью научной политики и дебатов в американском научном сообществе в борьбе с информационным кризисом и соперничеством США с СССР в космической гонке. Президент Кеннеди придавал огромное значение науке и ее роли в жизни общества, и его сильно беспокоили успехи СССР в космосе. В 1963 г. по решению Дж. Кеннеди была учреждена специальная комиссия под председательством д-ра А. Вайнберга (Alvin Weinberg) для оценки состояния научно-технической информации. Д-р А. Вайнберг — сын российских эмигрантов, участник Манхэттенского проекта и директор в течение 18 лет Окриджской национальной лаборатории (Oak Ridge National Laboratory) — был известен не только блестящими научными достижениями, но и удивительной способностью убеждения в общественной значимости науки для процветания страны. В результате работы комиссии появился знаменитый доклад «Наука, правительство, информация» (Science, Government, and Information), известный в мире как доклад А. Вайнберга. В докладе отмечалось, что ученые «похоронены под массивами информации» [Weinberg, 1963], с которыми

они не успевают ознакомиться, и информационный кризис рассматривался как угроза существованию самой науки, в то время как получение информации является интегральной и неотделимой частью научного процесса. В докладе комиссии, членом которой был Дж. Ледерберг, отмечалось: «Комиссия находится под впечатлением возможностей Science Citation Index. Мы хотели бы привлечь внимание научно-технического сообщества к этому новому, мощному, хотя пока мало используемому, поисковому инструменту. Комиссия полагает, что индексирование ссылок, в частности в сочетании с пермутационным индексированием заголовков, будет широко использоваться и его применение изменит способ нашего осмысления технической литературы и способ управления ею» [Weinberg, 1963]. Эта рекомендация явилась важным шагом в истории развития SCI, массивы которого послужили основой для развития наукометрии.

Некоторые выводы доклада Вайнберга были учтены и в СССР. В 1966 г. было принято Постановление Правительства СССР о создании Всесоюзного научно-технического информационного центра — аналога американского Clearing House (Центра отчетов, выполненных по контрактам с государством США) и об организации системы депонирования.

Нельзя сказать, что создание SCI и использование метода ссылок не вызывало критики. Любопытно, что в 60-х гг. в качестве аргумента противники метода использовали очень высокую цитируемость работ Т. Д. Лысенко. Отвечая на этот упрек, известный американский документалист Н. Каплан (N. Kaplan) отмечал, что ссылки устанавливают ассоциативную связь между двумя документами, а их количество показывает влияние ученого. Положительное

это влияние или отрицательное, на этот вопрос ответят эксперты, но, без сомнения, разрушительное влияние Лысенко на развитие советской генетики было огромным [Kaplan, 1965]. Негативное цитирование незначительно и является статистической случайностью, особенно в таких огромных массивах, как Web of Science. Выполненные в ISI исследования показали, что негативное цитирование в определенной области знания обычно не превышает 5 % (а чаще меньше). История, подобная Cold fusion (холодный синтез) в 1970-х гг., случается крайне редко и подтверждает скорее исключение из правил. Многочисленные исследования показали, что частота цитирования положительно коррелируется с признанием исследователя научным сообществом. Негативное цитирование — это всего лишь «шум», который в действительности не влияет на результаты анализа. Противники использования метода цитирования также указывают на «избыточное» цитирование методических и обзорных статей и самоцитирование или цитирование кругом ученых друг друга. Самоцитирование — это нормальная практика в публикациях, и его доля в биомедицинской литературе может достигать 25 %. Автор, позволяющий себе чрезмерное самоцитирование, может вызвать возражения у рецензентов статьи и у редактора научного журнала, если самоцитирование нанесло ущерб ссылкам на другие работы. Конечно, такой автор может опубликовать свою работу в журнале с низким импакт-фактором и с невысокими стандартами рецензирования. В таком случае он навредит самому себе, поскольку практически теряет возможность быть процитированным другими исследователями. Цитирование кругом ученых только друг друга напоминает

историю о мифическом единороге: каждый может его вообразить, но никто не может его воспроизвести. Распределение ссылок всегда чрезвычайно несбалансировано. Небольшая группа статей может получить большое количество ссылок, и значительное большинство публикаций может быть процитировано незначительно или совсем не быть процитировано. Это явление относится даже к нобелевским лауреатам. Такое распределение ссылок описывается кривой нормального распределения и достаточно распространено в тех случаях, где человеческий фактор играет роль. Разные области исследования имеют совершенно разные модели цитируемости, и разрыв между областями знаний может достигать соотношения 10:1. Так, средняя статья по молекулярной биологии и генетике за десятилетний период получит 40 ссылок, а средняя статья в журнале по математике или вычислительной технике соберет не более четырех ссылок. То же самое относится и к общественным наукам, в которых в первую очередь цитируются книги [Маркусова, Ярошевский, 1987].

В 1965 г. Ю. Гарфилд организует компьютеризированную службу Selective Dissemination of Information — SDI (избирательного распределения информации) — Automatic Subject Citation Alert (ASCA). Подписчики еженедельно получали распечатки библиографических описаний статей по интересующей их тематике из 2000 научных журналов. Каждый подписчик имел свой профиль интересов. Например, редакторы научных журналов могли получать списки статей, в которых были процитированы публикации из их журналов. Подписчик этой службы имел также возможность получать библиографические описания статей определенных авторов

или публикаций, в которых были процитированы работы пользователя. Таким образом, SCI и ASCA⁶ кроме информационной составляющей способствовали установлению прямых контактов между учеными разных стран мира. Сам Ю. Гарфилд внимательно следит, где и как цитируются его работы и насколько растет популярность SCI. В 1970 г. Ю. Гарфилд получает письмо от известного украинского науковеда к. т. н. А. А. Коренного⁷ и пишет об этом в апреле 1970 г. в Current Contents: «...недавно я начал очень интересную переписку с советским ученым, который активно занят исследованиями по науковедению и использует SCI для социометрических и других целей. Я никогда бы не имел случая установить такой контакт, если бы я не создал службу ASCA, позволяющую мне получать информацию о публикациях, в которых были процитированы мои работы» [Garfield, 1970].

Оперативно работающая система оповещения на основе процитированных публикаций Ю. Гарфилда позволила ему установить связь и с его дальневосточным коллегой, членом-корреспондентом РАН В. Е. Васьковским. В интервью (телефонный разговор 14.09.2014 г.) автору этой главы он вспоминал: «...в 1975 г. мы опубликовали статью по реактивам на фосфор и фосфолипиды [Vaskovskii et al., 1975]. Потом наша статья попала в список 50 наиболее цитируемых работ из СССР. Она цитируется

⁶В 1968 г. Постановлением Президента АН СССР академика М. В. Келдыша служба, аналогичная ASCA, была организована для научно-информационного обслуживания академиков в ВИНТИ.

⁷Коренной А. А. — ведущий научный сотрудник сектора по истории науки в Институте кибернетики АН УССР, соратник и друг Г. М. Доброва. См.: Коренной А. А. Science Citation Index: Применение в науковедении и информатике // НТИ. 1966. Сер. 2. № 10. С. 3—6.

до сих пор, набрала более 600 цитирований». Начинаясь эта статья так: «Метод определения фосфора является одним из самых популярных в химических и биохимических лабораториях», и была ссылка на Current Contents. Вскоре Васьковский получил письмо от Ю. Гарфилда, в котором он спрашивал, как в далекой Сибири узнали о Current Contents. Ю. Гарфилд писал: «...я был в начале 60-х в Москве и давно мечтал проехать на поезде через Сибирь». «Я побежал к Андрею Петровичу (член-корреспондент А. П. Капица был председателем Дальневосточного научного центра) с этим письмом, он дал указание Боре Славинскому⁸ — визит Ю. Гарфилда был включен в план на 1977 г.». Начиная с этого времени Ю. Гарфилд неоднократно приезжал в СССР и Россию как гость Академии наук. Последний его визит во Владивосток состоялся в 2007 г.

Практически во время всех приездов Ю. Гарфилд выступал на семинарах, посвященных информационным продуктам ISI и их использованию. Эти семинары обычно организовывались в Институте органической химии АН СССР или в Библиотеке по естественным наукам АН СССР. Я хорошо помню, как на семинаре в 1981 г. Ю. Гарфилд рассказывал о доступе в режиме онлайн к банку данных ISI и специально созданных для этой работы новых удаленных ресурсах: ISI BIOMED и ISI COMPUMAS. В то время никто из нас даже не видел персонального компьютера. Семинары Ю. Гарфилда освещались в «Вестнике АН СССР» и журнале «Научно-техническая информация» (Сер. 1), выпускаемом ВИНТИ АН СССР. В настоящее время эту традицию семинаров продолжает высококвалифи-

цированный персонал компании Thomson Reuters, которая купила ISI в 1992 г.

С 1972 г. в ISI начался выпуск Social Science Citation Index — SSCI (Указателя цитируемой литературы по общественным наукам), а в 1978 г. появился Arts & Humanities Citation Index — A&HCI (Указатель цитируемой литературы по искусству и литературе). Одним из важных этапов в развитии наукометрии стал ежегодный, начиная с 1975 г., выпуск Journal Citation Reports — JCR (Указателя цитируемости научных журналов). Первое издание JCR вышло в печатном виде и содержало информацию об импакт-факторах около 3000 научных журналов. В 1978 г. началась публикация JCR-Social Sciences (Указателя цитируемости научных журналов по общественным наукам). «Сейчас, — сказал Ю. Гарфилд, — многие ученые и редакторы делают ужасную ошибку, когда думают, что SCI был создан только для того, чтобы производить его побочный продукт — JCR. Основная цель этих ресурсов — информационное обеспечение исследователей, хотя массивы библиографических записей являются неоценимым источником наукометрической информации» [Garfield, 2011]. Впервые понятие «импакт-фактор» было предложено Ю. Гарфилдом совместно с доктором И. Шером (Irving Sher) в 1963 г. [Garfield, Sher, 1963]. Введение этого термина способствовало более качественному отбору научных журналов, осуществляемому библиотеками и информационными службами. Импакт-фактор быстро стал популярен как символ научного престижа журнала, хотя его значения сильно различаются в зависимости от области знания и его релевантности предметной области.

Использование SCI явилось, по существу, инновационным подходом к изучению фор-

⁸Славинский Б. В. был начальником отдела международных связей Президиума ДВНЦ.

мальных и неформальных научных связей авторов публикаций — одной из горячих проблем исследований в США. В период своей работы над проектом Welsh Ю. Гарфилд подружился с выдающимся психологом профессором У. Гарвеем (W. Garvey), работавшим в том же университете. В конце 1950-х гг. У. Гарвей и его коллега доктор Б. Гриффит в рамках проекта American Psychology Association (Американской ассоциации психологов — самой представительной по численности профессиональной научной организации США) выполнили выдающееся исследование по изучению системы научных коммуникаций. Следует сказать, что наличие и важность роли научных коммуникаций всегда признавалось в науке, однако именно У. Гарвей и Б. Гриффит впервые сумели построить всю цепочку получения и распространения нового знания по мере выполнения и завершения исследования. Группа У. Гарвея и Б. Гриффита выявила роль неформальных каналов научных коммуникаций и «invisible colleges» («незримых коллективов»), возникающих для оперативного распространения информации и обмена знаниями. Важный и нетривиальный вывод их исследования заключался в том, что научные коммуникации являются целостной системой, в которой формальные и неформальные методы распространения информации тесно связаны между собой. Эти исследователи отмечали, что система научных коммуникаций является открытой социальной системой, основанной на взаимодействии людей, и элементы этой системы связаны между собой таким образом, что изменение одного из них влияет на функционирование других [Garvey, Griffith, 1967]. Заслуги профессора У. Гарвея в исследованиях научной коммуникации и психологии научного творчества были от-

мечены в 1992 г. золотой медалью American Psychology Association.

В середине 1960-х гг. проф. Б. Гриффит возглавил кафедру в Дрексельском университете (Drexel University), и его интерес сместился к наукометрическим исследованиям. Изучая проблемы старения научной литературы, он увидел в этом процесс, в котором идеи или литература, казалось бы, изнашиваются или становятся менее полезными, отражая скорость роста и значимость знания. Он был ярким сторонником и пропагандистом использования метода цитирования для анализа формальной структуры науки, параллельной неформальной социальной структуре, которую он изучал вместе с профессором У. Гарвеем. Ю. Гарфилд отмечал важность вклада Б. Гриффита в 70-х гг. в разработку исследований по картированию науки. Именно он первым применил технику многовекторного шкалирования для построения сети когнитивных и социальных связей и научил этому других. В то время Г. Смолл (Henry Small — директор аналитического отдела ISI) занимался разработкой метода ко-цитирования. Б. Гриффит и Г. Смолл в 70-х гг. применили этот изысканный метод для создания карты всех основных, наиболее активно развивающихся специальностей в естественных науках [Garfield, 1969]. Хотя метод ко-цитирования был построен для создания связей между документами, он распространил эту идею для создания сетей ко-цитирования между авторами. Выступая на церемонии по поводу вручения профессору Б. Гриффиту в 1997 г. золотой медали имени Д. Прайса, профессор Г. Смолл подчеркнул, что «...несмотря на фантастические успехи в развитии технологии, принципы, стимулирующие научные коммуникации, остаются неизменными, и это исследование будет

служить моделью и основой для подобных исследований в электронном веке» [Small, 1997]. Исследование системы научных коммуникаций рассматривалось как одна из проблем науковедения, и работы этой группы были переведены и изданы Институтом истории естествознания и техники (ИИЕТ) в 1976 г. [Коммуникация..., 1976]. В предисловии к первому тому монографической серии Essays of an Information Scientist Ю. Гарфилда нобелевский лауреат Дж. Ледерберг писал: «...создание SCI — это лучший пример подхода Джина (Gene) к коммуникациям в науке: это не некоторый поддерживающий (вспомогательный) инструмент или смазочный материал, это важная часть самой науки, такая же, как работа за столом в лаборатории или с калькулятором» [Lederberg, 1977].

Профессор Гриффит вырастил плеяду талантливых учеников, двое из которых были удостоены медали Д. Прайса. Узнав о присуждении Б. Гриффиту золотой медали Прайса, выдающийся социолог науки Р. Мертон сказал: «Зная огромное восхищение Дерека преданностью Белвера науке и его вкладом в нее, я думаю, что он бы аплодировал этому решению. Если и есть наследник мантии Дерека в наукометрии, то это Белвер» [Small, 1997]. В память о профессоре Б. Гриффите как ученом и педагоге его вдова К. Адамс и доктор Ю. Гарфилд учредили для аспирантов Дрексельского университета стипендию имени В. Гриффита. Ученый с энциклопедическим образованием, друг и соратник профессоров Д. Прайса, У. Гарвея, Р. Мертона, заядлый яхтсмен, изучающий российскую историю до последних дней жизни, он имел открытый дом в Филадельфии для российских и европейских коллег.

Сразу после выпуска SCI Ю. Гарфилд проводит огромную работу по пропаганде ис-

пользования статистики SCI для изучения различных аспектов истории науки, научных открытий и темпов развития направлений науки. В 1970 г. Ю. Гарфилд снова пишет об использовании индекса для изучения истории науки «Индексирование ссылок и оценка научных статей». Он отмечает, что те, кто внес огромный вклад в развитие научных направлений, могут быть идентифицированы при анализе цитируемости их работ специалистами в данной области [Garfield, 1969]. Ю. Гарфилд обращает внимание научного сообщества на различие показателей цитируемости в зависимости от областей знания и призывает к осторожности в использовании примитивного подсчета ссылок при оценке исследователей. Он отмечает тот факт, что мониторинг цитируемости и его всплеск могут привлечь внимание к статьям, которые внезапно были открыты или переоткрыты через несколько лет. Ю. Гарфилд писал, что можно привести множество примеров важных открытий, которые имели малое воздействие на современные исследования. Широко известно, например, что значение работы Менделя не было замечено свыше 30 лет. Нобелевская премия по медицине была присуждена П. Роусу (Peyton Rous) в 1966 г. за исследование вируса рака, получившего название саркомы Роуса. Однако этот вирус он открыл в 1910 г.(!), и только после того, как вирус лейкемии был изолирован в 1951 г., работа Роуса была по достоинству оценена. Известный американской социолог науки профессор Г. Цукерман (Harriet Zuckerman, Columbia University, NY) указывала, что отсроченное признание случается во многих областях знания, включая работы, за которые впоследствии была присуждена Нобелевская премия. Историки и социологи могут изучать научное сопротивление на

примерах отсроченного признания статей, выявленных анализом сетей цитирования [Garfield, 1981]. В 1965 г. Ю. Гарфилд опубликовал исследование, в котором показал, что существует корреляция между высокой цитируемостью ученого и присуждением Нобелевской премии. В списке опубликованных им высокоцитируемых работ были как нобелевские лауреаты, так и те, кто получил эту премию позднее. Корреляция — это не причинно-следственная связь. По справедливому замечанию американского документалиста Б. Кронина (B. Cronin), «цитирование — это индивидуальный процесс, а результат этого процесса имеет огромное значение в сфере науки» [Cronin, 1981]. Будучи индивидуальным решением, указание на научный результат (публикацию) другого ученого выражает особенности мотивации того, кто на него ссылается. Премия — это результат выдающегося исследования, оцененного экспертами Нобелевского комитета. По данным Ю. Гарфилда, в среднем в 70-х гг. один современный ученый цитировался около восьми раз. Цитируемость нобелевских лауреатов в среднем в то время составляла около 150 раз в год. Его исследования показали, что цитируемость после присуждения Нобелевской премии резко возрастает. При этом он отмечает, что частота цитирования — это далеко не однозначный показатель применительно как к научному направлению, так и к отдельному ученому.

Изучение мотивов цитирования было предпринято рядом западных ученых, обративших внимание на различные виды этих мотивов. Ссылка фиксирует круг общения ученого. Но он может быть и оппонентным кругом, т. е. включать исследователей, с которыми автор полемизирует, подвергает критике их идеи и факты, противопо-

ставляя им собственные. Такая полемика также может оказать влияние на модель цитирования, притом не всегда в открытой форме. Однако небольшие отклонения в общепринятых нормах цитирования если случаются, то редко, поскольку этому препятствуют высокие этические стандарты научного сообщества [Маркусова, Ярошевский, 1987]. Сам Ю. Гарфилд неустанно призывает к осторожности в использовании данных цитирования, отмечая, что SCI «не счетчик ссылок» и поэтому как «любой инструмент — от ядерной энергии до молотка — должен быть правильно использован» [Garfield, 1985]. Позднее он писал: «...мы не хотели быть слишком агрессивными в его (SCI) другом применении, однако еще в 1965 г. мы с Шером писали о его использовании для оценки эффективности исследований, о новых факторах оценки научной литературы на основе анализа цитируемости» [Garfield, 1969].

Бесспорно, что отцом наукометрии является доктор Юджин Гарфилд, однако сам термин был предложен нашим соотечественником, выдающимся ученым, статистиком и философом профессором МГУ В. В. Налимовым в его ставшей бестселлером книге «Наукометрия», опубликованной в 1969 г. [Налимов, Мульченко, 1969]. В интервью⁹ Ю. Гарфилд вспомнил: «...профессор В. В. Налимов — блестящий статистик, выпустил книгу “Наукометрия”, которая была переведена каким-то правительственным агентством (может быть, ЦРУ) без указания названия организации, осуществившей это издание. Книга распространялась бесплатно, как технический отчет. Таким образом она попала ко мне в руки. В Венгрии с книгой

⁹Материал размещен на сайте Web of Stories (URL: <http://www.webofstories.com/play/eugene.garfield/57>).

ознакомился профессор Тибор Браун (T. Braun, Венгрия), который выбрал термин, предложенный В. В. Налимовым, и дал его в качестве названия нового журнала *Scientometrics*. По мнению Ю. Гарфилда, термин «*Scientometrics*» звучал по-английски значительно лучше, чем «науковедение» — «*Science of science*», и отражал сущность предмета — измерение науки количественными методами. Актуальность задач, поставленных В. В. Налимовым, начинает осознаваться только сейчас. Тому свидетельство, что в 2001 г. был опубликован специальный выпуск журнала *Scientometrics*, посвященный вкладу В. В. Налимова в становление и развитие наукометрии как научной дисциплины. Отмечая выдающийся вклад В. В. Налимова, следует рассказать и об истории развития этого направления в России.

История наукометрии в мире и в России неразрывно связана с историей науковедения. По мнению американского историка науки профессора Л. Грэхама (L. R. Graham) [Graham, 1998] и профессора П. Вoutersа (P. Wouters, Нидерланды) [Wouters, 1999], в начале XX в. русские были впереди Запада по проблемам истории науки. Выдающийся геохимик и философ профессор В. И. Вернадский начал исследования по этой тематике еще в 1893 г.! В 1902 г. в Московском государственном университете он прочитал первую лекцию «Взгляд на историю современного мира науки». После 1917 г. он вел активную кампанию по созданию института по истории науки и техники. Результатом явилось появление первого в мире института под названием «Комиссия академии наук по истории знаний». Как отмечает профессор Грэхам, «взгляды Вернадского на историю и социологию науки опережали его время, хотя современные историки могут

сказать, что он придавал слишком большое значение роли гениальных идей и недооценивал важность социального контекста и техники» [Graham, 1998]. Появилось в России понятие «науковедение», введенное И. Боричевским, и это направление привлекло внимание главного ученого секретаря АН СССР академика С. Ф. Ольденбурга. В 1929 г. академик В. И. Вернадский был отстранен от должности главы этой комиссии и заменен Н. И. Бухариным [Гиндилис, 2011].

В 1931 г. эта комиссия была преобразована в Институт истории науки и техники, и его директором стал Н. И. Бухарин. Возглавляемая им большая советская делегация приняла участие во Втором международном конгрессе по истории науки, проходившем в Лондоне в 1931 г. Западные исследователи отмечали [Wouters, 1999] то колоссальное впечатление, которое произвел на участников, в том числе на Дж. Бернала и Р. Мертон, доклад советского физика Б. Гессена о социальных корнях «Начал Ньютона». Однако в 1936 г. Б. Гессен и шесть других членов советской делегации были арестованы и погибли [Ярошевский, 1985]. В. В. Налимов провел 18 лет в ГУЛАГе и был освобожден в 1953 г. Он стал работать референтом в реферативном журнале ВИНТИ АН СССР.

В 1962 г. директором ИИЕТ АН СССР стал академик Б. М. Кедров. Книги Дж. Бернала и Д. Прайса привлекли большое внимание советского научного сообщества, и возник значительный интерес к проблемам науковедения. Активным пропагандистом науковедения стал профессор С. Р. Микулинский¹⁰, бывший в то время заместителем директора ИИЕТ. В июне 1966 г. во Львове-Ужгороде состоялся со-

¹⁰Микулинский С. Р. (1919–1991), член-корреспондент АН СССР, директор ИИЕТ с 1974 по 1987 г.

ветско-польский симпозиум по проблемам комплексного изучения науки, на котором была и наукометрическая секция. На этом симпозиуме впервые В. В. Налимов использовал термин «наукометрия» как термин, «характеризующий дисциплину, занимающуюся использованием количественных методов для изучения процесса развития науки». Нужно отметить, что еще в 1959 г. в журнале «Успехи физических наук» вышла статья В. В. Налимова «Научная и техническая информация как одна из задач кибернетики» [Влэдуч и др., 1959], посвященная химической кибернетике и наукометрии. В то время В. В. Налимов был внештатным сотрудником ВИНИТИ. Позднее он вспоминал: «Я познакомился с наукометрией совершенно случайно. В конце 50-х гг. я работал редактором реферативного журнала по физике в ВИНИТИ. Я работал в ВИНИТИ, поскольку я мог переводить с трех европейских языков. Однажды я получил статью Д. Прайса, и мне сказали, что я единственный специалист, который может ее перевести. Она была на итальянском языке. Мне понравилась эта статья. Она была посвящена экспоненциальному росту науки. Вместе с Г. Влэдучем и Стяжким я написал статью о связи между кибернетикой и информатикой. Она была опубликована в журнале “Успехи физических наук”. Директор института А. И. Михайлов был первым, кто прореагировал на нее: он пригласил моих соавторов и отругал. Обращение к кибернетике – вот что инкриминировалось двум моим коллегам. В то время это было против официальной идеологии» [Wouters, 1999].

Следует отметить, что в СССР в начале 60-х гг. образовались две школы по наукометрии. Одна, московская школа, не была организационно оформлена и явля-

лась «незримым коллективом». Инициатором и душой этой школы был профессор В. В. Налимов. С 1966 г. он начал проводить неформальный семинар по использованию метода цитирования в ИИЕТ АН СССР при активном участии и поддержке профессора С. Р. Микулинского. Постоянными участниками этого семинара и активными помощниками в исследованиях В. В. Налимова по наукометрии были И. М. Ориент — главный редактор одного из лучших отечественных журналов того времени «Заводская лаборатория», доцент Ю. В. Грановский (МГУ), З. Барина и др.

В интервью, данном в 1992 г., И. М. Ориент вспоминала: «...мы начали работать в 1966 г. Центром притяжения был Василий Васильевич Налимов. Тогда он получил от Ю. Гарфилда выпуск SCI за 1965 г. и пригласил людей, чтобы учиться работать с SCI (автора этой главы позднее учила И. М. Ориент). Сначала мы не отнеслись к этому серьезно, но позднее стали проводить много времени, работая с SCI. Мы “распространяли” SCI среди друг друга. Я выбрала аналитическую химию, Васильев изучал химическую физику, Ю. В. Грановский — неорганическую химию. Каждый из нас имел свое собственное направление» [Wouters, 1999].

В 1966 г. В. В. Налимов с группой соавторов опубликовал статью «Изучение научных журналов как каналов связи. Оценка вклада отдельных стран в мировой научный информационный поток» в журнале «Научно-техническая информация» [Оценка..., 1968]. В 1969 г. В. В. Налимов в соавторстве с З. М. Мульченко выпускает книгу «Наукометрия. Изучение развития науки как информационного процесса» [Налимов, Мульченко, 1969]. В этой книге наукометрия рассматривалась как раздел науковедения.

В книге было дано следующее определение этой дисциплине: «Будем называть наукометрией количественные методы изучения развития науки как информационного процесса» [Налимов, Мультченко, 1969]. Обсуждался следующий круг вопросов: информационная модель развития науки, рост информационных потоков, цитируемость литературы, изучение внутренних связей в науке по языку библиографических ссылок, «незримые коллективы», оценка вклада, вносимого отдельными странами в мировой научный информационный поток, процесс статистического анализа одного из научных направлений и логический анализ проблемы прогнозирования.

Опубликованная 45 лет назад, эта книга включала все актуальные проблемы исследований современного библиометрического сообщества. В 2001 г. международный журнал *Scientometrics* посвятил В. В. Налимову специальный номер, открывая который главный редактор журнала профессор Тибор Браун отметил, что В. В. Налимова вместе с Дерекком де Солла Прайсом можно считать основателями наукометрии. В 2010 г. в МГУ была проведена международная научная конференция, посвященная 100-летию со дня его рождения. Выпущен сборник трудов конференции, издана книга В. В. Налимова «Облик науки», которая впервые была опубликована в ISI Press Ю. Гарфилда в 1981 г. [Грановский, 2010]. В настоящее время его вдова и верная помощница философ Ж. Дрогалина готовит к переизданию книгу В. В. Налимова «Разбрасываю мысли», впервые опубликованную после его смерти.

Вторая школа, которая затем превратилась в специальный исследовательский центр под руководством Г. М. Доброва, действовала в Киеве, на Украине. Г. М. Добров

был очень энергичным человеком, хорошо вписавшимся в советский истеблишмент. Он был первым секретарем райкома комсомола в Киеве в конце 1950-х гг. Инженер по образованию, он окончил аспирантуру Института теплоэнергетики АН УССР (в настоящее время Институт технической теплофизики НАН Украины), защитив диссертацию по истории угольных комбайнов. Позднее в соавторстве он опубликовал двухтомную монографию по истории угольной промышленности.

В начале 60-х гг. его интересы сместились к проблемам кибернетики и теории управления. Монография Г. М. Доброва «Наука о науке», опубликованная в 1966 г. [Добров, 1966], стала значительной вехой в истории отечественного науковедения и наукометрии. В эту книгу был включен перевод книги Д. Прайса «Малая наука, большая наука». Эта книга Г. М. Доброва пользовалась огромной популярностью, дважды была переиздана (1970 и 1989 гг.). Ее опубликование ознаменовало начало развития науковедческих и наукометрических исследований на Украине. В это время Г. М. Добров уже руководит сектором в Институте кибернетики АН УССР. Возглавляемый им коллектив проводит исследования по оценке научного потенциала, прогнозированию науки. Проблемы управления наукой детально рассмотрены в монографиях «Науковедение как наука об эффективном управлении научной деятельностью: опыт, проблемы, перспективы», «Организация и эффективность науки» и в написанной в соавторстве с В. М. Глушковым и В. И. Терещенко монографии «Беседы об управлении». В 1969 г. Г. М. Добров начал выпуск периодического научного сборника «Науковедение и информатика» и был его ответственным редактором.

С 1979 г. Г. М. Добров был членом редакционной коллегии международного журнала *Scientometrics*, издаваемого профессором Т. Брауном в Будапеште. По приглашению Г. М. Доброва на Украину приезжали профессор Д. Прайс, Ю. Гарфилд. Активная научная деятельность и неутомимая энергия Г. М. Доброва привели к созданию в 1986 г. Центра исследований научно-технического потенциала и истории науки, который он возглавлял до конца жизни (1989 г.). Теперь этот центр носит его имя.

Два вопроса, о которых хотелось бы упомянуть, рассказывая об истории развития наукометрии в СССР. В СССР была проделана значительная работа по повышению квалификации библиотечных и информационных специалистов. Так, в 1966 г. в ВИНТИ были организованы специальные курсы для специалистов с отрывом от работы, с 3-месячным и 6-месячным обучением с шестичасовыми занятиями три раза в неделю. В 1971 г. эти курсы были преобразованы Постановлением Совета Министров СССР в Институт повышения квалификации информационных и библиотечных работников (ИПКИР) Министерства по науке и технике СССР. Несколько сотен тысяч библиотечных и информационных специалистов прослушали и сдали экзамен по циклу лекций об информационных продуктах ISI: *Current Contents*, *SCI* и *JCR*. Автор читала этот курс лекций с 1975 по 1992 г. Затем ИПКИР начал работать на коммерческой основе. В период с 2006 по 2011 г., до поглощения этого института Финансовой академией при Правительстве РФ (2012 г.), автор этой главы читала курс по информационным продуктам Thomson Reuters для специальной группы университетских специалистов. С 1967 г. (по настоящее время) в программу для сдачи экза-

мена по специальности «научно-техническая информация» для получения степени к. т. н. в ВИНТИ включен раздел о работах Д. Прайса, В. В. Налимова и Ю. Гарфилда. Наукометрии или библиометрии как специальности в номенклатуре ВАКа не существует до сих пор, и защиты диссертации по этой специальности относятся к библиотечным наукам с присуждением степени по педагогическим наукам.

По свидетельству Ю. В. Грановского [Грановский, 2010], В. В. Налимов уделял особое внимание процессу обучения специалистов, считая это необходимым для выполнения наукометрических исследований.

Можно констатировать, что постановление от 1 ноября 2013 г. Председателя Правительства РФ Д. Медведева № 979 «О внесении изменений в постановление Правительства Российской Федерации от 8 апреля 2009 г. № 312»¹¹, устанавливающее правила оценки результативности деятельности научных организаций, включающее ряд библиометрических показателей, будет способствовать растущей потребности в подготовке квалифицированных специалистов по этой тематике. Следует отметить, что пока работу по ознакомлению широкой научной общественности страны по работе и возможностям библиометрических информационных ресурсов выполняет группа высококвалифицированных специалистов компании Thomson Reuters.

Идея создания Указателя цитируемой литературы была очень популярна в СССР. В ВИНТИ АН СССР профессор Р. Гиляревский неоднократно предпринимал попытки создания экспериментального массива для будущего Индекса цитирования

¹¹Материал взят с сайта Российской академии наук (URL: <http://www.ras.ru/news/shownews.aspx?id=613a30f8-1475-4d9a-a6a3-75df1501be7a>).

отечественной литературы, но по не зависящим от него причинам эта работа не получила развития. Часть этого материала была использована В. А. Маркусовой для диссертационного исследования. В 1973 г. под руководством директора Азербайджанского института научно-технической информации и технико-экономических исследований Госплана Азербайджанской ССР А. Мехтиева были выпущены указатели цитируемой литературы (УЦЛ): УЦЛ-нефть, УЦЛ-экономика. Затем в 1974 г. был выпущен УЦЛ-информатика на основе ссылок статей, содержащихся в обеих сериях журнала «Научно-техническая информация» за период с 1964 по 1973 г. В массиве 1960 опубликованных статей содержалось 9535 ссылок на отечественные и зарубежные публикации [Маркусова, 1976]. К сожалению, эти работы не получили дальнейшего развития.

Однако успешная попытка создания указателя цитирования в России была осуществлена в 2001–2009 гг. Инициативная группа (к. т. н. А. Н. Либкинд, д. ф.-м. н. В. А. Минин и их младший коллега И. А. Либкинд) в 2001 г. предложила создать такой указатель на основе данных о публикациях участников проектов Российского фонда фундаментальных исследований (РФФИ). Тогдашний председатель РФФИ академик М. В. Алфимов поддержал эту идею. К ее реализации подключился еще ряд специалистов. Пилотная версия Указателя РФФИ была создана и продемонстрирована Ю. Гарфилду во время его приезда в Москву на конференцию, посвященную 50-летию ВИНТИ. Ю. Гарфилд с большим интересом ознакомился с Указателем РФФИ и отметил в нем ряд оригинальных решений и полезных опций, которые в то время отсутствовали в SCI. Однако по ряду причин финан-

сирование проекта прекращено. Но идея создания Российского указателя оставалась популярной.

Можно только поражаться огромной интуиции Ю. Гарфилда, идеи которого позднее были воплощены в новые информационные продукты. В 1973 г. Ю. Гарфилд публикует письмо к редактору журнала *Nature*, в котором обращает внимание на использование индекса как инструмента для предсказания будущих нобелевских лауреатов. «Анализ самых высокочитируемых авторов позволил в 1967 г. выявить двух нобелевских лауреатов среди списка наиболее цитируемых авторов. Теперь благодаря накопленным массивам статистики за 1961–1971 гг. (кроме массивов за 1962 и 1963 гг.) появилась возможность составить список высокочитируемых авторов. Из более чем 1,8 млн процитированных авторов только 42 тыс. были процитированы более 30 раз в каждом из обследованных лет. Однако только 2100 авторов были процитированы более тысячи раз. В этом списке оказались все нобелевские лауреаты за 1972 г. И также те (за небольшим исключением), кто получил эту премию в предшествующие годы» [Garfield, 1973]. Таким образом, заключает Ю. Гарфилд, нобелевские лауреаты принадлежат к элитной группе, состоящей из 0,1% всех цитируемых авторов.

5 декабря 1977 г. Ю. Гарфилд опубликовал статью, посвященную 250 наиболее высокочитируемым авторам за 1961–1975 гг. Авторы этих работ приглашались давать комментарии, почему, по их мнению, эти работы привлекли такое внимание. Этот раздел назывался «Citation Classics» (классика цитирования). Первый комментарий был написан автором самой высокочитируемой в мире статьи Оливером Х. Лоури (Oliver H. Lowry), опубликованной в 1951 г. Эта статья была

процитирована к 1976 г. более 58 тыс. раз¹². Этот раздел публиковался вплоть до 1992 г. Всего за этот период времени было опубликовано около 3000 комментариев авторов этих работ, из которых 18 принадлежали советским ученым. Только пять работ были опубликованы в советских научных журналах, остальные работы либо были опубликованы за рубежом, либо это было издание на английском языке монографий, первоначально опубликованных в СССР.

Отметим, что статья молодого советского физика А. А. Абрикосова¹³, опубликованная в 1957 г. в «Журнале теоретической и экспериментальной физики», попала в список самых цитируемых советских работ, опубликованных в советских журналах, за период с 1961 по 1972 г. и получила 730 ссылок [Garfield, 1975]. За эту работу в 2003 г. А. А. Абрикосов стал лауреатом Нобелевской премии по физике. Сведения о книгах попадали в SCI только в том случае, если книга была процитирована в журналах-источниках. Например, опубликованная академиком Л. М. Бrehовских в 1966 г. книга «Динамика жидких сред» стала бестселлером и чемпионом в вышеупомянутом разделе Citation Classics благодаря ее высоким показателям цитируемости в статьях, опубликованных в журналах-источниках. Позднее, в 1993 г., был выпущен том Citation Classics, состоящий из этих комментариев.

С накоплением массивов информации в базе данных (БД) SCI были выяснены временные характеристики в системе научных

коммуникаций. В конце 60-х было установлено, что средний возраст цитируемой статьи находится в пределах от пяти до пятнадцати лет, в зависимости от области знания. В то же время в области молекулярной биологии средний возраст около 25 % цитируемых публикаций составлял два года, другие же были намного старше. Эти данные позволили обратить внимание на очень небольшую группу статей, которая цитируется уже через несколько месяцев после опубликования. В 1975 г. впервые в нескольких выпусках библиографического указателя Current Contents были напечатаны списки ста наиболее цитируемых публикаций, опубликованных в этом же году. Два года спустя ISI приступил к периодической публикации списка наиболее цитируемых статей, так называемых «горячих публикаций» (hot papers). На основе выявленных фактов была реализована новая информационная услуга — публикация списков наиболее цитируемых статей текущего года. Эти списки «hot papers» обновляются каждые два месяца и доступны пользователям в настоящее время в Интернете.

Дальнейшее развитие наукометрии связано с разработками в аналитическом отделе ISI, который с начала 70-х гг. в течение почти 30 лет возглавлял Генри Смолл. В 1973 г. он разработал в ISI и осуществил практическое применение метода создания фронтов научных исследований или кластеров на основе ко-цитирования [Small, 1973]. Сила ко-цитирования может определяться количеством совместных ссылок на данную пару документов (публикаций) в других документах. Если предположить, что высокоцитируемые документы отражают важные понятия в какой-либо отрасли науки, то ассоциации ко-цитируемых документов можно рассматривать как классы или кластеры связанных

¹²Эта статья Лоури была процитирована более 300 тыс. раз.

¹³Академик А. А. Абрикосов, лауреат Нобелевской премии за основополагающие работы по теории сверхпроводников и сверхтекучих жидкостей. Разделил эту премию с академиком В. Л. Гинзбургом и англо-американским ученым лордом Anthony James Leggett в 2003 г.

понятий. Эти кластеры отражают основные направления современных исследований, в совокупности составляющие карты науки. Необходимо отметить, что одновременно с Г. Смоллом и независимо от него использование критерия ко-цитирования на небольшом массиве статей по лазерам было предложено И. В. Маршаковой (1973 г.) в ВИНТИ [Маршакова, 1973]. Разработка методов экспликации структуры современной науки посредством построения, анализа и оценки сетей цитирования, включая использование критерия ко-цитирования, позволила ISI начать в 1987 г. выпуск новой серии обзорно-справочных изданий, которые были названы Atlas of Science (Атласами науки) [Garfield, 1989]. Этот метод используется для построения кластеров направлений науки, входящих в состав БД Essential Science Indicators — ESI (Основные показатели науки и техники).

С конца 70-х гг. создание различных карт тематических направлений науки, а также кластеров авторов нашло широкое применение в разных странах мира. Создание таких карт науки позволило дать объективную и прозрачную оценку сильным и слабым сторонам научной деятельности, производить мониторинг развития научных направлений и науки в целом.

Уникальная по своим масштабам библиометрическая статистика SCI, SSCI, Arts&HCI, ESI и другие издания фирмы ISI стали богатейшим источником данных для наукометрических исследований, что послужило мощным толчком для их расширения и развития. С 1979 г. начал выходить двухмесячный международный журнал *Scientometrics*, который посвящен рассмотрению всех количественных аспектов науки о науке, коммуникации в науке и научной политики. В 1993 г. было организовано международное

общество по наукометрии и информетрии *International Society on Scientometrics and Informetrics*, конференции которого проходят раз в два года. В 2000 г. была создана международная организация научного сотрудничества в наукометрии, информетрии и вебометрии COLLNET, конференции которой проходят ежегодно. С 2007 г. общество COLLNET выпускает журнал *Collnet Journal of Scientometrics and Information Management*, публикуемый два раза в год.

Одним из важных моментов в становлении и развитии библиометрии явилось понимание того, что библиометрическая статистика может быть использована как показатель результативности научных исследований и эффективного использования денег налогоплательщиков.

Впервые статистические сведения о количестве публикаций — научной продуктивности и их цитируемости были использованы для оценки развития науки в США и в мире в отчете NSF США, выпущенном в 1972 г. под названием *Science Indicators*¹⁴ (показатели науки), который теперь называется *Science & Engineering Indicators*. С тех пор на протяжении более 40 лет пятая глава этих отчетов NSF используется для мониторинга тенденций развития мировой науки и ее прогнозирования. С 2008 г. в этом отчете публикуются данные о количестве высокоцитируемых работ и их доле в научной продуктивности каждой из обследуемых стран и в мировом потоке¹⁵.

До начала 80-х гг. основным способом оценки значимости намечаемых научных программ служила система рецензирования ведущими специалистами — мэтрами

¹⁴Science Indicators — 1972. NSB-73-1 // Washington, DC: U.S. Government Printing office, 1973.

¹⁵Материал взят с сайта National Science Foundation (www.nsf.gov).

науки. Выполненный в США по заказу NSF в конце 80-х гг. анализ показал, что при распределении грантов выбор грантополучателей определяется наполовину строго обоснованными рекомендациями, а наполовину — рекомендациями, высказанными наугад. В другом исследовании, выполненном по заказу той же организации, смешанную группу рецензентов просили ответить на вопросы, какой из двух заявок на грант они отдадут предпочтение: той, которая поступила из института, пользующегося мировой известностью, и основана на общепринятых представлениях, или, напротив, заявке из малоизвестного научного учреждения и содержащей радикальные идеи. Большинство респондентов предпочло первый вариант [Маркусова, 1998].

Этот пример показывает, что возникла потребность в данных, которые не могут быть получены от высококвалифицированных экспертов. Именно в этих случаях приходится обращаться к показателям развития науки. Такие показатели не служат заменой экспертизы, однако являются мощным информационным инструментом поддержки. Этот инструмент необходим при определении и мониторинге тенденций развития конкретной научной дисциплины на уровне страны или организации в сравнении с развитием этой научной дисциплины в мире (в сопоставлении с другими странами), оценки объема и других характеристик международного сотрудничества, роли развивающихся стран, влияния фундаментальных и прикладных исследований на развитие наукоемких технологий, структуры научных дисциплин и их взаимоотношений с другими областями знаний.

Рост конкурентоспособности научных коллективов и борьба за дополнительные инвестиции в науку привели к развитию

сильных наукометрических коллективов в США и Европе. В Великобритании это был коллектив под руководством доктора С. Катца (Sylvan Katz) и Б. Мартина (B. Martin) в Университете графства Сассекс, а в Нидерландах — под руководством профессора А. Ван Раана (Antony Van Raan) в Университете Лейдена. Коллектив А. Ван Раана начал работать с 1980 г. в Лейденском университете. В 1989 г. на базе этого коллектива был создан Centre for Study Science and Technology — CWTS (Центр по анализу науки и техники). За 25 лет своего существования этот центр превратился в учреждение мирового уровня и является лидером научных исследований по разработке и применению библиометрических индикаторов науки для оценки эффективности научной деятельности. Огромную поддержку этому коллективу в период его становления оказал один из видных специалистов по научной политике, директор Национального фонда по технологиям, физик по образованию доктор Кейс Ле Пеир (C. Le Pair), возглавлявший в начале 1970-х гг. наукометрическую группу в Национальном научном фонде по физике в Нидерландах. В настоящее время CWTS руководит профессор П. Боутерс. Сильные наукометрические коллективы работают в Бельгии под руководством профессора В. Глэнцела (W. Glänzel, Университет Каролинка) и профессора Л. Лейдесдорфа (L. Leidesdorf, Amsterdam University), в Центре по изучению науки и техники (Observatoire des Sciences et des Techniques — Франция) под руководством доктора М. Зитта (M. Zitt), коллектив под руководством доктора Л. Батлер (Linda Butler, Australian National University) в Австралии. Существуют многочисленные библиометрические коллективы в Индии, Китае, Бразилии и Испании.

Активные исследования в наукометрии с начала 70-х гг. проводились и продолжают вестись в Библиотеке АН Венгрии под руководством профессора Т. Брауна. Т. Браун — признанный специалист в области радиоаналитической химии, работавший в Университете Етвос (Будапешт) и одновременно создавший и возглавлявший группу по наукометрии в Библиотеке АН Венгрии. Эта группа выполнила и опубликовала ряд интересных исследований по научной продуктивности и влиятельности (импакту) различных стран мира. Члены этого коллектива, тогда совсем молодые исследователи, В. Глэнцел и А. Шуберт (А. Schubert) стали признанными специалистами в мировом наукометрическом сообществе. Профессор Т. Браун стал одним из первых, награжденных медалью Д. Прайса (1986 г.). Позднее эту премию получили профессора А. Шуберт и В. Глэнцел. Именно Т. Браун начал выпускать в 1979 г. *Scientometrics* и был его главным редактором более 30 лет. Теперь эту эстафету принял В. Глэнцел. В Чехословакии очень интересные исследования по наукометрии проводились специалистом в области ядерной физики доктором Я. Влахи, работавшим в различных институтах чехословацкой АН. Я. Влахи опубликовал четыреста работ по количественным аспектам исследований в области физики. Он практически работал и писал один. Только две его статьи были опубликованы при сотрудничестве с доктором Х. Ю. Червоном из Академии наук ГДР. В 1987 г. Я. Влахи был награжден медалью Д. Прайса. Позднее его интересы сместились к истории чешского сопротивления нацистам во время Второй мировой войны. Отец Я. Влахи, активный участник сопротивления, был пойман и расстрелян в тюрьме в Берлине

в 1943 г. Я. Влахи скоропостижно скончался в Берлине в 2012 г.

В настоящее время в России существует довольно большое библиометрическое сообщество, проводящее исследования в разных организациях: в МГУ им. М. В. Ломоносова (доцент Ю. В. Грановский), в ВИНИТИ РАН (к. т. н. А. Н. Либкинд, д. пед. н. В. А. Маркусова), в Институте философии РАН (д. ф. н. И. В. Маршакова), в Институте биологии моря ДВО (профессор А. И. Пудовкин), в Высшей школе экономики (к. ф.-м. н. В. В. Писляков, И. А. Стерлигов и др.), в Институте проблем развития науки РАН (профессор Л. Э. Миндели), в ИИЕТ (к. псих. н. А. А. Алахвердян), в Санкт-Петербургском университете (О. В. Москалева), в Уральском федеральном университете (Центр мониторинга науки и образования, М. А. Акоев), в СО РАН (к. т. н. Н. А. Мазов). Огромную роль в изучении мотиваций цитирования сыграл выдающийся советский психолог профессор М. Г. Ярошевский, возглавлявший отдел психологии научного творчества с середины 60-х до середины 90-х гг. прошлого века в ИИЕТ РАН. В октябре 2013 г. в ИПРАН при участии Thomson Reuters была проведена представительная международная конференция «Проблемы наукометрии. Состояние и перспективы развития», на которой были представлены доклады отечественных специалистов и их коллег из Белоруссии, Великобритании, Китая, Нидерландов и Франции. Начиная с 1991 г. в ИНИОН РАН проводится ежемесячно научный семинар по науковедению и наукометрии под руководством известного философа и историка науки профессора А. И. Ракитова.

Важным этапом в развитии наукометрии как дисциплины, позволяющей изучать эффективность научных исследований,

следует считать первый международный симпозиум по оценке научных исследований, состоявшийся в 1987 г. в Лондоне. Кроме ведущих в мире специалистов в области наукометрии Ю. Гарфилда, Г. Смолла, А. Ван Раана и др., в работе принимал участие и профессор А. Вайнберг, один из крупнейших в США администраторов науки, который одним из первых осознал важность использования наукометрических данных для более эффективного распределения инвестиций в науку. На этом симпозиуме внимание было сосредоточено на двух важнейших этапах оценки научных исследований и инвестиций в них: уже выполненных и тех, чье финансирование находилось в процессе обсуждения, таких как «Геном человека» или «Создание коллаидера» [The Evaluation..., 1989]. Как мы знаем, эти международные проекты были осуществлены.

Мониторинг библиометрических показателей широко используется NSF США, Organization for Economic Co-operation and Development (OECD), Европейской Комиссией, Национальным институтом по политике в области науки и техники и Министерством экономики, торговли и промышленности Японии, а также в системе оценки качества исследований университетов Research Excellence Framework в Великобритании. В 1992 г. Ю. Гарфилд пишет о возросшем спросе на использование статистики цитируемости для представителей правительства, промышленности и оценки научной деятельности университетов и публикует результаты исследования в ISI сравнительного анализа влияния (импактов) стран Большой семерки [Garfield, 1992].

Огромное внимание научной общественности в 1997 г. привлекла статья президента Королевского общества Великобри-

тании лорда Р. Мэя (R. May) о взаимосвязи научных инвестиций и научной продуктивности на основе статистики SCI за 1981–1994 гг., опубликованная в журнале Science [May, 1997].

В ходе работы Всемирной конференции по науке, проводившейся в 1999 г. в Будапеште, был поставлен вопрос о необходимости доступа стран к достоверным статистическим данным в области науки и техники, имеющим отношение к разработке политики в этом направлении. На конференции было отмечено, что «правительствам следует содействовать дальнейшему развитию и созданию национальных статистических служб, способных обеспечивать потребности в достоверных данных. Настало время для новых усилий в деле достижения всемирного прогресса в области статистики науки и техники, чтобы определить те трудности, с которыми в настоящее время сталкиваются страны при сборе и использовании таких данных, а также для определения их информационных потребностей...» [Декларация..., 1999].

Именно в этой ситуации Статистический институт ЮНЕСКО, созданный на Генеральной конференции в 1999 г., приступил к проведению фундаментального международного обзора политики в сфере информационных потребностей и работы со статистическими данными в научно-технической области. С одной стороны, задача сводилась к определению приоритетов и разработке стратегических подходов для повышения актуальности, доступности и качества статистики по науке и технике во всех странах мира, с другой — к помощи в определении новой статистической программы ЮНЕСКО в области науки и техники. В обзоре отмечалось, что «печатные работы являются основным показателем производства нового

знания». В качестве параметров измерения были выбраны сами публикации и ссылки на них [ЮНЕСКО]. Таким образом, библиометрические показатели были официально закреплены как количественные показатели развития науки. Конечно, упомянутые показатели не отменяют важности сбора и анализа таких традиционных параметров, как размеры финансирования науки, количество ученых, занимающихся исследованиями и разработками, и т. д. В зависимости от конкретной задачи могут использоваться различные показатели. В упомянутом обзоре подчеркивалось, что на протяжении последних двадцати лет правительства нуждаются во все более точных и объективных оценках исследовательской деятельности.

В 2004 г. широкое внимание специалистов по научной политике привлекла статья лорда Д. Кинга (D. King), в то время президента Королевского общества Великобритании, *Scientific Impact of Nations* [King, 2004]. Лорд Д. Кинг отмечал, что «...правительству необходима реальная оценка научного вклада национальной науки в мировую». Анализ был выполнен для правительственной службы Office Science and Technology — OST (Офис по науке и технике) Великобритании. Целью исследования было выявление эффективности национальных исследований и сопоставление науки страны с наукой других 31 стран. В статье были проанализированы затраты на исследования и разработки, научная продуктивность и цитируемость стран Большой восьмерки, 15 стран ЕС, Австралии, Австрии, Бельгии, Южной Африки, Китая, Сингапура, Южной Кореи, Тайваня, Индии, Ирана. Оказалось, что 98 % затрат на науку в мире приходилось на 31 страну, а затраты остальных 62 стран составили всего 2 %. Поскольку цитируемость резко

различается в разных областях знаний, то для сопоставительного анализа был использован агрегированный показатель цитируемости, который позволил проводить сравнения разных дисциплин. Ранг страны оценивался по доле страны в массиве высокоцитируемых статей¹⁶. Были сопоставлены затраты валового национального продукта на одного человека (*wealth intensity*) с цитируемостью. Для выявления дисциплинарной направленности научных исследований в каждой из анализируемых стран была использована классификация OST и построена карта национальных приоритетов Большой восьмерки. Оказалось, что в европейской науке важную роль играют небольшие страны, такие как Бельгия, Дания, Финляндия, Дания, Швеция и Швейцария. Все эти страны вместе с населением 53 млн имели очень высокий импакт за период с 1997 по 2001 г., и их цитируемость составила 12,7 % от цитируемости мирового потока, в то время как для Великобритании эта цифра составила примерно ту же величину — 12,8 %. При этом ВВП этих стран составил всего 6 %, значительно меньше, чем Великобритании. Автор отмечал, что перед лицом глобального терроризма и угрозы глобального потепления правительства всех стран должны обратиться к науке.

В настоящее время библиометрические показатели стали частью протокола отчетности о деятельности научных организаций, университетов, исследовательских групп и индивидуальных специалистов. К библиометрическим показателям

¹⁶Высокоцитируемые статьи составляют 1 % самых цитируемых публикаций за 11-летний период по каждой предметной области по классификации ESI. Для каждого года и каждой предметной области берется свой порог цитируемости.

телям относятся количество публикаций; их цитируемость; импакт-фактор научного журнала, в котором они были опубликованы; количество высокоцитируемых статей; индекс Хирша и различные нормализованные показатели цитируемости. Эти показатели должны использоваться с учетом количества грантов, научных премий, участия в составах редколлегий ведущих научных журналов. Именно совокупность всех этих критериев положена в оценку деятельности исследователя или научного коллектива, занимающегося фундаментальными исследованиями в США и других развитых странах.

ISI последовательно расширял спектр информационных продуктов. С 1990 г. помимо научных журналов стали также индексироваться выпуски научных конференций Conference Proceedings (CPCI-S и CPCI-SSH). Создание универсальной базы данных цитированной литературы шло в ногу с развитием информационных и коммуникационных технологий. С 1981 г. был организован доступ в режиме онлайн к БД SCI через информационную систему DIALOG. Начиная с 1988 г. указатели научного цитирования стали выходить на CD-ROM.

В 1992 г. ISI был приобретен компанией Thomson, которая в 2008-м слилась с информационным агентством Reuters и образовала крупнейшую международную информационную компанию Thomson Reuters.

Принципиально новые возможности для доступа к указателям научного цитирования и работы с ними открылись с появлением сети Интернет. В 1997 г. SCI, SSCI, AHCI, CPCI-S и CPCI-SSH, а также JCR — Science Edition, JCR — Social Science Edition были объединены в единую политематическую реферативную базу данных научных публикаций с указанием цитирования — Web

of Science Core Collection¹⁷. В 2008–2010 гг. была существенно расширена география научных журналов, индексируемых в Web of Science Core Collection. В настоящее время в базе данных индексируются приблизительно двенадцать с половиной тысяч лучших международных научных журналов по всем отраслям знания.

Новым этапом в развитии платформы Web of Science стало размещение на ней региональных указателей научного цитирования. В 2009 г. на платформе был размещен китайский указатель Chinese Science Citation Database, в 2013 г. — латиноамериканский указатель SciELO Citation Index и в 2014 г. — корейский указатель Korean Citation Index (KCI). Региональные указатели научного цитирования индексируют наиболее значимые источники научных публикаций в своих регионах.

Библиометрическая статистика такого уникального ресурса, как Web of Science, позволяет проследить быстро меняющийся ландшафт научного мира и оценить место национальных научных достижений на карте мировой науки. Построенные на пионерской идее Ю. Гарфилда информационные ресурсы компании Thomson Reuters высоко оценены мировым научным сообществом.

Важным этапом в использовании библиометрических ресурсов стало создание мировых рейтингов университетов [Маркусова, 2010]. Наиболее известными рейтингами университетов стали британский рейтинг — World University Rankings, выпускаемый журналом Times Higher Education; рейтинг Шанхайского университета — Academic Ranking of World Universities (с 2003 г.) и Лейденский (Leiden Ranking), подготавливаемый CWTS. Развитие Интернета привело к созданию

¹⁷До 2014 г. основная база данных Web of Science Core Collection именовалась просто Web of Science.

и вебометрического рейтинга Webometrics, выполняемого с 2004 г. испанской лабораторией по вебометрике (Cybermetrics). Во всех этих рейтингах, за исключением рейтинга Webometrics, библиометрические показатели (количество статей, опубликованных профессорско-преподавательским составом университета, их цитируемость, импакт-факторы) являются важной — а в Лейденском рейтинге единственной — составляющей средневзвешенной оценки.

Рост внимания к оценке эффективности научных исследований в организациях и университетах сопровождался разработкой библиометрическим сообществом нормализованных указателей, позволяющих сопоставлять различные области знания. В 2005 г. профессором Х. Хиршем (J. Hirsch) был предложен индекс Хирша (ИХ), позволяющий оценивать одновременно качественную (цитируемость) и количественную (количество опубликованных работ) деятельность исследователей. Х. Хирш полагал, что для физиков значение $ИХ = 12$ является достаточным основанием для получения пожизненной позиции ассистента в университете США, а $ИХ = 18$ может обеспечить должность пожизненного профессора. Индекс Хирша — это один из наиболее обсуждаемых показателей, поскольку два самых мощных информационных ресурса мира Web of Science и Scopus предоставляют пользователю возможность автоматически получить значение индекса Хирша. Ученый имеет индекс равный h , если h его N_p статей имеют по меньшей мере h ссылок каждая, а каждая другая из статей ($N_p - h$) имеет менее чем h ссылок. В июне 2008 г. International Mathematical Union (Международный математический союз) опубликовал обстоятельный доклад, выполненный Комитетом по количественной оценке ис-

следователей, в котором выступил против использования библиометрических показателей и особенно против ИХ. В значительной степени это выступление математиков, мало осведомленных о зависимости цитируемости от области знания, связано с низкими показателями цитируемости в математике¹⁸.

Появились новые показатели информационной значимости журналов, такие как Eigenfactor, предложенный группой специалистов из Университета Вашингтона в 2007 г. (подробнее см. раздел 2.5. «Взвешенные» индикаторы. Собственный фактор (Eigenfactor) и индекс влияния статьи (Article Influence)), и показатель SJR, разработанный известной испанской группой Scimago Lab, работающей по контракту с компанией Эльзевир (Elsevier) на основе статистики БД Scopus.

Развитие сервисных услуг для пользователей WoS привело к созданию уникальных опций по анализу отдельных массивов информации, таких как Analyze Results и Citation Reports. В августе 2008 г. появилась новая опция, позволяющая оценивать деятельность научных фондов и агентств, финансирующих научные исследования на конкурсной основе, — Funding Agency. В 2009 г. появился новый аналитический инструмент — InCites, позволяющий проводить сопоставительный анализ стран, организаций и индивидуальных исследователей, работающих в разных областях знания. Для этой цели были разработаны специальные агрегированные показатели. В сентябре 2014 г. запущено новое поколение Web of Science

¹⁸Материал взят с сайта International Mathematical Union (IMU) (URL: <http://www.mathunion.org/publications/report/citationstatistics0/>).

и InCites, включающее также библиометрическую статистику ESI и JCR¹⁹.

Использование библиометрической статистики администраторами науки и финансирующими организациями для оценки инвестиций в научные исследования приняло глобальный характер. Как отмечал Ю. Гарфилд, «мы являемся свидетелями превращения библиометрических исследований в новую отрасль индустрии – оценку результативности научных исследований, выполняемых в университетских и научных коллективах» [Garfield, 2011]. К сожалению, иногда происходит то, что Ю. Гарфилд называл уже в 60-х гг. прошлого века «злоупотреблением цитированием». С одной стороны, это связано с легкостью манипулирования массмедиа этими показателями, с другой стороны — с безответственностью малокомпетентных «специалистов» и администраторов науки, с легкостью берущихся за выполнение такой оценки. Это происходит одновременно при глобальном росте потребностей в использовании библиометрических показателей как инструментов оценки. Недовольство некоторых издателей научных журналов в связи с использованием импакт-факторов привело к декларации по оценке науки — Declaration on Research Assessment (DORA), опубликованной издательской компанией Royal Society Publishing²⁰.

В ответ на этот вызов научного сообщества на прошедшей в сентябре 2014 г. международной научной конференции Science and Engineering Indicators в Лейдене была проведена широкая дискуссия об ответственности наукометрического сообщества

за надежное (лучшее) использование индикаторов наукометрии. В день закрытия конференции профессор А. Ван Раан, президент Европейской ассоциации наукометрических институтов (European Network of Indicators Designers), согласился взять на себя координирующую роль в подготовке и опубликовании консенсусной версии манифеста профессионального библиометрического сообщества, связанного с разработкой стандартов методики оценки и этических норм при их выполнении.

Мы полностью разделяем тревогу международной библиометрической общественности в связи с качеством библиометрической оценки деятельности научных организаций и сектора высшей школы. В свете ряда постановлений Правительства РФ и президента РФ по оценке научной деятельности Федеральное агентство по научным организациям выпустило весьма странное распоряжение «О представлении поквартальных сведений(!!!) о публикационной активности и цитируемости организаций».

Будем надеяться, что Министерство образования и науки РФ сыграет координирующую роль в разработке научных программ в университетах для ознакомления студентов и аспирантов с основами использования библиометрической статистики информационных ресурсов Thomson Reuters, что будет способствовать повышению качества образования в целом и качеству подготовки специалистов по библиометрии.

Пионерские исследования Ю. Гарфилда привели к использованию методологии цитирования и появлению соперника SCI — БД Scopus, созданной компанией Elsevier в 2004 г. Ассоциативные связи на основе цитирования привели к развитию различных алгоритмов информационного поиска, включая HITS and Pagerank. Оба алгоритма

¹⁹Материал взят с сайта URL: <http://wokinfo.com/sci-anniversary.html#timeline>.

²⁰Материал взят с сайта Royal Society Publishing (URL: <http://royalsocietypublishing.org/citation-metrics#question2>).

используют структурные связи между веб-сайтами через гиперссылки.

По мнению профессиональных специалистов по библиометрии, наш XXI в. стал веком библиометрии, воплотившей в себя все идеи создателя SCI доктора Ю. Гарфилда, обладателя множества наград и премий мирового научного сообщества.

Влэдун Г. Э., Налимов В. В., Стяжкин Н. И. Научная и техническая информация как одна из задач кибернетики // Успехи физических наук. 1959. Т. 69. № 1. С. 13–56.

Гиляревский Р. С., Маркусова В. А., Черный А. И. Слово о Юджине Гарфилде // Науч.-техн. информация. Сер. 2. 1995. № 12. С. 23–28.

Гиндилис Н. Л. Становление науковедения в СССР (Середина 60-х годов XX в.) // Науковедческие исследования: сб. науч. трудов / РАН. ИНИОН. Центр науч.-информ. исслед. по науке, образованию и технологиям; отв. ред. А. И. Ракитов. М., 2011. С. 217–272.

Грановский Ю. В. Трудная судьба науковедения в России // Науковедческие исследования: сб. науч. трудов / РАН. ИНИОН. Центр науч.-информ. исслед. по науке, образованию и технологиям; отв. ред. А. И. Ракитов. М., 2010. С. 115–124.

Декларация о науке и использовании научного знания // Всемирная конференция «Наука двадцать первого столетия: Новая ответственность». Будапешт, Венгрия, 26 июня — 1 июля 1999 / под эгидой Образовательной, научной и культурной организации ООН (ЮНЕСКО) и Международного совета научных союзов (ISU).

Добров Г. М. Наука о науке: Введение в общее наукознание. Киев: Наукова думка, 1966.

Коммуникация в современной науке: сб. переводов // под ред. Э. М. Мирского, В. Н. Садовского. М.: Прогресс, 1976. С. 438.

Маркусова В. А. Первый советский указатель библиографических ссылок по информатике // Науч.-техн. информация. Сер. 1. 1976. № 2. С. 30–32.

Маркусова В. А., Черный А. И. Информационная продукция и технология ее подготовки в Институте научной информации США // Науч.-техн. информация. Сер. 1. 1985. № 12. С. 6–15.

Если бы существовали Нобелевские премии по информатике, наукометрии и социологии науки, то Ю. Гарфилд должен был бы стать первым лауреатом такой премии за огромный вклад в развитие этих наук, значение которых в складывающемся информационном мире трудно переоценить.

Маркусова В. А., Ярошевский М. Г. Компьютер и этика цитирования // Природа. 1987. № 9. С. 100–107.

Маркусова В. А. Типичные ошибки в заявках на получение гранта // Науч.-техн. информация. Сер. 1. 1998. № 5. С. 40–42.

Маркусова В. А. Позиции отечественных вузов в мировых рейтингах // Экономика образования. 2010. № 2. С. 35–46.

Маршакова И. В. Система связей между документами, построенная на основе ссылок (по указателю Science Citation Index) // Науч.-техн. информация. Сер. 2. 1973. № 6. С. 3–8.

Михайлов А. И., Черный А. И., Гиляревский Р. С. Основы научной информации. М.: Наука, 1965. С. 435.

Михайлов А. И., Черный А. И., Гиляревский Р. С. Научные коммуникации и информатика. М.: Наука, 1976. С. 435.

Налимов В. В., Мульченко З. М. Наукометрия. М.: Наука, 1969. С. 192.

Оценка научных журналов как каналов связи / З. Б. Барина и др. // Научно-техническая информация. Сер. 2. 1968. № 12. С. 1–11.

ЮНЕСКО [Сайт]. URL: <http://www.uis.unesco.org>.

Ярошевский М. Г. Репрессированная наука / под ред. М. Г. Ярошевского. М.: Политиздат, 1985. С. 431.

Adair W. C. Citation Indexes for Scientific Literature? // American Documentation. 1955. Vol. 6. № 1. P. 31–32.

Broad W. J. Librarian Turned Entrepreneur Makes Millions off Mere Footnotes // Science. 1978. Vol. 202. № 664. P. 853–857.

Cronin B. The need for a theory of citing // Journal of Documentation. 1981. Vol. 37 (1). P. 16–24.

Garfield E. Bernal J. D. The Sage of Cambridge. 4S Award Memorializes His Contribution to the Social Studies of Science // Essays of an information scientist.

Garfield E. Citation Indexes for Science // Science. 1955. Vol. 122. № 3159. P. 108–111.

Garfield E. Citation Index in Sociological and Historical research // Current Contents. 1969. № 9. August 26. P. 42–46.

Garfield E. SCl and ASCA Promote International Contacts // Current Contents. 1970. № 15. April 15. P. 100–101.

Garfield E. The Mystery of the Transposed Journal Lists – Wherein Bradford's Law of Scattering is Generalized According to Garfield's Law of Concentration // Current Contents. 1971. № 17. August 4. P. 222–223.

Garfield E. More on Forecasting Nobel Prizes and the Most Cited Scientists of 1972! // Current Contents. 1973. № 40. October 3. P. 5–6.

Garfield E. Highly Cited Articles. 20 Articles from Russian Journals // Current Contents. 1975. № 45. November 10. P. 374–377.

Garfield E. Citation indexing: its theory and application in science, technology, and humanities. New York: Wiley, 1979.

Garfield E. Information science and technology have come of age – organizational names should show it // Essays of an information scientist. Vol. 3. 1977–1978. Philadelphia, PA: ISI Press, 1980. P. 449.

Garfield E. Premature Discovery or Delayed Recognition – Why? // Essays of an information scientist. Vol. 4. 1979–1980. Philadelphia, PA: ISI Press, 1981. P. 488–493.

Garfield E. Uses and Misuses of Citation Frequency // Essays of an Information Scientist. 1985. Vol. 8. P. 403–409.

Garfield E. Launching the ISI Atlas of Science; For the new year, a new generation of reviews // Essays of an information scientist. Vol. 10. 1987. Peer review, refereeing, fraud and other essays. Philadelphia, PA: ISI Press, 1989. P. 1–6.

Garfield E. Contract research Services at ISI-Citation Analysis for Governmental, Industrial, and Academic Clients // Essays of an information scientist: Of Nobel Class, Women in Science, Citation Classics and Other Essays. Vol. 15. 1992–1993. Philadelphia, PA: ISI Press, 1992. P. 1–6.

Garfield E. A Century of Citation Indexing. Key note address // 12th COLLNET Meeting, September 20–23, 2011, Istanbul Bilgi University. Istanbul, 2011.

Garfield E., Sher I. H. New factors in the evaluation of scientific literature through citation indexing // American Documentation. 1963. № 14. P. 195–201.

Garvey W. D., Griffith B. C. Communication in a science: The system and its modification // Communication in science: Documentation and automation / ed. by A. de Reuck J. Knight. London, 1967.

Graham L. R. What have we learned about science and technology from the Russian experience? Stanford, California: Stanford University Press, 1998.

Hyslop M. R. Documentalists consider machine techniques // Special Libraries, 1953. Vol. 44. № 5. May-June. P. 196–198.

Kaplan N. The norms of citation behavior: Prolegomena to the footnote // American Documentation. 1965. № 16. P. 179–187.

King D. A. Scientific Impact of Nations // Nature. 2004. № 430. 15 July. P. 311–316.

Lederberg J. Foreword // Essays of an information scientist. Vol. 1. 1962–1973. Philadelphia, PA: ISI Press, 1977. P. 10–15.

Lederberg J. How the Science Citation Index got started // The Web of Knowledge. A Festschrift in Honor of Eugene Garfield / ed. by B. Cronin, H. B. Atkins. Medford, NJ: ASIS. 2000. P. 25–64.

Libkind I., Libkind A., Markusova V., Minin V., Jansz M. Index of RFBR as Scientific Information Search and Analysis Tool // The 9th International Conference on Science and Technology Indicators. Leuven, Belgium. 2006. September 7–9. P. 229–233.

May R. M. The Scientific Wealth of Nations // Science. 1997. Vol. 275. P. 793–796.

Price D. J. de S. Science since Babylon. New Haven. Yale University Press, 1961.

Price D. J. de S. Networks of Scientific Papers // Science. 1965. Vol. 149. № 3683. P. 510–515.

Price D. J. de S. Little Science, Big Science // New York, Columbia U.P. 1963. См. рус. перевод: Прайс Д. де Солла. Малая наука, большая наука // Наука о науке. М.: Прогресс, 1966. С. 281–384.

Small H., Belver C. Griffith and John Irvine and Ben R. Martin (as a team) win the 1997 Derek de Solla Price Award // Scientometrics. 1997. № 40(3). P. 357–362.

Small H. Co-citation in the scientific literature: a new measure of the relationship between two documents // Journal of the American Society of Information Science. 1973. Vol. 24. № 4. P. 265–269.

Small H. G., Griffith B. C. The structure of scientific literatures: I. Identifying and graphing specialties. Science Studies. 1974. Vol. 4. P. 17–40.

The Evaluation of Scientific Research // John Wiley & Sons. – CIBA Foundation Conference. Chichester, New York, Brisbane, Toronto, Singapore. 1989. P. 276.

The social function of science, 1939 (переизд. 1946) // Наука и общество: сб. ст. и выступлений. М., 1953.

Vannevar Bush. Science, the Endless Frontier: A Report to the President // U.S. Government Printing Office. 1945.

Vaskovskii V. E., Kostetskii E. Y., Vasendin I. M. Universal Reagent for Phospholipid Analysis // Journal Chromatography. 1975. № 114. P. 129–141. Vol. 5. 1981–1982. Philadelphia, PA: ISI Press, 1983. P. 511–520.

Weinberg A. Science, Government and Information. A report of the President's Science Advisory Committee // The White House, Washington DC. January 10, 1963. P. 55.

Wouters P. The Citation Culture // Amsterdam University. 1999. P. 279.

НАУКА, ТЕХНОЛОГИЯ И ОБЩЕСТВО

SCIENCE, TECHNOLOGY AND SOCIETY

DOI 10.15826/B978-5-7996-1352-5.0004



М. А. Акоев
Заместитель
директора Центра
мониторинга на-
уки и образования
Уральского федера-
льного универси-
тета имени первого
Президента России
Б. Н. Ельцина.

This chapter examines the opportunities and limitations of scientometrics as a tool for R&D resource allocation. The author describes various approaches of using the quantitative and qualitative indicators of research productivity. Universities, research and government organizations demonstrate a growing demand for scientometric resources and tools. The reasons for this demand are considered and the role of scientometric specialists in research organizations is examined. The link between the research and technological processes is demonstrated and the life cycle of research and technological innovations is explored. The author considers various aspects of resource allocation for research and development, and applications of scientometric methods for assessment of resource allocation effectiveness. Lastly, the author considers the time lag between resource allocation and production of measurable output in research and development.

В главе рассматриваются возможности и ограничения наукометрии в процессах принятия решений об эффективности выделения ресурсов для поддержки научной работы. Описаны подходы к построению экспертных оценок количественных и качественных показателей научной продуктивности. Рассмотрены причины роста интереса и востребованности методов наукометрии со стороны администрации научных заведений и государственных органов управления. Рассмотрена также роль специалистов по наукометрии в научной организации. Показана связь между научным и технологическим процессами и представлен жизненный цикл научной и технологической продукции. Рассмотрен вопрос учета численности научного персонала и ресурсного обеспечения научной деятельности для оценки эффективности научного и технологического процессов, а также вопрос численности ученых и вариантов измерения научной продуктивности. Представлена картина взаимосвязей и задержки времени появления измеримых артефактов научной и технологической продукции.

Скажите мне, как вы измеряете мою деятельность, и я скажу вам, как буду себя вести. Если вы измеряете мою деятельность нелогично, не жалуйтесь на нелогичное поведение.

Ильяху Голдратт, автор Theory of Constraint

Наукометрия вызывает противоречивые мнения в обществе. Представители органов власти видят в наукометрии высшее мерило вклада любого ученого или научной организации, а ученые видят в ней в луч-

шем случае веяния, призванные усложнить их жизнь. При этом высказываемые мнения поляризуются от обоснованного указания на особенности отдельных наук, делающих применение методов наукометрии сложными или

невозможными [Игра в цифри, 2011], до сожаления, например, что Хорхе Хирш, автор индекса Хирша, не утонул в молодости (от известного астрофизика и писателя Н. Н. Горькавого [Горькавый, 2014]).

Полярность мнений в отношении возможности использования наукометрии для оценки научной деятельности наблюдается давно, что связано с конфликтом интересов участников и сложностью в достижении компромисса. Противоречие можно было бы преодолеть, если бы существовала процедура, позволяющая оценить качество научного результата без задержки в годы или даже десятилетия от момента получения результата до последствий его внедрения в практику. Суть противоречия в том, что использование наукометрических индикаторов вносит формальный элемент в процесс распределения ресурсов на научные исследования. Этот элемент, с одной стороны, слабо поддается непосредственному манипулированию со стороны соискателя, а с другой стороны, может применяться с существенными искажениями. В действительности наукометрия может быть использована как одна из составляющих процесса распределения ресурсов, причем не только лицами, принимающими решения, но и соискателями. Результат конкурентного распределения ресурсов всегда будет порождать проигравших, но соответствующая организация процесса распределения может привести к тому, что проигравшие будут считать этот результат справедливым [Ловушки мышления, 2014, с. 242–254].

Два ключевых фактора повышают справедливость процесса. Речь идет о полноте и доступности информации и о соблюдении правил [Медоуз, 2010, с. 254–258]. Наукометрия, с некоторыми ограничениями, может сделать более доступной информацию

в процессах принятия решения, но она не может быть применена для всех областей научной деятельности. Например, в процитированном выше высказывании Николая Николаевича Горькавого указывается на неприменимость наукометрических показателей в деятельности NASA, непосредственные результаты работы которой являются прикладными и могут не отражаться наукометрическими индикаторами [Горькавый, 2014]. Кроме того, наукометрия не может гарантировать соблюдение правил, как, собственно, и экспертные оценки, которые будут рассмотрены далее в главе. Если правила, по которым производится распределение ресурсов, сначала декларируются, а потом явно или неявно нарушаются, то применение наукометрических методов не повысит качество принимаемых решений, поскольку нарушение правил будет подталкивать участников процесса к непосредственному воздействию на лицо, принимающее решение. Это относится как к той части процесса, которая обеспечивает выигрыш, так и к части, связанной с представлением результата.

Может быть, если процесс распределения ресурсов настолько сложен, то имеет смысл исключить его и предоставлять всем ученым ресурсы в соответствии с их потребностями, диктуемыми проводимыми исследованиями? Два фактора препятствуют удовлетворению произвольного научного любопытства за счет налогоплательщиков: ограниченность выделяемых ресурсов и управление направлениями исследований.

Число ученых в любой стране должно соответствовать потребности в результатах их деятельности и возможности экономики страны обеспечивать ученых ресурсами. Если направления исследований неизменны, а экономика настолько стабильна, что позволяет выделять достаточное число ресур-

сов, то распределение может быть однократно зафиксировано и необходимость участия в нем не будет обременять ученых. Идиллию стационарного распределения нарушают два фактора: необходимость воспроизводства ученых вследствие естественной убыли и задержка во времени между выделением ресурса на научное исследование и оценкой полученного результата. Не существует способа, который позволяет достоверно оценить способности будущего ученого, кроме как на основании уже выполненных им научных работ. То есть нет возможности, создав набор тестов, выделять из массы абитуриентов тех, кто не просто способен стать ученым, а кто станет ученым в конкретной области в соответствии с прогнозируемой естественной убылью. Подготовка новых поколений ученых является конкурентным процессом, в результате которого ученым становится не столько тот, кто обладает способностями, сколько тот, кто может реализовать свои способности, демонстрируя качественный научный результат. Важно, что для подтверждения способности стабильно получать качественный научный результат дается не одна, а несколько попыток продемонстрировать результаты научных исследований. Не все попытки будут удачны, даже у тех, кто в дальнейшем станет ученым. Однако без соревновательности в отборе будущих ученых нет гарантии, что наукой будут заниматься те, кто готов добиваться качественных результатов.

Конкурентность отбора в научной среде можно оценить по соотношению числа тех, кто стал ученым, к тем, кто пытался им стать. По эмпирическому наблюдению, 75 % авторов опубликуют за все время научной карьеры одну или две статьи, что составляет 25 % общего числа статей [Price, 1986, p. 223]. Отсюда следует, что как минимум только один

из четырех авторов продолжает научную карьеру (более подробное обсуждение см. в разделе 3.6. Показатели для ученых и организаций). Доля студентов, выбравших научную карьеру, из числа всех студентов, получивших диплом, еще меньше.

Кроме того, в описанной выше схеме стационарного распределения ресурсов существует скрытая угроза, что ученый, получив один раз гарантию выделения ресурсов, перестанет вкладывать усилия в достижение новых результатов, а в лучшем случае будет работать над улучшением результата, полученного ранее, что далеко не всегда требуется для дальнейшего развития научного направления или практического использования результата. Нужно сделать еще раз оговорку, что значимость полученных научных результатов достоверно оценивается через много лет и только после применения на практике. Необходимость регулярно доказывать важность пролонгирования выделения ресурсов на продолжение исследований по научному направлению способствует оценке ученым важности проводимого исследования извне, а не изнутри коллектива единомышленников с учетом всех указанных выше особенностей процесса принятия решений.

В обсуждениях в качестве аргумента против регулярной оценки часто приводится пример о времени, затраченном Г. Я. Перельманом на доказательство гипотезы Пуанкаре. В качестве возражения против использования данного примера отметим следующее: во-первых, пример Г. Я. Перельмана скорее доказывает, что лишь очень небольшое число людей готовы концентрировать все свои силы только на решении проблемы и достигнуть результата; во-вторых, не обязательно повторять предыдущий путь, можно работать лектором и точно так же концентрироваться

на сложной задаче, причем никто от вас не будет требовать промежуточных отчетов. Например, в 2013 г. лектор университета Нью-Гэмпшира Итан Чжан (Yitang Zhang) опубликовал в *Annals of Mathematics* статью, позволяющую существенно продвинуться в доказательстве гипотезы о простых числах-близнецах. Свою первую работу [Klarreich, 2013] он опубликовал в возрасте 50 лет, и это был его первый научный результат после получения степени PhD в 1991 г.

Рассмотрев ограничения идеальной модели распределения ресурсов на научные исследования, вернемся к реальности, в которой ресурсы, выделяемые на науку, не только зависят от состояния экономики страны, но и меняются в зависимости от приоритетов развития. По образному выражению В. Л. Кожевникова, директора ИХТТ УрО РАН, состояние науки и технологий — это производная от уровня развития страны [Расчет на одного ученого, 2012]. Каждый новый год научное сообщество может столкнуться с сокращением финансирования, как следствие, часть исследований придется выполнять за меньшие суммы или сокращать персонал. В случае регулярного сокращения финансирования ожидать получения высоких результатов на прежнем уровне не приходится. Стоимость получения научных результатов, которые позже могут быть опубликованы в научных журналах, сопоставима в развитых странах (подробнее см.: [Кожевников, Поляков, 2010] и далее в данной главе). Негативным следствием длительного сокращения финансирования или систематического недофинансирования научных исследований является невозможность быстрого воспроизводства необходимого числа ученых, которые смогут генерировать научный результат высокого уровня, даже в случае кратного роста финансирования. Аналогичная про-

блема медленного роста числа ученых возникает и в случае необходимости быстрого развития приоритетных направлений. Наука как система инерционна, обсуждение сложностей управления инерционными системами выходит за рамки данной работы, и для дальнейшего чтения рекомендуется обратиться к книге: [Медоуз, 2010].

1.1. Процесс оценивания

При принятии решений о распределении ресурсов на научные исследования возникает ситуация, когда уровень компетенции по тематике конкретного исследования лица, принимающего решения, становится недостаточным для единоличного решения. Лицо, принимающее решение, привлекает себе в помощь специалиста, знакомого с предметом исследования, или того, кто может сформулировать аргументированный вариант решения. Привлечение к процессу оценивания решения эксперта в предметной области — наилучший выбор при прочих равных условиях. Замечу, что, по сути, наукометрия, применяемая даже без привлечения экспертов в предметной области, тоже содержит результаты экспертного знания в двух аспектах. Первый аспект — наукометрические показатели требуют для своего расчета привлечения грамотных специалистов. Второй аспект — данные показатели рассчитываются на массиве данных, отражающих результаты научной деятельности с задержкой во времени (подробнее обсуждение см. в разделе 1.5 и в главе 3). В случае привлечения экспертов нужно ответить на три вопроса: кто может выступить экспертом, как сформулировать вопрос эксперту и кто формулирует тематику конкурса по распределению ресурсов.

Если в стране работает небольшая группа специалистов, то отбор экспертов легко мо-

жет быть проведен путем опроса знакомых. А если ключевых специалистов организовать в формальный клуб, то можно решить и проблему экспертизы новых членов для восполнения пула (корпуса) экспертов, необходимого в силу естественной убыли экспертов. Первые академии наук, по сути, являлись клубами экспертов, привлекаемыми для ответов на вопросы, которые волновали монархов. Если ученых становится больше, то можно отобрать пул экспертов и просить их порекомендовать экспертов, которые наилучшим образом могут ответить на вопросы при принятии решения о распределении ресурсов в узкой предметной области. Можно не создавать свой корпус экспертов, а пользоваться существующими, например, можно использовать «Корпус экспертов по естественным наукам»¹. При привлечении экспертов нужно помнить, что возможен конфликт интересов. В силу узости предметной области можно не найти ни одного эксперта, который бы не был связан с претендентами на ресурсы, либо сами претенденты будут единственными экспертами в предметной области. Подробно, с примерами, социальный аспект привлечения экспертов разобран в лекции, прочитанной М. М. Соколовым (текст лекции см.: [Соколов, 2011]). Важно помнить, что экспертная оценка позволяет принять самое точное и быстрое решение, но часто сопровождается существенными затратами на поиск экспертов и организацию процесса опроса. Одна из причин востребованности наукометрических методов оценки как дополнения к экспертным решениям — это возможность быстрее и дешевле получить доступ к уже сформулированным результатам экспертной оценки с учетом всех ограничений на получение и интерпретацию результата.

¹ <http://www.expertcorpus.ru/>

Если мы посмотрим динамику развития корпуса экспертов как элитного клуба, то выявятся две сложности: отбор новых членов только на основе рекомендательной системы очень быстро понижает общий уровень членов клуба, и члены клуба, сменив область своих научных интересов, могут перестать следить за новыми работами в области, по которой были отобраны. Анализ разных вариантов решений в области сохранения качества элитных клубов в популярной форме приведен в соответствующей литературе [Ефимов, 1988]. Основным условием уверенности в должном уровне экспертов является возможность дополнить мнение коллег внешними признаками, свидетельствующими в пользу сохранения высокого уровня эксперта в своей предметной области. В качестве подобного признака можно было бы использовать ссылки на данного эксперта в учебниках, но довольно часто эти ссылки появляются уже после смерти ученого. Приемлемой заменой может служить отбор экспертов на основе наукометрических показателей. Можно пойти дальше и отбирать экспертов только с использованием формальных показателей. В этом случае мы не связаны необходимостью поддерживать элитный клуб экспертов, поскольку будем отбирать нужных экспертов под конкретную задачу. Сложность реализации данной схемы в том, что эксперты в клубе не только выполняют роли оракулов, выносящих решение по заданному вопросу, но также предоставляют информацию о том, кто из членов клуба лучшим образом может ответить на поставленный вопрос.

Наилучшим решением является компромиссный подход между моделью клуба и моделью поиска экспертов по формальным показателям. Так, эксперты, будучи специалистами в предметной области, смогут

преодолеть ограничение отбора экспертов только на основе формальных признаков, а формальные показатели, подкрепляющие мнение экспертов, будут способствовать повышению объективности при отборе. По описанной схеме работают научные журналы. Роль клуба в этом случае играет редакционная коллегия или совет, а рецензенты отбираются на основе мнения членов клуба с использованием при необходимости наукометрических показателей.

О чем мы должны спрашивать экспертов при распределении ресурсов? Если вопрос будет представлять собой задачу распределения всех ресурсов между несколькими претендентами, то эксперты могут провести распределение несколькими способами — и все они будут обоснованы и, возможно, справедливы. Но если мы, распределив ресурсы, хотим получить результат и можем описать желаемый образ будущего, в котором результат реализован, то вопросы экспертам должны формулироваться в терминах желаемого результата — цели конкурса. Отметим, что цели не обязательно формулируются в терминах получения конкретного результата исследования. При распределении ресурсов могут ставиться цели, связанные с поддержкой существующих научных коллективов, с развитием научных групп, созданием нового направления исследований. Основное требование — цель должна быть сформулирована до проведения опроса экспертов и представлена в форме, доступной экспертам, например в форме текста с описанием желаемого результата, в форме открытых вопросов или в форме анкеты. Для выполнения экспертизы должна быть собрана определенная информация об оцениваемых претендентах, причем важно как собрать данные, которые позволят эксперту в соответствии с целью

высказать свое мнение, так и гарантировать верифицируемость этих данных (подробнее экспертная оценка рассматривается в книге: [Хаббард, 2009]).

Вводя понятие цели в процесс экспертного оценивания, нужно понимать, что не бывает целей без заинтересованных лиц. Цели всегда связаны с лицами, получающими выгоду от результатов достижения цели. Не может быть одной цели у организации (например, у университета) и у группы лиц, которые считают необходимым достижение цели. Кроме того, с этой группой лиц должны быть согласны те, кто может заблокировать достижение цели своими действиями. Вопрос о целях экспертизы и консенсус в формулировке цели есть тот пробный камень, который позволяет получить наиболее качественное распределение ресурсов.

При обсуждении формулировок целей для принятия решений по итогам экспертизы важно затронуть два аспекта, а именно: что делать, если претендентов, достойных поддержки, больше, чем ресурсов, и, обратный вопрос, что делать, если претендентов недостаточно для распределения всех ресурсов. Два решения в этом случае могут быть реализованы: либо изменение объема выделяемых ресурсов, либо изменение порогов отбора. Важно при этом учитывать, что выбор между двумя вариантами решений — это чаще всего выбор между кратковременным исчерпанием ресурсов и воспитанием «иждивенческих» ожиданий. Например, используя политический ресурс, можно лоббировать решение, выделяя ресурсы для групп со сниженными показателями в том же объеме, что и для продуктивных групп. Если подобное решение будет принято, то мы либо подкрепим уверенность в том, что высокие показатели — не цель, либо снизим показатели у остальных,

так как, понизив планку в одном месте, мы создаем ожидание повсеместного понижения планки. В случае если остаются нераспределенные ресурсы, более продуктивно вкладывать их в мероприятия по развитию уровня претендентов, исключая «недостойных» претендентов из системы распределения ресурсов.

Последний вопрос в привлечении экспертов: кто формулирует тематику конкурса? С одной стороны, тематика конкурса не может быть задана очень детально, если это не конкурс на выполнение опытно-конструкторской разработки, а с другой стороны, она не может быть сформулирована как решение глобальной задачи, так как в этом случае мы рискуем не дожидаться результатов. Также редко встречаются задачи, для которых практическое применение результатов находится в горизонте планирования от трех до пяти лет. Скорее всего, у такой «краткосрочной» задачи есть вполне конкретный потребитель и коллектив ученых, которые уже работали над задачей, что позволяет видеть контуры практики за теоретическими конструкциями. Кто может сформулировать постановку задачи лучше, чем ученый, уже работающий в предметной области и понимающий возможные связи результатов с потребностями ученых, работающих в других предметных областях? Кажется, что в предложенной схеме существует некий порочный круг: ведь соискатель может сначала сам предложить тему исследования, а затем подать заявку на конкурс и выиграть его. Однако в описанной ранее схеме, схеме самоподстройки, нет конфликта интересов, так как сначала именно эксперты оценивают предложенную тему на предмет перспективности ее разработки, возможности завершить работу и получить результат, и только на следующем этапе ав-

тор темы соревнуется с коллегами за право выполнить работу.

По сути, описанная выше схема является компромиссом, который позволяет снизить затраты на привлечение ресурсов традиционными способами, описанными в форме пяти тактик привлечения ресурсов [Латур, 2013, с. 179–197]. Прочитанную книгу Бруно Латура стоит прочитать в дополнение к рекомендованной выше книге Дерика Прайса, так как они взаимно дополняют друг друга, отражая как количественный аспект деятельности науки, так и социальный контекст, в рамках которого происходит научная деятельность.

Скрытая особенность схемы самоподстройки — это возможный отход от потребности практики как критерия полезности результатов научной работы. Можно нарисовать мрачную картину, когда все ученые — участники процесса начнут «играть в бисер» и оторвутся от практических потребностей. Внешне процесс будет реализовываться, однако цель процесса — практическое применение — не будет достигаться. Участники научного процесса признают, что описанная схема содержит элементы риска, но ее главное преимущество — ускорение научного процесса — окупает возможные издержки.

В чем цель ускорения научного процесса? Почему недостаточно для этого уже задействованных ресурсов? Известно, например, что только в странах OECD в 2011 г. в научных исследованиях было занято 4,3 млн человек и численность ученых в мире за последние полвека только возрастала [OECD, 2014]. Основная причина вложения ресурсов в научные исследования состоит в том, что результаты исследований, будучи примененными на практике, позволяют изменять окружающий мир, приспособливая

его к потребностям общества. Однако для применения на практике научный результат должен быть уже получен. Это означает, что начать работать над его получением необходимо существенно раньше.

1.2. Целеполагание в научном процессе

Воздействие результатов науки и технологии на экономику как систему достаточно подробно рассмотрено в соответствующей литературе [Мокир, 2012, с. 45–106]. Статистическую значимость связи между вложениями в науку и технологию и уровнем экономического развития стран читатель может установить сам, обратившись к наиболее репрезентативному набору данных Организации экономического сотрудничества и развития, доступному по подписке [OECD, 2014]. Книга Джоэля Мокира [Мокир, 2012] является последней в триаде книг, которые дают исчерпывающее описание для понимания всех аспектов наукометрии: книга Дерека Прайса, описывающая базовые подходы к индикаторам; книга Бруно Латтура, описывающая социальный аспект, в рамках которого возникают артефакты, измеряемые наукометрией; книга Джоэля Мокира, из которой читатель почерпнет понимание целей научного и технического процесса.

Дж. Мокир предлагает модель связи научного и технического процессов [Мокир, 2012, с. 31–44], которая в обобщенном виде представлена на рис. 1. На схеме учтено представление Б. Латтура [Латур, 2013, с. 257] о жизненном цикле развития науки. Для описания модели использована нотация системной динамики (обсуждение применимости моделей системной динамики см. в [Price, 1986]), которая позволяет представить влияние элементов системы

друг на друга с указанием обратных связей и задержек в системе [Медоуз, 2010]. Вводится деление на два вида знаний: Ω -знание, представляющее сумму знаний об устройстве мира и способах его познания, и λ -знание, отвечающее на вопрос, как изменить окружающий мир с использованием технологий. Ω -знание создает основу для возможности наращивать λ -знания, что отражено управляющей стрелкой к «крану», ведущему к «бассейну» с именем элемента. Знак «плюс» рядом с управляющей стрелкой обозначает положительную связь: чем больше накапливается в мире Ω -знаний, тем больше технологических знаний мы можем создать. Двойное подчеркивание стрелки означает, что процесс накопления происходит с существенной задержкой; например, от момента определения причины заболевания до момента нахождения способа лечения (технология, направленная на изменение человека) проходит много лет, а для некоторых заболеваний поиск способов лечения может затянуться и на десятилетия.

Технологии, составляющие основу λ -знаний, не ограничиваются только инженерными технологиями, а включают всю совокупность приемов по изменению мира, в том числе общества как составной части нашего мира. Помимо технологий, λ -знания включают в себя и набор «рецептов», которые могут быть фиксацией наблюдений и статистических закономерностей без оснований в форме Ω -знаний. Отсутствие таких оснований не позволяет понимать границы применения рецептов и причины, по которым они работают или могут перестать работать. Например, область медицины, несмотря на большие ресурсы, которые в нее вкладываются, до сих пор является скорее набором высокоуровневых «рецептов», чем

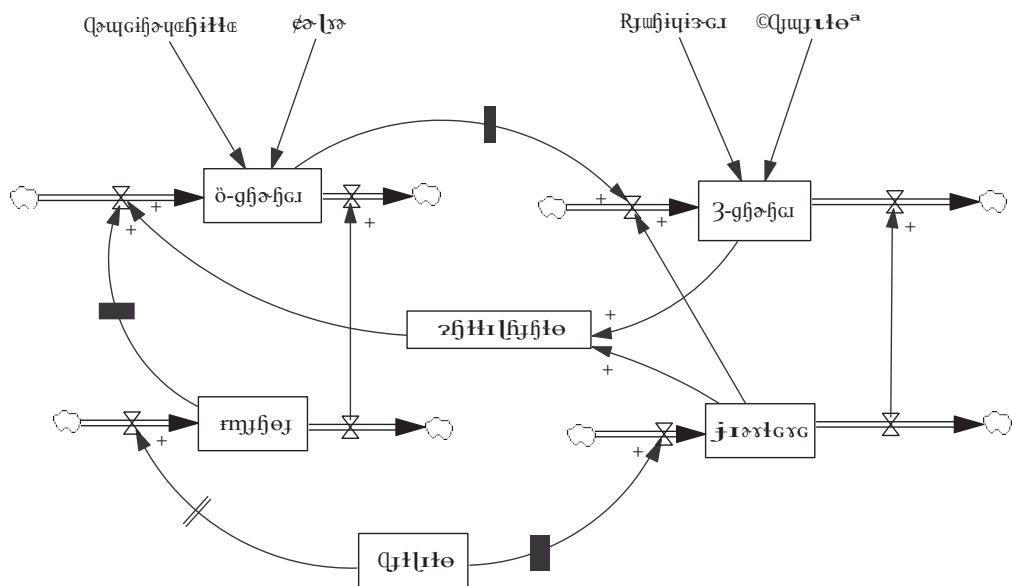


Рис. 1. Модель связи научного и технического процессов по Дж. Мокиру

результатом понимания процессов жизненного цикла организма [Мокир, 2012, с. 232–255]. Важным элементом Ω -знаний является рациональный метод как совокупность приемов, лежащих в основе научного метода [Less Wrong, 2006]. Эти приемы позволяют минимизировать искажения (biases) в процессе рассуждений о внешнем мире.

Каждый следующий шаг по изменению Ω -знаний — это более компактное, непротиворечивое и полное описание мира по сравнению с предыдущим описанием. Другая формулировка данного принципа дана Д. Гильбертом: «Значение научной работы можно измерить числом предыдущих публикаций, чтение которых становится ненужным после этой работы» [цит. по: Нейгебауэр, 1968, с. 147]. То есть каждый новый научный результат должен сокращать затраты на ознакомление с пре-

дыдущими обобщенными в полученном факте научными результатами либо сокращать наши усилия по получению фактов непосредственно из природы. Из данной формулировки можно вывести критерий оценки научной результативности как число фактов, на которое сокращается доступ при знакомстве с данным результатом. Косвенный показатель, достигаемый при публикации результата и позволяющий судить о степени сокращения числа фактов, это число цитирований данного документа как в форме цитат, так и в форме указаний на общеизвестный факт.

При рассмотрении вопроса приращения λ -знаний нужно принять во внимание, что разные научные направления порождают разные дисциплины λ -знаний. Наиболее распространенные λ -знания представляют инженерные и медицинские дисциплины.

В дальнейшем при описании λ -знаний будут приводиться примеры из области инженерного знания, неявно распространяющие описываемые особенности на все виды λ -знаний. Различие в дисциплинах с точки зрения наукометрических исследований несущественно, так как для исследователя важно понимать, что λ -знания служат для удовлетворения существующих потребностей человека. Социальные и гуманитарные науки могут накладывать свои особенности на λ -знания, базирующиеся на них. Эти особенности необходимо учитывать при наукометрических исследованиях данных наук.

В противоположность научному инженерный метод как способ накопления и обращения знаний основан на использовании «рецептов» — эвристик как способа решения задачи, когда наука не представила теорию, обеспечивающую быстрое и точное решение задачи [Коеп, 2003]. Инженер должен решать задачу, не дожидаясь, пока наука предоставит развитую теорию. Показателен пример французских ученых, которые разрабатывали представление о сопротивлении материалов, а в это время английские инженеры строили во Франции мосты (часть этих мостов эксплуатируется до сих пор, пройдя проверку временем).

Цикл обратной связи замыкается по линии от λ -знания через измерительные инструменты к Ω -знаниям, что в теории обеспечивает бесконечное приращение научных знаний (подобные циклы называются усиливающими) через развитие технологий и разработку все более совершенных измерительных приборов. В области действия данного усиливающего цикла различия между инженерными и научными методами стираются, и процессы исследований и разработок не могут быть надежно разделены для целей сбора статистики. Поэтому, на-

пример, при сборе статистических данных о науке и технологии, из общего объема НИОКР не могут быть выделены объемы финансирования только на науку.

Особенность как Ω -знаний, так и λ -знаний в том, что доступ к ним невозможен без людей, которые прошли специальную подготовку и достаточно долго практиковали работу со знаниями. Любой желающий, потратив некоторое время и, возможно, сумму денег, сможет получить доступ практически к любой опубликованной статье. Однако для понимания содержимого статьи необходимо не только быть знакомым с основными цитируемыми источниками, но и владеть всеми базовыми знаниями по предметной области, в том числе и теми, которые явно не отражаются в письменных источниках. В естественных науках, по оценке Д. Прайса, для надежного обоснования представленной в публикации информации необходимо привлечь в среднем 12 предыдущих работ, с которыми нужно ознакомиться для понимания сути опубликованной работы [Price, 1975, p. 125]. Значение публикаций и ссылок в построении сети аргументации достаточно подробно рассмотрено в литературе [Латур, 2013, с. 46–110].

При обсуждении важности построения и поддержания сети аргументации можно услышать контраргумент, что для некоторых областей науки эта процедура является достаточно сложной либо вообще невозможной в силу малого числа исследователей, работающих в одной области. Р. Коллинз в своем исследовании социологии философии показал, что процессы в философии подобны процессам в естественных науках с двумя особенностями, которые заключаются в существенно замедленном темпе развития и иной, чем в естественных на-

уках, форме построения сети аргументации [Collins, 2000, с. 65–205].

Для общества публикации играют роль внешних признаков, повышающих уверенность в том, что научная работа действительно ведется, а наукометрические показатели играют роль сигналов об уровне выполненной работы. Однако наличие артефактов, функция которых понятна, а содержание недоступно для большинства населения, может порождать недоверие к работе ученых в долгосрочной перспективе. Для предотвращения снижения доверия сами ученые предпринимают меры для популяризации своих достижений. Такие меры (например, сайт научной организации) увеличивают открытость информации о проводимых исследованиях и научных достижениях. Снижение уровня доверия к работе ученых в обществе потенциально опасно тем, что часто приводит к сокращению финансирования, которое перераспределяется в пользу других интересов общества. Регулятором в усиливающем цикле выступает уровень ресурсов, выделяемых на поддержание численности ученых и инженеров.

В λ -знаниях существует аналогичная ситуация с доступом к ним, с той лишь разницей, что не требуется обеспечивать доступность документации для общества, скорее ставится цель усложнить доступ к технологическим секретам для поддержания конкурентного преимущества. Вопрос повышения доверия к инженерам как носителям λ -знаний тоже не ставится, так как выгода от вложений в технологию верифицируется существенно проще, чем в науку.

Когда мы обсуждаем уровень Ω -знаний в обществе, мы должны себе отдавать отчет в том, что это только те знания, доступ к которым можно получить через ученых. Отметим, что, по оценке Д. Прайса, порядка 10 %

статей не читаются никем, кроме авторов [Price, 1986, p. 108]. Никакой трагедии в том, что часть информации не востребуется учеными, нет. Однако необходимо помнить, что даже если результаты некоторых исследований зафиксированы в виде публикации и эти публикации доступны, то нет гарантий, что для использования отраженного в статьях знания не понадобится времени, сверх необходимого для чтения и понимания написанного в работах.

Если для доступа к Ω -знаниям требуются ученые, то как на их основе инженеры (практики) могут способствовать росту λ -знаний? Мир ученых и мир инженеров не изолированы друг от друга, а возможность сформулировать вопрос и понять ответ гарантируется общей образовательной базой. Для создания инновационных технологических решений подготовку инженеров необходимо проводить совместно с подготовкой будущих ученых. Три из четырех авторов, которые опубликовали одну или две работы в естественных науках и не продолжили в дальнейшем карьеру ученого, скорее выбирают карьеру инженера (обсуждение этого факта см. ранее в главе). Совместная подготовка ученых и инженеров и вовлечение их в научный процесс гарантирует, что выпускники, выбравшие инженерную карьеру, будут не только лучше понимать язык науки и будут способны читать научные работы, но и на практике получают представления о результатах научной деятельности, в которой участвовали в качестве помощников. Предположение о том, что университет, в котором выполняются научные исследования, позволяет наилучшим образом подготовить выпускников к практической деятельности, лежит в основе того, что многие индикаторы рейтингов вузов являются оценкой

уровня именно научных исследований (обсуждение рейтингов см. в части 3.12. Использование библиометрических данных при построении рейтингов вузов и научных организаций).

Должны ли все вузы обязательно вести научную работу, выпуская специалистов высокого качества? Ответ, скорее, нет, не должны. Только порядка 175 университетов в США являются исследовательскими [Розовски, 1995]. Бюджеты исследовательских университетов существенно больше, чем в вузах, которые занимаются только преподаванием. Отвечая на поставленный вопрос, можно предположить, что некоторое время достаточно сосредоточиться только на выпуске практиков, не развивая в вузах науку и экономя деньги. Однако неизбежно наступает момент, когда уровень знаний преподавателей в вузах настолько отстает от текущего уровня развития науки, что уровень практиков в стране становится неконкурентным по сравнению с развитыми странами мира.

Отметим, что еще один аргумент в пользу исследовательских университетов состоит в том, что их выпускники являются лучшим каналом трансфера технологий. Это связано с тем, что они не отягощены грузом существующих в практике представлений, а получили знания о передовых исследованиях, которые будут доступны в практике уже после того, как выпускники начнут свою практическую деятельность. Слабым местом исследовательских институтов, не ведущих подготовку студентов, является то, что для воплощения их результатов в практику требуется содержать структуру, занимающуюся прикладными и опытно-конструкторскими работами, а это замедляет практическое применение научных результатов.

Что является ограничителем роста знаний в модели связи научного и технического процессов (рис. 1)? Это естественная убыль ученых и практиков. Если соответствующая область знаний не находится в области внимания ученых и практиков, то с утратой исследователей и/или инженеров есть риск быстро не восстановить уровень владения знаниями, и часто цена восстановления будет существенной. Риски утраты знаний стали значимы к концу XX в., вследствие того что циклы жизни артефактов, полученных в процессе применения λ -знаний, стали сопоставимы с продолжительностью человеческой жизни. Во врезке приведены примеры утраты знаний, причем важно отметить, что ни в одном случае утрата не была фатальной. Случай с каиновой кислотой скорее иллюстрирует пример утраты знаний в организациях, имеющих доступ к исходному сырью, из которого производился продукт, однако издержки всех участников процесса не становятся от этого меньше. Можно ли средствами наукометрии выявлять области с риском утраты знаний? Ответ — нет, можно только выявить области, которые прекратили развиваться как научные дисциплины, но без экспертов определить причину прекращения развития практически невозможно. В связи с этим нужно отметить два момента: подобный мониторинг с привлечением экспертов был бы очень дорогим (так как эксперты выключаются из процесса производства новых знаний); кроме того, мониторинг не гарантирует полноты нахождения всех случаев потенциальной утраты знаний.

Если в отдельной стране вследствие убыли ученых или практиков утеряны отдельные области знаний, то всегда есть возможность привлечь нужного специалиста из-за

границы. Для этого достаточно поддерживать внутри страны необходимое число экспертов, которые могут идентифицировать носителя требуемых знаний. Однако это не всегда возможно как по экономическим, так и по политическим причинам. Поэтому в стране должен поддерживаться как минимум уровень владения всеми видами знаний, которые необходимы для воспроизводства критических технологий. Максимум ученых и практиков определяется скорее возможностью системы образования воспроизводить кадры и возможностью экономики поддерживать процесс воспроизводства знаний.

1.3. Выделяемые ресурсы

Для составления оценки научной деятельности помимо экспертных и наукометрических оценок, основанных на библиометрических показателях, необходимо учитывать, какие ресурсы выделялись на проведение исследований и сколько людей работало на получение результата.

Оценим, как много денег выделяется на проведение научных исследований. Рассмотрение ресурсов ограничивается, как правило, только финансовыми показателями. Это связано с недоступностью информации о материально-техническом обеспечении научных исследований на

Каиновая кислота — биологически активное вещество, выделенное из морских водорослей *Digeneasimplex* японскими исследователями S. Murakami, T. Takemoto и Z. Shimizu в 1953 г. [Nitta, 1958]. *Digeneasimplex* используется в традиционной японской медицине. Название кислота получила от японского слова *Kaininso* — призрак моря. С 1970-х гг. кислота широко используется в нейробиологии как активатор определенного типа глутаматных рецепторов [Kainic acid, 2014]. Химическая структура кислоты приведена на рис. 2.

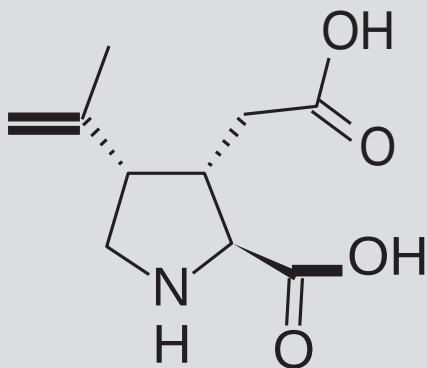


Рис. 2. Химическая структура каиновой кислоты

В конце 1990-х гг. каиновая кислота начала исчезать из каталогов всех фирм-поставщиков, а когда снова появилась в каталогах, то стала стоить значительно дороже [Kainic acid, 2014]. Оказалось, что вне зависимости от этикеток все мировые запасы каината производил некий японец, который ушел на пенсию, и способ выделения каината из водорослей в промышленном масштабе был утерян. Прошло несколько лет, прежде чем фирмы стали производить синтетический каинат, который был значительно дороже выделенного из водорослей [Idelsong, 2009].

Корпоративная память и обратная контрабанда.

В конце 2011 г. в сети Интернет был опубликован анонимный текст, описывающий утерю конструкторской документации на нефтехимический завод, построенный в начале 1980-х гг. [Institutional memory..., 2011].

Завод все время до написания текста работал, проблема возникла в тот момент, когда было принято решение о его модернизации. В тексте описаны три стороны проблемы: физическая утрата документации (в том числе и в электронном виде); неполное отражение в документации информации о принятых при проектировании решениях и ограниченная доступность специалистов, которые могут по работающему заводу восстановить проектную документацию, необходимую для его реконструкции.

уровне подавляющего большинства стран, например, ее нет в статистических сборниках OECD [OECD, 2014]. В сборниках на уровне некоторых стран подобная информация доступна. Например, детальная информация о ресурсах, финансовых и оборудовании, привлекаемых для выполнения научных исследований в вузах и научных организациях Министерства образования и науки, собирается Северо-Западным научно-методическим центром (СЗНМЦ) при СПбГЭТУ «ЛЭТИ» им. В. И. Ульянова (Ленина) и публикуется ежегодно в форме статистических сборников. В этих сборниках публикуются, например, показатели, агрегированные по всем вузам страны [Научный потенциал вузов и научных..., 2014], а также доступны отдельные выпуски по каждому федеральному округу с информацией по каждому вузу. Для организаций РАН подобная информация

доступна в изданиях Института проблем развития науки РАН².

На рис. 3 приведено распределение доли от бюджета на НИОКР в некоторых странах мира в зависимости от тематики по данным OECD [OECD, 2012]. Данные по другим странам и периодам могут быть получены из того же источника, который регулярно обновляется. Важно, что представленная статистика не учитывает деление на исследования и разработки, что является результатом сложности создания и использования методики разделения на чистые исследования и разработку в любой НИОКР. Тематики НИОКР классифицированы по шести направлениям: естественные, технические, медицинские, сельскохозяйственные, социальные и гуманитарные. Распределение на диаграмме

²<http://www.issras.ru/>

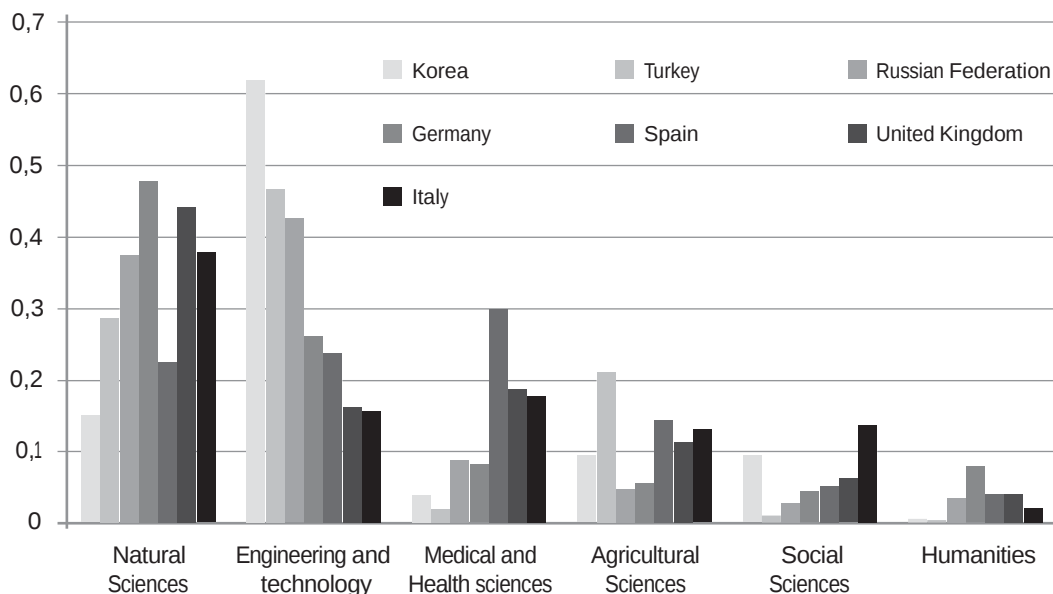


Рис. 3. Распределение доли финансирования НИОКР по отдельным странам и тематикам в 2011 г. (источник: [OECD, 2012])

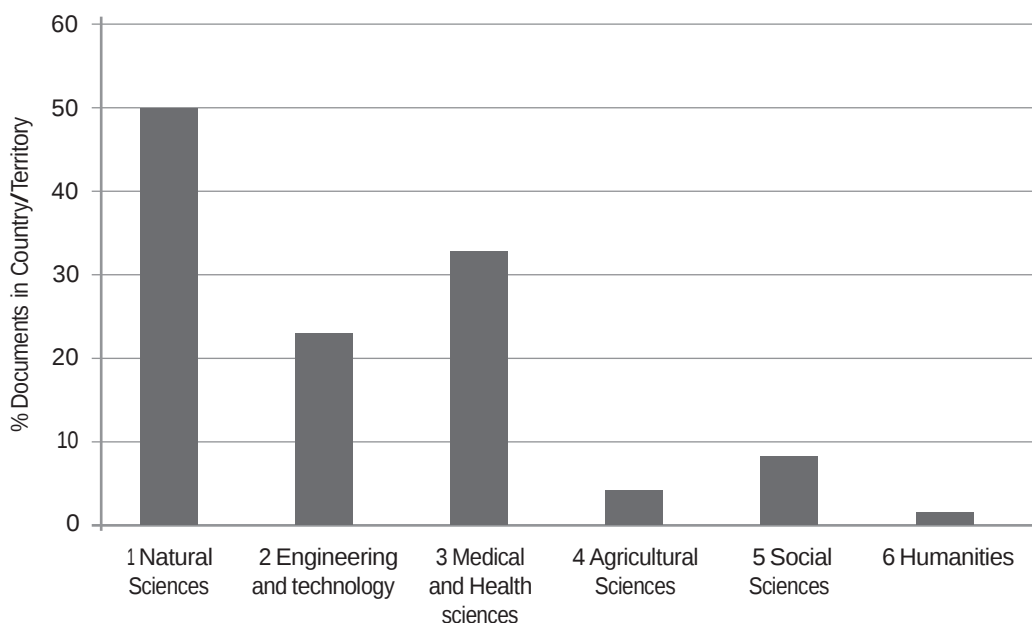


Рис. 4. Распределение публикаций всех стран за период 2011–2013 гг. по тематикам OECD.

Источник: InCites, Thomson Reuters (2012)

(Data Processed: March 31, 2014. Data Source: Web of Science CC)

имеет явно выраженное смещение расходов в область естественных и технических НИОКР. В среднем по миру расходы также смещены в область медицинских исследований, но в большей степени смещение обеспечено за счет расходов отдельных развитых стран, прежде всего США.

Если мы посмотрим на распределение доли публикаций по тематикам OECD за 2011–2013 гг. (рис. 4), то увидим повторение картины распределения финансирования с поправкой на медицинские исследования. Выбор периода агрегирования данных в три года начиная с 2011 г. определяется тем, что финансирование 2011 г. влияет на число публикаций с задержкой. Корректнее было бы сопоставить финансирование и публикации за более длинный интервал вре-

мени, но картина распределения не меняется на протяжении последних 30–40 лет.

В чем причина неравномерности распределения финансирования? Во-первых, разные страны специализируются на тех тематиках, которые дают максимальный вклад в решение задач, наиболее востребованных в экономике страны. Например, Турция отдает существенно больший приоритет исследованиям в сельском хозяйстве, чем медицине. Южная Корея основные ресурсы вкладывает в инженерные НИОКР. Во-вторых, сдвиг в первые три тематические области определяется тем, что в каждой из них имеются структуры, концентрирующие существенные финансовые ресурсы, и существует традиция вкладывать финансы в получение научного результата

с последующей отдачей в практической деятельности. Для естественных и технических НИОКР такими структурами являются государство, обеспечивающее обороноспособность и безопасность страны, а также промышленность, использующая полученные результаты для создания инновационных продуктов. Вопросы распределения ресурсов достаточно подробно рассмотрены в литературе [Латур, 2013, с. 261–280)].

Нужно отметить, что, несмотря на существенные затраты на финансирование медицинских исследований и большое число публикаций в данной области, прогресс здесь больше носит характер накопления отдельных «рецептов» [Мокир, 2012, с. 232–255]. Также необходимо отметить, что именно статьи по медико-биологическим

исследованиям содержат больше всего не-верифицируемых исследований.

Может быть, смещение затрат в область технических НИОКР диктуется тем, что большая часть затрат попадает на этап опытно-конструкторских разработок? Исследования показывают, что доля расходов на фундаментальные исследования, которые дают больше всего статей в научных журналах, в структуре НИОКР стабильна и составляет порядка 10 %. Остальные расходы распределяются между прикладными исследованиями и разработками [Латур, 2013, с. 270].

Также распределение выделяемых ресурсов в течение жизненного цикла НИОКР можно проиллюстрировать (рис. 5), разделив период работы по уровням готовности технологии к использованию на практике

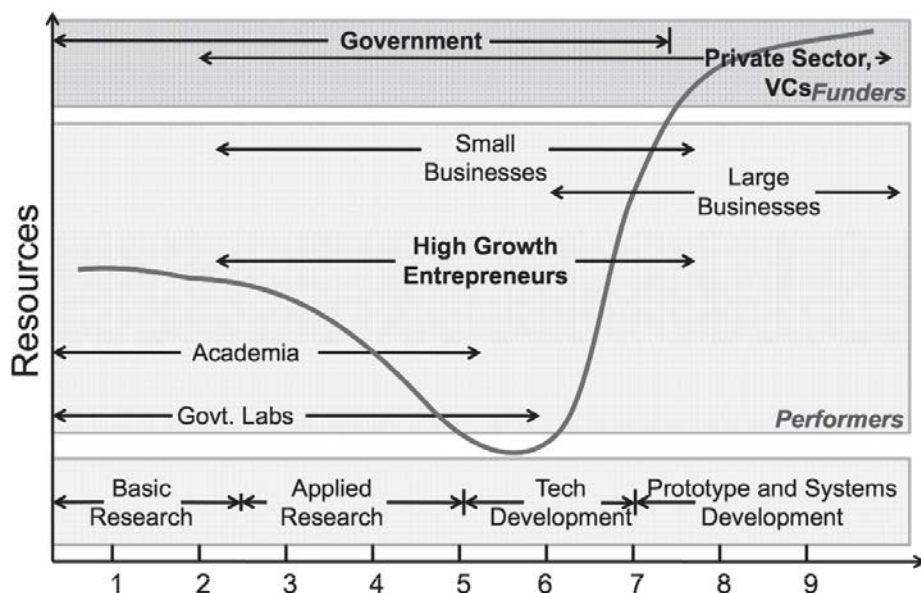


Рис. 5. Распределение затрат ресурсов на выполнение НИОКР. Цифры означают уровень готовности технологии TRL. В блоке Funders указано распределение источников финансирования по секторам экономики. В блоке Performers указано распределение исполнителей.

Иллюстрация из: <http://wp.ifi.uzh.ch/leitner/?p=346>

(Technology readiness level – TRL) [Technology readiness level, 2014]. На рис. 5 видно, что наиболее затратные этапы наступают после этапа фундаментальных исследований (уровни готовности 1–2). Также необходимо отметить, что этап фундаментальных исследований финансируется только государством как наиболее рисковый и с существенной задержкой в получении практических результатов.

Оценим, как много было потрачено денег на производство видимого результата — научной публикации. При оценке нужно выбирать, за какой период провести измерение. Как правило, выбирают пять лет как минимальный срок, за который завершаются исследовательские программы и сглаживаются естественные изменения финансирования. В какой валюте считать? В разные годы курсы валют отличаются, также отличаются и уровни жизни в разных странах. Для проведения межстрановых сравнений OECD использует курс национальных валют

к доллару США по паритету покупательной способности (PPP \$), который рассчитывается по соотношению суммы, которую необходимо заплатить за одинаковый набор товаров в сравниваемых странах. Результаты расчета стоимости за пятилетний период приведены в табл. 1. Отдельно в таблице выделено значение по всем странам OECD. Отметим, что расходы в вузах³ по аналогичной методике оцениваются в диапазоне от 50 до 150 тыс. PPP \$, что объясняется большей сосредоточенностью вузов на исследованиях, чем на разработках. Также стоит отметить, что высокая стоимость одной публикации в России и Китае объясняется большей долей расходов на технологические разработки, чем на исследования.

Аналогичным образом оценим количество статей, опубликованных на одного исследователя в течение пяти лет. Результаты оценки приведены в табл. 2. Методика расчета

³Отчет НИР для разработки плана развития науки в УрФУ, первый этап. Thomson Reuters, 2013.

Таблица 1

Стоимость одной публикации за период в пять лет в PPP \$*

	2000– 2004	2001– 2005	2002– 2006	2003– 2007	2004– 2008	2005– 2009	2006– 2010	2007– 2011	2008– 2012
Великобритания	342,1	351,9	358,9	362,1	365,2	366,2	359,6	349,7	337,8
Франция	597,4	607,3	609,6	609,7	613,5	622,5	631,4	639,3	645,1
Германия	682,3	695,2	713,3	730,6	753,1	771,5	791,4	811,0	828,3
--OECD	839,1	853,1	866,6	887,4	909,9	924,5	932,4	935,6	934,1
США	864,1	877,1	890,6	915,9	944,5	967,0	978,6	984,3	987,0
Южная Корея	973,2	942,2	933,4	956,9	976,4	983,5	993,3	996,3	991,5
Россия	483,9	544,5	621,2	706,7	784,3	887,2	971,5	1018,9	1071,6
Китай	1102,4	1101,2	1099,4	1099,5	1106,0	1141,4	1183,2	1217,7	1249,2
Япония	1186,0	1239,1	1300,4	1375,8	1452,0	1496,1	1527,3	1547,6	1553,4

* Приведены расчеты автора, значения упорядочены в порядке возрастания по последнему пятилетнему периоду, число статей определено по Global Comparisons. InCites, Thomson Reuters (2012). Data Processed: June 1, 2014. Data Source: Web of Science CC, расходы на НИОКР по [OECD, 2014].

Таблица 2

Число публикаций в период пяти лет на одного занятого в исследованиях (FTE)*

	2000– 2004	2001– 2005	2002– 2006	2003– 2007	2004– 2008	2005– 2009	2006– 2010	2007– 2011	2008– 2012
Россия	0,29	0,29	0,29	0,30	0,31	0,32	0,33	0,33	0,35
Китай	0,32	0,36	0,41	0,46	0,51	0,50	0,53	0,53	0,54
Япония	0,70	0,71	0,75	0,72	0,72	0,69	0,68	0,68	0,71
Ю. Корея	1,08	0,98	1,06	1,10	1,17	1,11	1,10	1,09	1,14
--OECD	1,28	1,27	1,30	1,28	1,31	1,32	1,33	1,35	1,36
Франция	1,77	1,75	1,71	1,70	1,69	1,76	1,75	1,72	1,74
США	1,68	1,67	1,67	1,61	1,70	1,76	1,77	1,82	1,79
Германия	1,62	1,61	1,65	1,68	1,73	1,78	1,79	1,80	1,80
Великобритания	2,58	2,45	2,32	2,20	2,17	2,07	2,11	2,21	2,30

* Приведены расчеты автора, значения упорядочены в порядке возрастания по последнему пятилетнему периоду, число статей указано по Global Comparisons. InCites, Thomson Reuters (2012). Data Processed: June 1, 2014. Data Source: Web of Science CC, исследователи FTE на первый год в периоде пяти лет [OECD, 2014].

числа публикаций идентична использованной при построении табл. 1. Число исследователей дано на начальный период пяти лет и приведено к эквиваленту по полной занятости (FTE). Отдельно фоном выделены общие значения по всем странам OECD. Одна из самых низких продуктивностей по числу опубликованных за пять лет статей у России. Причина такого положения в том, что большая часть государственных расходов в нашей стране достается промышленности, при этом государство — основной источник расходов на НИОКР. Для Китая низкие значения показателя являются результатом активного роста вложений в науку при отставании роста числа публикаций.

1.4. Особенности анализа количественных данных

При проведении наукометрического исследования необходимо учитывать, для кого проводится анализ, кто его проводит и с какой целью, какие используются источники

данных. По сути, наукометрия является формой количественного анализа на основе специализированных источников данных. Количественный анализ возможен только при наличии источников данных, релевантных решаемым задачам. Собственно доступность массивов данных относительно научной деятельности и делает возможным проведение наукометрических исследований без наличия большого штата технических специалистов по поиску, извлечению и обработке первичных данных. Представление о количественном анализе как методе поддержки принятия решений изложено в книге [Дэвенпорт, Ким, 2014], для углубленного изучения аспектов количественного анализа можно обратиться к рекомендованным в ней источникам.

Рассмотрим аспекты проведения количественного анализа данных, которые позволяют эффективно выполнять наукометрические исследования. Первое, с чего нужно начать, это определить, для кого вы

будете готовить анализ, если, конечно, вы не проводите научное исследование с целью подготовки публикации. Потребители вашего анализа определяют цель проводимого анализа (например, в виде вопросов исследования) и форму представления результатов. С вопросами «что?» и «для кого?» нужно определиться перед началом исследования, а не после получения результатов, так как можно потратить много времени впустую, собирая и анализируя данные, которые не дают ответа на поставленные вопросы. Довольно часто потребители сами не могут четко сформулировать задачу. Рассмотрим две возможные причины: недостаток знаний о возможностях наукометрии и нечеткая формулировка решаемой проблемы. В первом случае необходимо проводить обучение, на примерах показывая возможности и ограничения методов наукометрии. Во втором случае рекомендуется сначала обдумать проблему, провести экспресс-анализ и на примерах уточнить поставленную задачу. Проще выполнить несколько предварительных анализов, последовательно уточняя преследуемую цель, чем получить ответ на бесполезный вопрос.

Самый главный аспект деятельности аналитиков, выполняющих количественный анализ, состоит в том, что бесполезно ограничиваться представлением фактов, которые следуют из обработанных данных. Нужно представить варианты решений, возможно даже рассмотрев сценарии реализации того или иного решения. Представление вариантов решения вместе с результатами анализа не заменяет для потребителей необходимость принимать решение самостоятельно. Однако предложенные вами варианты должны очерчивать потенциально реализуемые возможности с анализом

негативных последствий. Совершенно недопустимо представлять только один вариант действий по результатам анализа, минимальное число представляемых вариантов — два. Чем больше разных, но реализуемых вариантов вы сможете предложить, тем адекватнее будет принятое решение [Ловушки мышления, 2014, с. 40–56].

Важным моментом при выработке рекомендаций является необходимость понимания процессов, которые приводят к наблюдаемым данным. Ориентируясь только на статистические данные, легко дать вполне выполнимые рекомендации, которые, однако, не приведут к получению желаемых результатов. Например, многие публикации, выполненные российскими авторами совместно с зарубежными, обладают сравнительно более высокими оценками качества, чем выполненные только россиянами. Дать рекомендацию публиковать больше статей с зарубежными соавторами — это путь к увеличению совместных публикаций с иностранцами, а не путь повышения качества публикаций.

Перед представлением результатов необходимо обязательно проверить восприятие итогового документа на коллегам, это позволит избежать досадных ошибок в анализе. Чтобы выполнить и представить анализ наилучшим образом, рекомендуется обратить внимание на методику, используемую одной из ведущих консалтинговых фирм мира [Инструменты McKinsey, 2009]. К числу наиболее частых недочетов в представляемых анализах относятся отсутствие либо самой методики, по которой был выполнен анализ, либо ее обоснования и отсутствие объяснения выбора бенчмарков, т. е. референтных примеров. Подробнее остановимся на необходимости объяснения выбора бенчмарков, которые позволяют провести

сравнительное исследование. Если вы сравниваете результаты деятельности ученых, то как минимум должны позаботиться о том, чтобы выбранные для сравнения коллеги работали в близкой предметной области, а в идеале были сравнимы по времени научной карьеры и типу организации. Если не ставится цель проведения глобальных сравнений, то желательно учитывать и регионы мира, в которых работают сравниваемые ученые. Выбор корректных бенчмарков — это достаточно сложная задача, но, будучи выполненной один раз, она позволяет в дальнейшем использовать созданный набор элементов сравнения, обязательно каждый раз проверяя, не нарушилось ли условие корректности сравнения.

Для представления результата желательно разработать единообразный стиль оформления для повторяющихся результатов анализа, что позволит обратиться к предыдущим выводам при проведении сравнений. Это также уменьшит затраты потребителей на понимание логики изложения в повторяющихся документах. Обязательно сохраняйте все рабочие материалы анализа, ведь это позволит в будущем восстановить важные детали, которые быстро забываются при выполнении последующих работ. Метод ведения дневника проекта — наилучший метод для повышения качества проведенного анализа. В этом дневнике нужно кратко записывать принятые решения, чтобы была возможность через две недели после сдачи аналитического отчета провести разбор полученного результата с целью понять причины совершенных ошибок и отметить удачные находки, которые стоит применять в дальнейшем.

Без доступа к источникам данных количественный анализ невозможен. Каждый источник обладает своим набором харак-

теристик, которые нужно учитывать при выполнении анализа данных. При обработке информации нужно больше внимания уделять возможностям автоматизированной обработки данных с использованием программ, а не ручному манипулированию сырыми данными. Использование автоматизации при обработке данных позволяет снизить число ошибок, получаемых вследствие рутинной ручной обработки, и больше сосредоточиться на решаемой задаче. Также стоит обратить внимание на доступность аналитических баз данных, в которых информация предварительно обработана и данные подготовлены для сравнительного анализа. Затраты на использование подобных баз данных окупают себя при необходимости выполнения большого объема аналитической работы.

Сложность выполнения наукометрических исследований связана с тем, что при анализе данных нужно понимать контекст проводимых исследований. Если проводить анализ без учета тематики исследуемых предметных областей, то ценность результатов будет подобна ценности информации о средней температуре тела по больнице (не позволяющей выявить, чья температура ниже/выше нормальной). В этом случае возможны два варианта решений. Во-первых, можно создать пул консультантов в предметной области, с которыми проводить обсуждение проводимых исследований. Во-вторых, можно воспитать наукометрических специалистов из ученых, работающих по тематике исследований, и это наилучшее решение. Являясь также активными учеными, наукометрические специалисты, с одной стороны, смогут учитывать контекст проводимых исследований, а с другой — будут лучшими проводниками результатов анализа.

1.5. Задержки и артефакты научного и технологического процессов

В процессе выполнения наукометрических исследований необходимо учитывать три ключевых аспекта научного и технологического процессов: 1) кто генерирует научные результаты; 2) недоступность в общем случае измерения влияния научных результатов на практику; 3) задержку между моментом появления идеи и получением первых измеримых результатов.

При анализе наукометрических данных важно видеть коллективы ученых, работающих над получением результатов, которые вы измеряете. Иначе возникает соблазн воспринимать артефакты научной деятельности как набор событий, которые демонстрируют нестабильные закономерности и которыми можно управлять, увеличивая предоставляемые ресурсы. Связь с практикой, обсуждаемая в разделе 1.2, не порождает видимых артефактов и может быть оценена только по косвенным признакам. Однако наибольшие трудности создают задержки в процессе, порождающем научные артефакты.

Структуру, в рамках которой развивается наука, можно определить как локальную научную группу численностью до 15 человек. Объединяющим началом для группы является научный руководитель, который определяет направление исследований группы. Вместе группу держит несколько моментов. Во-первых, это возможность эффективно использовать выделенные ресурсы (площади, научное оборудование, музейные коллекции и т. п.). Во-вторых, это возможность сокращать затраты членов группы на ознакомление с новыми научными результатами, полученными вне группы. И в-третьих, это возможность проводить обсуждения полученных результатов перед

их публикацией с той целью, чтобы поддерживалась общая репутация группы и чтобы сообщения этой группы не игнорировались внешними коллегами. Структура научной группы подробно исследовалась в литературе [см.: Латур, 2013].

Успешность функционирования научной группы определяется способностью руководителя привлекать финансирование для выполнения исследований. Важно, чтобы это финансирование было регулярным и в объеме, достаточном для функционирования группы. Регулярность привлечения финансирования свидетельствует о том, что лидер понимает интересы лиц, распределяющих ресурсы, и способен привлечь ресурсы внешних групп для обсуждения и совершенствования результатов работы своей группы. Способность группы преобразовывать привлекаемые ресурсы (свои и других групп) в измеримые артефакты научного и технологического процесса и оценивается в конечном счете при анализе наукометрических показателей.

Например, если мы видим стабильный рост публикаций по узкой тематике в организации, это означает, что группы, генерирующие научный результат, смогли получить существенное финансирование. Аналогично: если в организации увеличивается доля совместных работ с учеными других организаций по узкой тематике, то весьма вероятно, что увеличили финансирование на научные обмены, но для выяснения, в какой организации произошло увеличение финансирования, нужны дополнительные данные.

Говоря о научной группе, нужно понимать, что можно повысить эффективность ее работы, сократив издержки группы на обсуждение полученных ею результатов. Это можно реализовать, сконцентрировав в одном месте научные группы, работающие

над своими проблемами в рамках большой темы. Эффект от концентрации научных групп и его причины рассмотрены в литературе [см.: Яблонский, 2001, с. 228–229]. Любая научная структура, которая специализируется на узкой научной тематике, организует деятельность научных групп внутри себя, будет эффективнее и результативнее, чем разрозненные научные группы. Наиболее известный пример концентрации научных групп для решения сложной проблемы — это Манхэттенский проект.

Специализированная научная структура, например научно-исследовательский институт, если ее существование не связано с уникальным оборудованием, будет выключена из процесса воспроизводства научных кадров, что затрудняет распространение знаний, которые в ней генерируются, и повышает риск застоя. Распределение научных групп по университетам позволяет преодолеть подобное ограничение, но ценой замедления процессов обсуждения, так как коллеги могут находиться территориально далеко. Как эффективно организовать обсуждение в распределенных коллективах? Несмотря на развитость современных средств коммуникации, требуется личное присутствие на таких обсуждениях, так как это позволяет гарантировать концентрацию участников данного процесса на обсуждаемой задаче, а фиксированные даты обсуждения обеспечивают регулярность подготовки отчетов о проделанной работе.

При оценке влияния научных результатов на практику нужно понимать, что патенты не являются артефактами практической деятельности, а возникают в результате работы над НИОКР. Патенты предназначены для фиксации прав собственности и для оповещения о владении технологией, но не для непосредственной их реализации

на практике. Отметим, что для многих областей практической деятельности патенты не играют существенной роли, а права собственности эффективней охраняются секретами производства. Скудность информации о деталях технологического процесса объясняется тем, что результаты практики в основном генерируются на предприятиях, которые противятся раскрытию деталей своей работы (рис. 5). Можно предложить три параметра оценки результативности технологической деятельности: 1) объемы договоров на исследования со стороны хозяйствующих субъектов; 2) востребованность выпускников, которые участвовали в научном процессе во время учебы; 3) контент-анализ публикаций в журналах, публикующих результаты прикладных работ. Ни один из параметров, приведенных выше, не позволяет провести достоверную количественную оценку, скорее мы получим качественные характеристики.

В качестве вспомогательного инструмента при проведении прикладных наукометрических исследований можно использовать схему, представленную на рис. 6 (см. на предыдущей странице). На схеме изображены развернутые во времени артефакты, получаемые в процессе реализации НИОКР. Стрелки указывают на связь элементов процесса с порождаемыми артефактами. Артефакты над осью процесса практически не поддаются оценке, так как чаще всего недоступны. Артефакты под осью обладают разной степенью доступности, но, как правило, обнаруживаются в разных источниках и в какой-то степени являются измеримыми. Например, данные о части оборудования, установленного в лабораториях Европы, доступны на портале MERIL⁴, а при поиске по полным текстам публикаций можно найти

⁴<http://portal.meril.eu/>

информацию о том, в каких работах данное оборудование было задействовано.

Отличие предложенной схемы научного процесса от традиционных схем (например, [How science works, 2014]) в том, что здесь представлена связь между скрытыми элементами научного процесса и артефактами научного и технологического процессов. Особое внимание уделено аспекту измеримости и задержкам в получении результатов. Данная схема акцентирует внимание на том, что от момента появления идеи до публикации первых результатов может пройти много времени (по разным оценкам в среднем от года до трех лет). В течение этого времени высказанная идея не столько оформляется, сколько проходит проверку в процессе экспериментов, внутренних обсуждений и обсуждений с привлечением коллег (например, на конференциях) и т. д. Как следствие, к моменту публикации идея не только оформлена должным образом, но и вплетена в систему представлений коллег об исследуемой области. Раннее ознакомление с идеями редуцирует отторжение нового в научном сообществе, например, обсуждаемый в разделе 3.3 эффект Матфея скорее свидетельствует о том, что если ученые представляют свои идеи на уровне страны, не вовлекая иностранных коллег в обсуждение своих новых результатов, то доверие

к этим результатам меньше. Пояснения многих понятий, представленных на диаграмме, будут даны в следующих главах.

Рассмотренные в главе два «полярных» подхода к оценке научной результативности, а именно экспертное оценивание и опора на наукометрию, не являются взаимоисключающими. Каждый из них обладает своими преимуществами. Только объединив оба подхода в рамках процесса принятия решения и вовлекая в процесс все заинтересованные стороны, при условии максимальной прозрачности методов, которыми получены оценки, можно достичь результата, способствующего развитию науки и технологии для общественного блага. Сложившиеся в странах с развитой наукой механизмы конкурсного распределения ресурсов на научные исследования обладают достаточной степенью свободы для поиска нового, а связь с практикой через определение приоритетов настраивает процесс на достижение полезных результатов для всего общества. Наукометрические инструменты являются в данном случае средством повышения качества принимаемых решений, но не могут стать единственным определяющим критерием. Иначе вместо системы развития научного знания мы получим систему, в которой «ученые» преследуют цели достижения измеримых показателей, а не создания новых знаний.

Горькавый Н. Н. don_beaver – HACA не ценит индекс Хирша. 2014. URL: <http://don-beaver.livejournal.com/137590.html> (дата обращения: 11.06.2014).

Дэвенпорт Т., Ким Д. Х. О чем говорят цифры. Как понимать и использовать данные. М.: Манн, Иванов и Фербер, 2014. 224 с.

Ефимов А. Н. Элитные группы, их возникновение и эволюция // Знание – сила. 1988. № 1. С. 56–64.

Игра в цифры, или как теперь оценивают труд ученого (сборник статей о библиометрике). М.: Изд-во МЦНМО, 2011. 72 с.

Инструменты McKinsey. Лучшая практика решения бизнес-проблем / И. М. Расиел, П. Н. Фрига. М.: Манн, Иванов и Фербер, 2009. 224 с.

Кожевников В. Л., Поляков Е. В. Сравнительная эффективность научного труда в некоторых национальных академиях наук // Вестник УрО РАН. 2010. № 4(34). С. 3–8.

- Латур Б. Наука в действии: Прагматический поворот. СПб.: Изд-во Европ. ун-та в Санкт-Петербурге, 2013. 416 p. (Latour B. Science in Action: How to Follow Scientists and Engineers Through Society. Cambridge, Mass.: Harvard University Press, 1988. 288 p.)
- Ловушки мышления. Как принимать решения, о которых вы не пожалеете / Д. Хиз, Ч. Хиз. М.: Манн, Иванов и Фербер, 2014. 336 с.
- Медоуз Д. Х. Азбука системного мышления. М.: БИНОМ. Лаборатория знаний, 2010. 344 с.
- Мокир Дж. Дары Афины. Исторические истоки экономики. М.: Изд-во Института Гайдара, 2012. 408 с. (Mokyr J. The Gifts of Athena: Historical Origins of the Knowledge Economy. Princeton N.J.: Princeton University Press, 2004. 384 p.)
- Научный потенциал вузов и научных организаций Министерства образования и науки Российской Федерации. 2013 / под ред. О. А. Ладного. СПб.: СПбГЭТУ, 2014. 210 с.
- Нейгебауэр О. Точные науки в древности. М.: Наука, 1968. 224 с.
- Расчет на одного ученого / В. Фигнер, М. Акоев, Е. Ульянова. 2012. № 27 (518). URL: <http://expert.ru/ural/2012/27/raschet-na-odnogo-uchenogo/> (дата обращения: 10.06.2014).
- Розовски Г. Университет. Пособие для владельца М.: Иерусалим Мосты Культуры, Гешарим, 1995. 414 с.
- Соколов М. Как управляют научной продуктивностью: лекция. Прочитана 25 ноября 2010 г. в рамках проекта «Публичные лекции Полит.ру». 2011. URL: <http://polit.ru/article/2011/03/05/sokolov/> (дата обращения: 10.06.2014).
- Сообщение idel song 13 ноября 2009. URL: <http://ivanov-petrov.livejournal.com/1314669.html?thread=63853933> (дата обращения: 24.06.2014).
- Хаббард Д. Как измерить все, что угодно. Оценка стоимости нематериального в бизнесе М.: Олимп-Бизнес, 2009. 320 с.
- Яблонский А. И. Модели и методы исследования науки: Философия России XX века. М.: Едиториал УРСС, 2001. 400 с.
- Collins R. The Sociology of Philosophies: A Global Theory of Intellectual Change Belknap Press of Harvard University Press, 2000. 1098 p. (Русский перевод: Коллинз Р. Социология философий: глобальная теория интеллектуального изменения. Новосибирск: Сибирский хронограф, 2002. 1280 с.)
- How science works: The flowchart / University of California Museum of Paleontology. 2014. URL: <http://undsci.berkeley.edu/article/scienceflowchart> (дата обращения: 24.06.2014).
- Institutional memory and reverse smuggling [анонимный текст]. URL: <http://wrttn.in/04af1a04.12.2011> (перевод на русский язык: URL: <http://ahitech.livejournal.com/171492.html>)
- Kainic acid // Wikipedia, the free encyclopedia. 2014. URL: https://en.wikipedia.org/wiki/Kainic_acid
- Klarreich E. Unheralded Mathematician Bridges the Prime Gap // Quanta Magazine. 2013. URL: <http://www.quantamagazine.org/20130519-unheraldedmathematician-bridges-the-prime-gap/> (дата обращения: 13.06.2014).
- Koen B. V. Discussion of the Method: Conducting the Engineer's Approach to Problem Solving-New York : Oxford University Press, 2003. 276 p.
- Less Wrong is a community blog devoted to refining the art of human rationality. 2006. URL: <http://lesswrong.com/> (дата обращения: 13.06.2014). [Русские переводы части материалов доступны по адресу: URL: <http://lesswrong.ru/>]
- Nitta I. Watase H. Tomiie Y. Structure of Kainic Acid and its Isomer, Allokainic Acid // Nature. 1958. Vol. 181. № 4611. P. 761–762.
- OECD «Main Science and Technology Indicators», OECD Science, Technology and R&D Statistics (database). [2014]. DOI: 10.1787/data-00182-en (дата обращения: 12.06.2014).
- OECD «Research and Development Statistics: Gross domestic expenditure on R-D by sector of performance and field of science», OECD Science, Technology and R&D Statistics (database). [2012]. DOI: 10.1787/data-00187-en (дата обращения 28.04. 2014).
- Price D. Little Science, Big Science and beyond. New York : Columbia University Press, 1986. 301 p. (Перевод первого издания 1963 года на русский язык: Прайс Д. Наука о науке // Наука о науке. М.: Прогресс, 1966. С. 236–254 ; Он же. Малая наука, большая наука // Наука о науке. М.: Прогресс, 1966. С. 281–384.)
- Price D. Science Since Babylon: Enlarged Edition. New Haven: Yale University Press, 1975. 232 p.
- Technology readiness level // Wikipedia, the free encyclopedia. 2014. URL: http://en.wikipedia.org/wiki/Technology_readiness_level (дата обращения: 13.06.2014).

БИБЛИОМЕТРИЧЕСКИЕ ИНДИКАТОРЫ В РЕСУРСАХ THOMSON REUTERS^{1, 2}



В. В. Писляков
Заместитель
директора
библиотеки
Национального
исследовательско-
го университета
«Высшая школа
экономики»,
кандидат физико-
математических
наук.

BIBLIOMETRIC INDICATORS IN THE THOMSON REUTERS INFORMATION RESOURCES

DOI 10.15826/B978-5-7996-1352-5.0005

This chapter examines bibliometric indicators related to citedness of journals, authors, research groups, institutions and whole countries. The introductory section deals with the basics of bibliometric analysis and features of citation databases. The author discusses the usage of various bibliometric indicators: the impact factor, average citedness, share of uncited papers, Eigenfactor and Article Influence Scores, Hirsch and Hirsch-type indices, and others. A special section investigates indicators of chronological distribution of references. Particular attention is paid to normalized indicators, including indicators normalized by research disciplines, as well as by publication sources. The final section emphasizes the importance of informed and reasonable use of bibliometric indicators in research policy-making, funding allocation, and faculty and research personnel recruitment.

В главе рассматриваются библиометрические индикаторы, оценивающие цитируемость журналов, авторов, научных коллективов, организаций и целых стран. Дается определение и обсуждается использование импакт-фактора и его вариаций, относительной цитируемости, ранговых метрик, коэффициента нецитируемости, «взвешенных» индикаторов (собственный фактор, индекс влияния статьи), индекса Хирша и ряда «Хирш-подобных» показателей и др. Специальный раздел посвящен показателям, характеризующим хронологическое распределение библиографических ссылок. Особенное внимание уделено индикаторам нормализованной цитируемости, с нормализацией по областям науки и по журналам, а также их совокупному анализу. Глава содержит вводную часть, в которой излагаются основы библиометрического анализа и особенности баз данных научного цитирования. Финальный раздел подчеркивает необходимость грамотной и аккуратной трактовки библиометрических индикаторов при принятии административных решений, распределении грантов, осуществлении кадровой политики.

ВВЕДЕНИЕ

Вданной главе мы рассмотрим ряд библиометрических индикаторов, которые либо со-

держатся в готовом виде, либо могут быть рассчитаны при помощи баз данных научного цитирования компании Thomson Reuters. Все

¹Вданной главе использованы материалы издания: Писляков В. В. Библиометрические индикаторы: практикум. М.: НФПК, Инфра-М, 2014.

²Глава публикуется с сохранением авторской пунктуации.

эти индикаторы связаны с цитируемостью научных публикаций³.

Для каждого показателя в доступной форме будет представлен алгоритм его расчета и объяснен «физический смысл»: что именно он измеряет, когда может быть применен (а когда — нет), какие существуют потенциальные опасности и ограничения у предложенных методик.

Структура настоящей главы: в вводном разделе 2.1 рассказывается, что содержится в базах данных научного цитирования, каково наполнение и функции инструментов, предоставляемых Thomson Reuters, какие существуют обозначения, терминологические и методологические договоренности в литературе по библиометрии. Раздел 2.2 посвящен импакт-индикаторам, измеряющим среднюю цитируемость в расчете на одну статью; здесь же затрагивается вопрос о самоцитировании в научной литературе. В разделе 2.3 идет речь о подходах к кросс-дисциплинарному сравнению библиометрических характеристик — прежде всего, при помощи нормализованных индикаторов с различными способами нормализации, а также с использованием ранговых методов. В разделе 2.4 вводится понятие о показателях «экстремальной» (предельной) цитируемости, к которым отнесены доля высокоцитируемых статей и коэффициент нецитируемости. «Взвешенные» индикаторы, учитывающие при подсчете цитируемости научный уровень цитируемых журналов, освещаются в разделе 2.5. Индексу Хирша

и некоторым его модификациям посвящен раздел 2.6. В разделе 2.7 исследуются показатели, характеризующие хронологическую структуру распределения ссылок и, наконец, Заключение завершает данную главу.

2.1. Библиометрические инструменты. База данных Web of Science Core Collection и аналитические надстройки

Начнем мы с описания инструмента Web of Science Core Collection (Web of Science CC), продукта компании Thomson Reuters. Web of Science CC относится к «библиометрическим базам данных» или, как еще говорят, «базам данных (научного) цитирования» (citation database, citation index). Web of Science CC размещается на более широкой платформе Web of Science, включающей, помимо Web of Science CC, специализированные базы данных по патентам, ряд тематических реферативных баз данных, национальные (страновые) базы научного цитирования и др.

2.1.1. Принципы организации библиометрических баз данных.

Методологические замечания

Из чего состоят базы научного цитирования, и что именно можно в них найти? Основной контент библиометрических баз данных — научные журналы. При этом в базе данных не содержится полного текста статей этих журналов. О каждой статье хранится (и может быть выдана пользователю), как правило, следующая информация:

³Необходимо заметить, что английское «citation», как справедливо указывалось, например, А. В. Полетаевым или Р. С. Гиляревским, означает не «цитирование» или тем более «цитату», а лишь библиографическую ссылку. При строгом подходе следует говорить не о цитируемости, а о числе полученных ссылок, «citation index» переводить как «указатель ссылок» и т. д. Тем не менее в силу ограниченности языка, необходимости использовать синонимы и трудноискренимой практики, уже сложившейся в русскоязычной литературе, в рамках настоящей главы «цитирование» будет эквивалентом «ссылке». При этом, конечно, не подразумевается, что если одна статья «цитирует» другую, то в первой содержится выдержка, фрагмент из второй (то, что обозначается английским словом «quotation»). Речь всего лишь о том, что вторая статья фигурирует в списке использованной литературы первой статьи.

— библиографические сведения о статье («выходные данные»: автор(ы), название статьи, название журнала, год выхода, том, номер, страницы);

— аннотация статьи (реферат) — в том случае если она имела в исходном тексте публикации; базы цитирования не составляют аннотации тех произведений, в которых они изначально отсутствуют;

— ключевые слова; иногда это несколько наборов ключевых слов — например, слова, приписанные статье в оригинальном

тексте публикации («авторские ключевые слова»); ключевые слова, «назначенные» базой данных цитирования исходя из ее внутреннего тезауруса ключевых слов и автоматических алгоритмов; или ключевые слова, заимствованные из тематических реферативных баз данных, где есть та же самая статья;

— тематика (рубрика), приписанная статье, и тип публикации (см. далее);

— организации, в которых работают авторы, т. н. «аффилиации», с почтовыми

UvrD facilitates DNA repair by pulling RNA polymerase backwards

By: Epshtein, V (Epshtein, Vitaly)^{1,1}; Kamarthapu, V (Kamarthapu, Venu)^{1,2,1}; McGary, K (McGary, Katelyn)^{1,1}; Svetlov, V (Svetlov, Vladimir)^{1,1}; Ueberheide, B (Ueberheide, Beatrix)^{1,1}; Proshkin, S (Proshkin, Sergey)^{3,1}; Mironov, A (Mironov, Alexander)^{3,4,1}; Nudler, E (Nudler, Evgeny)^{1,2,1}

NATURE

Volume: 505 Issue: 7483 Pages: 372+

DOI: 10.1038/nature12928

Published: JAN 16 2014

[View Journal Information](#)

Abstract

UvrD helicase is required for nucleotide excision repair, although its role in this process is not well defined. Here we show that *Escherichia coli* UvrD binds RNA polymerase during transcription elongation and, using its helicase/translocase activity, forces RNA polymerase to slide backward along DNA. By inducing backtracking, UvrD exposes DNA lesions shielded by blocked RNA polymerase, allowing nucleotide excision repair enzymes to gain access to sites of damage. Our results establish UvrD as a bona fide transcription elongation factor that contributes to genomic integrity by resolving conflicts between transcription and DNA repair complexes. Furthermore, we show that the elongation factor NusA cooperates with UvrD in coupling transcription to DNA repair by promoting backtracking and recruiting nucleotide excision repair enzymes to exposed lesions. Because backtracking is a shared feature of all cellular RNA polymerases, we propose that this mechanism enables RNA polymerases to function as global DNA damage scanners in bacteria and eukaryotes.

Keywords

KeyWords Plus: NUCLEOTIDE EXCISION-REPAIR; ESCHERICHIA-COLI-UVRD; TRANSCRIPTION ELONGATION; HELICASE-II; GENOME INSTABILITY; GENE-PRODUCT; IN-VITRO; E. COLI; DAMAGE; PROTEIN

Author Information

Reprint Address: Nudler, E (reprint author)

+ NYU, Sch Med, Dept Biochem & Mol Pharmacol, New York, NY 10016 USA.

Addresses:

+ [1] NYU, Sch Med, Dept Biochem & Mol Pharmacol, New York, NY 10016 USA.

+ [2] NYU, Sch Med, Howard Hughes Med Inst, New York, NY 10016 USA.

[3] State Res Inst Genet & Select Ind Microorganisms, Moscow 117545, Russia

+ [4] Russian Acad Sci, Engelhardt Inst Mol Biol, Moscow 119991, Russia

E-mail Addresses: evgeny.nudler@nyumc.org

Рис. 7. Запись в библиометрической базе данных на статью из журнала Nature.
Скриншот интерфейса базы данных Web of Science CC

адресами мест работы и, иногда, электронными адресами авторов;

— список цитируемой в статье литературы — именно это поле является ключевым для базы данных и делает ее «библиометрической» базой данных;

— различные второстепенные поля: номер ISSN журнала, язык оригинального документа, название и адрес издательства и др.

Таким образом, если журнал «расписывается» (индексируется) библиометрической базой данных, это означает, что о каждой публикации этого журнала мы сможем узнать из базы данных сведения, перечисленные выше. Полного текста статей при этом в базе не будет. В качестве иллюстрации на рис. 7 приведен пример библиографической записи Web of Science CC на некую журнальную статью, а на рис. 8 — список

Cited References: 54
(from Web of Science Core Collection)
From: UvrD facilitates DNA repair by pulling RNA polymerase backwards ...[More](#)

☐ Select Page [Save to EndNote online](#) ▾ [Add to Marked List](#)

☐ 1.

A physical interaction of UvrD with nucleotide excision repair protein UvrB
By: Ahn, B
MOLECULES AND CELLS Volume: 10 Issue: 5 Pages: 592-597 Published: OCT 31 2000
 [View Abstract](#)

☐ 2.

TRANSCRIPTIONAL CONTROL OF THE UVRD GENE OF ESCHERICHIA-COLI
By: ARTHUR, HM; EASTLAKE, PB
GENE Volume: 25 Issue: 2-3 Pages: 309-316 Published: 1983

☐ 3.

A ratchet mechanism of transcription elongation and its control
By: Bar-Nahum, G; Epshtein, V; Ruckenstein, AE; et al.
CELL Volume: 120 Issue: 2 Pages: 183-193 Published: JAN 28 2005
 [View Abstract](#)

☐ 4.

Damage recognition in nucleotide excision repair of DNA
By: Batty, DP; Wood, RD
GENE Volume: 241 Issue: 2 Pages: 193-204 Published: JAN 11 2000
 [View Abstract](#)

☐ 5.

TRANSCRIPT CLEAVAGE FACTORS FROM ESCHERICHIA-COLI
By: BORUKHOV, S; SAGITOV, V; GOLDFARB, A
CELL Volume: 72 Issue: 3 Pages: 459-466 Published: FEB 12 1993
 [View Abstract](#)

Рис. 8. Список цитируемой в статье литературы.
Скриншот интерфейса базы данных Web of Science CC

литературы этой статьи (он открывается при нажатии соответствующей ссылки).

Помимо журнального контента, библиометрические базы данных могут включать в себя также некоторое количество трудов конференций и, с недавнего времени, книг. При этом поля, попадающие в базу данных, аналогичны тем, которые перечислены для журналов. Только в случае трудов конференций структурной единицей будет не статья, а доклад на конференции; в случае книг — глава из книги. В случае сборников трудов конференций речь идет именно о сборниках текстов докладов, а не о кратких тезисах. Обычно текст такого доклада занимает от 4 страниц и более и в обязательном порядке имеет список использованной литературы. За исключением своего предназначения, он ничем не отличается от статьи в журнале (и впоследствии часто становится статьей). В серьезных научных монографиях также имеется список литературы по главам. Это дает возможность базам данных трактовать доклады и главы в полном подобии статьям и индексировать эти документы на тех же принципах (но также без размещения полного текста).

В базах научного цитирования каждый журнал имеет «тематическую привязку», он отнесен к той или иной научной дисциплине (может быть отнесен более чем к одной). При этом при расчете практически всех индикаторов, о которых пойдет речь в настоящей главе, тематическая рубрика статьи определяется в базе данных цитирования на основании тематической рубрики журнала, в котором она опубликована. Все статьи одного журнала имеют одну и ту же рубрику (рубрики). Мы будем говорить, что статья/журнал относятся к той или иной дисциплине или научной

области. Два последних понятия используются нами как синонимы.

Помимо дисциплины, каждой публикации в библиометрических базах данных присваивается тип документа. Это может быть научная статья (Article), научный обзор (Review), заметка редактора (Editorial), письмо (Letter), книжная рецензия (Book Review) и др. При расчете ряда библиометрических индикаторов могут учитываться не все, а лишь некоторые типы публикаций — чаще всего это Article и Review. Кроме того, замечено, что разные типы публикаций обычно получают разное число ссылок. Так, документы типа Review, «научный обзор», в среднем цитируются более активно, чем типа Article, «научная статья». Тем не менее словом «статья» мы будем обозначать, как часто это делается в обыденной практике, любую публикацию в журнале. В тех разделах, где будет важно отличать «публикацию» от «статьи» (т. е. от публикации типа «научная статья»), мы будем прибегать к английскому «Article» — так избежим двусмысленности.

В библиометрии статью считают публикацией той или иной лаборатории/организации/страны на основании институциональных адресов, которые указаны авторами в статье и, соответственно, перенесены в поле «address» (или «affiliation» и т. д.) базы данных. В данном случае при анализе не интересуются деталями авторской биографии — как он менял места работы, одновременно совмещал их и т. д. «То, что указано в статье» — это окончательный ответ на вопрос, в какой организации работал автор, ее написавший. При этом автор, конечно, может указать в статье сразу несколько мест работы — и в таком виде эта информация будет перенесена в базы цитирования.

Про отнесение публикаций к авторам, организациям и странам следует пояснить еще один момент, связанный с соавторством. Существует несколько способов счета статей, написанных несколькими авторами. Самый простой из них — «полный счет», whole (или total) counting, который предполагает, что каждому из соавторов засчитывается по одной статье. Написал автор статью в одиночку или в сотрудничестве с 20 коллегами — это все равно плюс одна публикация в его статистике. То же самое для организаций и стран: если организация (страна) хотя бы один раз указана в списке мест работы авторов статьи, эта публикация считается как целиком принадлежащая ей (и всем остальным организациям/странам соавторов наравне).

Помимо этого есть несколько способов «дробного счета», fractional counting. Если авторов в статье n , то считается, что каждый написал $1/n$ статьи, и сложением именно таких долей находится суммарная публикационная активность автора по всем его работам. То же самое для организаций: каждая статья распределяется по всем организациям в равных долях или в долях, пропорциональных числу авторов из соответствующей организации. Дробный счет применяется как для публикаций, так и для цитирований, которые они получили.

Во всех индикаторах, которые будут рассмотрены в настоящей главе, будет использоваться только метод whole counting. Его отличает простота, однако следует помнить, что он не обеспечивает аддитивность большинства показателей. Например, отвечая на вопрос «сколько всего статей написали авторы организации», нельзя приводить арифметическую сумму числа статей, написанных каждым автором этой организации. Если два или более ученых

из данного института написали статью в соавторстве, то каждый из них получит по одной статье в свою публикационную статистику. Таким образом, простая сумма по индивидуальным показателям авторов даст двойку или больше, в то время как для статистики организации как целого это, очевидно, одна статья. Кроме того, даже если автор один, но он указал, например, два места работы в двух разных странах (это вполне возможно), то такая публикация при использовании whole counting одновременно поднимет на единицу статистику двух разных лабораторий, двух разных организаций и двух разных стран.

2.1.2. Web of Science Core Collection

Итак, Web of Science CC является базой данных цитирования, «библиометрической» базой данных. Она содержит семь блоков (или «изданий», editions) и в сумме охватывает более 12 000 научных журналов, 12 000 серий трудов конференций и более 50 000 книг. Web of Science CC является ядром библиометрических продуктов Thomson Reuters, их главной частью. Блоки, из которых состоит Web of Science CC, содержат различные типы документов по различным разделам науки (в самом широком смысле этого слова):

- Science Citation Index Expanded (SCIE) — журналы по естественным, техническим, медицинским наукам (физика, химия, математика, биология, информатика и вычислительная техника, медицина и т. д.); архив с 1898 г.
- Social Sciences Citation Index (SSCI) — журналы по общественным наукам (экономика, менеджмент, социология, право, политология, история и т. д.); архив с 1898 г.
- Arts & Humanities Citation Index (A&HCI) — журналы по гуманитарным наукам (исто-

рия, литературоведение, искусствоведение, религиоведение и т. д.); архив с 1975 г.

- Conference Proceedings Citation Index – Science — труды конференций, посвященных естественным, техническим, медицинским наукам; архив с 1990 г.
- Conference Proceedings Citation Index – Social Science & Humanities — труды конференций, посвященных общественным и гуманитарным наукам; архив с 1990 г.
- Book Citation Index – Science — книги, посвященные естественным, техническим, медицинским наукам; архив с 2005 г.
- Book Citation Index — Social Sciences & Humanities — книги, посвященные общественным и гуманитарным наукам; архив с 2005 г.

В журнальных блоках есть издания, дублирующиеся в разных индексах, — речь идет о мультидисциплинарных журналах, относящихся к различным типам наук. Например, журнал *Scientometrics* отнесен к рубрике «Information Science & Library Science» в SSCI и к «Computer Science Multidisciplinary Applications» в SCIE.

В настоящее время Web of Science CC индексирует около 150 российских изданий в SCIE, 3 журнала в SSCI и 5 — в A&HCI.

2.1.3. Аналитические надстройки:

Journal Citation Reports, Essential Science Indicators, InCites

Повторим, что для каждой публикации в Web of Science CC можно узнать, какие документы она цитирует и, наоборот, какие документы цитируют ее. Следовательно, можно собирать статистику цитируемости. Для комплексных задач сделать это нелегко, поэтому Web of Science CC дополняет аналитические инструменты, в частности базы данных Journal Citation Reports (JCR), Essential Science Indicators (ESI) и InCites. Все

эти продукты используют Web of Science CC как свою основу и лишь собирают, агрегируют из Web of Science CC данные по различным «информационным единицам» (журналам, ученым, организациям, странам и т. д.).

JCR — база данных по библиометрическим показателям журналов как целого. Публикуются данные о количестве выходящих в журнале статей, числе полученных журналом ссылок, хронологическом распределении сделанных/полученных ссылок, импакт-факторе журнала (см. далее п. 2.2.1) и др.

ESI — база данных по библиометрическим индикаторам авторов, организаций, стран, журналов. Публикуются данные о количестве вышедших у автора/организации/страны/журнала статей и их цитируемости (по журналам информация в ESI менее подробная, чем в JCR). Ограничения: в ESI включены только те организации и авторы, которые попали в 1 % наиболее цитируемых хотя бы в одной научной дисциплине; и только те журналы и страны, которые попали в 50 % наиболее цитируемых хотя бы в одной научной дисциплине. Помимо этого в ESI имеется специальный раздел по мировым высокоцитируемым статьям (попавшим в 1 % самых цитируемых среди тех, которые вышли в заданном году в заданной научной области, см. далее п. 2.4.1), среднемировым библиометрическим показателям («baselines») и перспективным научным «фронтам», которые определяются специальной библиометрической процедурой (об этом подробнее см. на с. 224).

InCites — аналитический инструмент, позволяющий проводить детальный и глубокий анализ библиометрических показателей организаций, отдельных ученых, стран. Единственный из перечисленных продуктов использует нормализацию цитируемости по областям науки и по журналам. Во

второй половине 2014 г. вышла новая, расширенная версия продукта и, поскольку изменения коснулись не только интерфейса, но и методологической основы используемых индикаторов, мы будем обозначать две версии как InCites1 и InCites2.

Подчеркнем еще раз, что все данные, приведенные в JCR, ESI и InCites, рассчитаны по Web of Science CC и, строго говоря, содержатся в Web of Science CC. Однако вручную их крайне сложно оттуда извлечь, поэтому готовую агрегированную информацию по журналам, авторам, организациям и странам приводят для удобного использования в специальных аналитических «надстройках» JCR, ESI и InCites.

2.2. Индикаторы влиятельности статей (импакт-индикаторы)

К этому наиболее распространенному классу индикаторов цитируемости относятся показатели, оценивающие число ссылок, полученных в среднем одной статьей, входящей в некоторое заданное множество публикаций. Это могут быть статьи из какого-либо журнала или работы определенного автора, исследовательского коллектива, организации, целой страны и т. д.

Чаще всего исследуется множество статей, опубликованных за определенный фиксированный промежуток времени. Например, это могут быть статьи, написанные в некотором исследуемом университете U за пятилетний промежуток 2009–2013 гг. Этот интервал, в течение которого выходили оцениваемые статьи, называется «публикационным окном» (publication window). Не менее важен временной интервал, в течение которого выходили те статьи, ссылки из которых мы учитываем при подсчете импакт-индикатора. Например, это могут быть статьи, вышедшие в 2013 г. Этот промежуток

времени называется «окном цитирования» (citation window).

При подсчете импакт-индикаторов (и большинства библиометрических индикаторов в принципе) необходимо четкое обозначение как публикационного окна, так и окна цитирования, — в противном случае методика, определяющая измеряемый индикатор, будет неполна, а корректный и однозначный подсчет показателя будет невозможен. Итак, в приведенном нами примере мы бы считали среднее число ссылок, присутствующих в статьях 2013 г. выхода, на работы, опубликованные в 2009–2013 гг. учеными, работающими в вузе U. «Среднее» обозначает здесь усреднение по числу оцениваемых статей: найденное абсолютное число ссылок должно быть поделено на число статей в публикационном окне (т. е. на число статей, опубликованных учеными вуза U в 2009–2013 гг.).

Рассмотрение импакт-индикаторов мы начнем с журнальных показателей, поскольку впервые они были введены именно для журналов и на журнальной литературе легче всего проиллюстрировать их закономерности.

2.2.1. Импакт-фактор журнала

Как сказано выше, импакт-индикаторы являются наиболее распространенными среди показателей цитируемости. А самым известным и широко используемым среди них является импакт-фактор журнала (journal impact factor). При его подсчете используется двухлетнее публикационное окно и однолетнее окно цитирования.

Импакт-фактор журнала изменяется из года в год, поэтому рассчитывается для конкретного (отчетного) года. Для года Y импакт-фактор журнала равен отношению числа всех ссылок, полученных в году

У статьями данного журнала, вышедшими в годах $Y-1$ и $Y-2$, к числу этих статей (т. е. к числу статей журнала, вышедших в годах $Y-1$ и $Y-2$). Таким образом, публикационное окно — два года [$Y-1$; $Y-2$] (оценивается средняя цитируемость этого множества статей журнала), окно цитирования — один год [Y] (учитываются цитирования, сделанные в этом году).

Другими словами, импакт-фактор характеризует среднее число ссылок, полученных в отчетном году статьями журнала, опубликованными в течение двух предыдущих лет. Проиллюстрируем это конкретным примером. В 2013 г. во всех журналах, охваченных Web of Science CC, было сделано 125 ссылок на статьи журнала «Успехи химии», опубликованные в 2012 г., и 154 ссылки на его статьи, опубликованные в 2011 г. Итого статьи «Успехов», вышедшие в 2011–2012 гг., получили в 2013 г. $125 + 154 = 279$ цитирований. При этом в 2012 г. журнал опубликовал 54 статьи, в 2011 г. тоже 54, итого за промежуток 2011–2012 гг. (публикационное окно)

вышло 108 статей. Таким образом, импакт-фактор журнала «Успехи химии» в 2013 г. равен отношению $279 : 108 = 2,58$.

Все эти данные (как итоговый импакт-фактор, так и слагаемые его числителя и знаменателя) можно найти в базе данных JCR. Там они публикуются ежегодно, данные за очередной год обычно появляются в июне-июле следующего года (т. е. импакт-факторы 2014 г. можно узнать летом 2015 г. и т. д.). На рис. 9 приведен скриншот базы данных JCR с журналами по акустике, которые упорядочены по убыванию импакт-фактора. Кроме того, некоторые издательства размещают текущие значения импакт-факторов своих журналов на их веб-страницах, в открытом доступе. Нередко они сопровождаются ранговыми показателями (см. далее, п. 2.3.4).

На данный момент (конец 2014 г., последнее вышедшее издание — JCR-2013) максимальный импакт-фактор имеет журнал CA — A Cancer Journal for Clinicians, $IF = 162,5$. Это совершенно невероятный показатель (журнал публикует малое число статей, но получает много ссылок), и ближайший «конкурент», журнал New England Journal of Medicine, отстает от него ровно в три раза, $IF = 54,4$. Как и многие другие распределения в библиометрии, распределение журналов по импакт-фактору сильно сдвинуто относительно нормального: существует несколько журналов-лидеров и обширное множество изданий с низким импактом. В итоге медиана распределения импакт-факторов в журналах, входящих в SCIE (т. е. журналов по естественным, техническим, медицинским наукам), равна 1,41. То есть среди всех 8474 журналов половина имеет импакт-фактор выше 1,41, а половина — ниже.

Обратим внимание: JCR не публикует импакт-факторы гуманитарных журналов (т. е.

	Full Journal Title	Total Cites	Journal Impact Factor ▼
1	ULTRASCHALL IN DER MEDIZIN	1,332	4.645
2	JOURNAL OF VIBRATION AND CONTROL	3,288	4.355
3	ULTRASONICS SONOCHEMISTRY	5,796	3.816
4	ULTRASOUND IN OBSTETRICS & GYNECOLOGY	8,105	3.140
5	IEEE Transactions on Audio Speech and Language Processing	3,492	2.625
6	ULTRASOUND IN MEDICINE AND BIOLOGY	7,532	2.099
7	JOURNAL OF SOUND AND VIBRATION	21,330	1.857

Рис. 9. Журналы по акустике, в порядке убывания импакт-фактора. Скриншот интерфейса базы данных JCR (Thomson Reuters)

тех, которые входят только в базу A&HCI). Факторов, обуславливающих это, несколько. Прежде всего, в гуманитарных науках низкий уровень цитируемости статей, в том числе из-за того, что большую роль играет книжная литература, ее относительная роль выше, чем в общественных и, тем более, естественных/технических/медицинских науках. Кроме того, в большей мере цитируются «старые» материалы, опубликованные вне рамок двухлетнего окна цитирования, используемого при подсчете импакт-фактора. Все это делает практически невозможным сравнение научного уровня гуманитарных журналов при помощи импакта. Чтобы не вводить ученых и издателей в заблуждение, Thomson Reuters следует политике вообще не публиковать импакт-факторы для изданий по гуманитарным наукам.

2.2.2. Пятилетний импакт-фактор журнала, индекс оперативности

Когда говорят о «классическом» импакт-факторе, или «гарфилдовском» импакт-факторе, или просто об импакт-факторе без уточнений, то имеют в виду показатель, введенный в п. 2.2.1. Он учитывает цитирования, которые статьи журнала получают в течение двух лет после своего выхода. Однако в ряде научных дисциплин, особенно в социальных и гуманитарных науках, профессиональное сообщество не успевает в полной мере воспринять новое знание за столь короткий срок, как два года, и целесообразно использовать показатель с более широким публикационным окном. Поэтому в базе данных JCR также публикуются значения пятилетнего импакт-фактора журналов.

Для года Y пятилетний импакт-фактор журнала равен отношению числа всех ссылок, полученных в году Y его статьями, вышедшими в годах с $Y-1$ по $Y-5$, к числу этих

статей (т. е. к числу статей журнала, вышедших в годах с $Y-1$ по $Y-5$). Таким образом, импакт-фактор характеризует среднее число ссылок, сделанных в отчетном году на статьи журнала, опубликованные в течение пяти предыдущих лет.

Если упорядочить журналы по убыванию пятилетнего импакт-фактора, результат будет отличаться от ранжирования по обычному импакту: журналы, чьи материалы устаревают медленнее, т. е. те издания, на статьи которых делается значительное число ссылок даже через пять лет после их публикации, будут иметь преимущество перед журналами, получающими основное число ссылок на материалы менее чем трехлетней давности. Известно, что скорость «старения» публикуемого журналом знания в первую очередь зависит от его дисциплинарной области (подробнее см. далее, раздел 2.7).

Максимальный пятилетний импакт-фактор для журналов из SCIE равен 107,7, медиана — 1,50. Она несколько выше медианы двухлетнего импакт-фактора (1,41), это свидетельствует о том, что пятилетний показатель более полно охватывает цитирования, полученные журналами.

Заметим теперь, что при вычислении как импакт-фактора, так и пятилетнего импакт-фактора не учитываются ссылки, сделанные на те статьи журнала, которые вышли непосредственно в отчетном году. Они «пропадают» для журнала, если мы используем только эти два показателя. Тем не менее такие цитирования встречаются и, более того, их число имеет сейчас тенденцию к увеличению по причине все более частого размещения препринтов в Интернете, открытия специальных разделов на сайтах издателей, где публикуются предварительные версии статей, принятых в печать, а также общего ускорения производствен-

ного цикла научных издательств. Показатель, фиксирующий цитирования «того же года», также публикуется в базе данных JCR и называется immediacy index. Будем называть его «индексом оперативности» (другой вариант перевода, также предлагавшийся в литературе, — «индекс немедленного цитирования»).

Индекс оперативности предполагает одностороннее публикационное окно и одностороннее окно цитирования, причем они совпадают — это отчетный год Y . Для вычисления индекса необходимо разделить число всех ссылок, полученных в году Y статьями журнала, вышедшими в том же году Y , на число этих статей. Индекс оперативности показывает, насколько быстро ученый мир реагирует на статьи журнала, как скоро воспринимает его тексты и использует их при воспроизводстве научного знания.

Самый большой индекс оперативности в JCR-2013 (у того же журнала CA — A Cancer Journal for Clinicians) — 27,76. Медиана по журналам из SCIE предсказуемо мала — 0,27. Практически для каждого журнала индекс оперативности — самый низкий показатель из трех рассмотренных импакт-индикаторов (однако встречаются и исключения, например Physics of Life Reviews или Annals of Family Medicine, где индекс оперативности, наоборот, выше двух- и пятилетнего импактов).

2.2.3. Средняя цитируемость

Наряду с импакт-фактором, рассчитываемым для журналов, можно ввести аналогичный по сути показатель для оценки средней影响力的 статей отдельного ученого, или лаборатории/факультета, или организации, или целой страны. Необходимо лишь зафиксировать, как и для любого импакт-индикатора, публикационное окно и окно

цитирования исходя из следующей логики. Если изучать только самые свежие статьи, то многие из них еще не наберут достаточное число ссылок, в полной мере отражающее их научный уровень. С другой стороны, если сделать акцент на статьях, вышедших давно, то не удастся оценить текущее состояние исследований автора/организации. Как правило, приемлемым решением является использование пятилетнего публикационного окна и совпадающего с ним окна цитирования. В этом случае считается среднее число ссылок, полученных за последние полные пять лет теми статьями, которые вышли в течение тех же последних пяти лет (в расчете на одну статью).

В ESI информация дается по последним десяти годам, плюс прошедшая часть текущего года (с задержкой в 2 месяца), но существует возможность просмотра отдельно каждого пятилетнего интервала, входящего в десятилетку.

Заметим, что для бурно развивающихся областей, где знание быстро устаревает (таких как нанотехнологии, онкология, прикладная физика), можно выбрать временной интервал меньше пяти лет; для областей, где старение знания происходит особенно долго (математика, зоология, социология), имеет смысл увеличить оба окна. В InCites есть возможность посчитать для организации или автора показатель «cites per document» (цитируемость в расчете на одну статью) на любом временном промежутке.

2.2.4. Совокупный и средневзвешенный импакт-фактор

Как следует из определения, импакт-фактор является не характеристикой журнала как целого, а показателем средней影响力的, среднего уровня одной статьи

в журнале. При этом реальный уровень статей, измеренный в полученных ими ссылках, обычно сильно разнится от статьи к статье даже в рамках одного журнала. В журнале с импакт-фактором 2,0 могут встретиться как те статьи, которые в течение двух лет после своего выхода не получили ни одной ссылки, так и те, что получили 20 ссылок (напомним: из определения импакта следует, что в среднем в таком журнале статья получает 4 ссылки за два года, следующих за годом публикации). Импакт-фактор отражает среднюю цитируемость, но не может учесть колебания вокруг среднего.

Именно истинное, реальное число ссылок, полученных статьей, отражает ее подлинную влияние, а импакт-фактор издания, в котором она опубликована, делает это лишь косвенно и недостоверно. Поэтому в п. 2.2.3 мы считали именно реальную цитируемость. Однако у методик, построенных на измерении истинного, «наблюдаемого» числа ссылок, есть серьезный недостаток: для адекватной оценки исследуемых работ необходимо, чтобы после их опубликования прошел значительный промежуток времени. Статьи должны успеть получить то число ссылок, которое отражает их реальный уровень, — в п. 2.2.3 мы рекомендовали с этой целью пятилетний интервал.

Ввиду этого как некоторая оценка именно оперативного состояния научной деятельности организации и — реже — ученого используется показатель, базирующийся на импакт-факторе тех журналов, где выходят их статьи. Такой индикатор может быть посчитан сразу после выхода соответствующей публикации (а в известном смысле даже раньше, в момент принятия статьи в печать). Корректное полное название данного индикатора — «совокупный импакт-фактор журналов, в которых опубликованы статьи

организации/ученого». Рассчитывается он за некоторый промежуток времени, для оперативного оценивания логично брать один (завершившийся) год. Показатель равен сумме импакт-факторов тех журналов, в которых публиковались статьи организации/ученого; если есть несколько статей из одного и того же журнала — соответствующее слагаемое умножается на число статей, вышедших в данном журнале.

Совокупный импакт-фактор характеризует (как и следует из его названия) деятельность организации/ученого «интегрально». Если необходимо ввести оценку в расчете на одну опубликованную статью, используют термин «средневзвешенный импакт-фактор журналов, в которых опубликованы статьи организации/ученого». Последний равен отношению совокупного импакта, который введен выше, к общему числу опубликованных за рассматриваемый промежуток времени статей.

Средневзвешенный импакт-фактор показывает средний уровень статей в тех журналах, в которых публикуется организация/автор, и в какой-то мере позволяет предсказать дальнейшую цитируемость работ организации/автора.

Следует иметь в виду, что совокупный и средневзвешенный импакт-факторы сильно зависят от области знания, в которой работает ученый / функционирует организация. Про нормализацию по областям науки речь пойдет в разделе 2.3. Если же ограничиться рамками одной дисциплины (например, исследовать деятельность отдельного ученого или узкотематической лаборатории), то полученное значение средневзвешенного импакта интересно сравнить с показателем по всем журналам данной дисциплины. Так можно определить, в «сильных» или «слабых» журналах публикуется в сред-

нем ученый/лаборатория. Для иллюстрации можно представить, что организация/автор публикуются все время в одном и том же «среднем» для себя журнале, тогда именно средневзвешенный импакт-фактор отражает место этого «среднего» журнала в иерархии научных периодических изданий соответствующей дисциплины¹⁴.

2.2.5. Роль самоцитирования.

Индикаторы, характеризующие самоцитирование

Заметим, что на все показатели, разобранные в настоящем разделе, оказывают влияние и цитирования, сделанные самими «исследуемыми объектами» на себя. В расчет числителя импакт-фактора журнала включаются в том числе ссылки, полученные им из его же статей. При оценке средней цитируемости автора или организации учтены и цитирования автором/организацией своих публикаций и т. д. В своем стандартном определении все рассмотренные в настоящем разделе показатели включают самоцитирование.

Нельзя сказать, что учет самоцитирования — это недостаток разобранных метрик или, наоборот, их сильная сторона. Необходимо знать об этой особенности методики и иметь ее в виду. Полезно сравнивать показатели, посчитанные с включенным самоцитированием и с исключенным. Например,

¹⁴Как упоминалось в п. 2.2.1, распределение импакт-факторов сильно асимметрично, с небольшим числом журналов с высокими значениями импакта и множеством изданий с низкими показателями. Это верно в той или иной мере для любой дисциплины. Часто в таких случаях более корректным показателем является медианный импакт-фактор журналов, в которых публикуется ученый/лаборатория, вместо среднего (средневзвешенного) импакта. Соответственно, сравнивать его следует с медианным значением по всей дисциплине.

в JCR для журналов приводится и импакт-фактор (классический, двухлетний), из числителя которого убраны ссылки журнала на самого себя — Impact Factor Without Journal Self Cites.

Для оценки уровня самоцитирования журналов обычно используют два показателя. В числителе обоих индикаторов — число ссылок, полученных журналом из статей, опубликованных в нем самом. Число «самоцитирований». В знаменателе первого показателя число всех ссылок, полученных журналом. Этот показатель называется коэффициентом самоцитируемости, он показывает долю во всех ссылках, полученных журналом, ссылок, полученных им из него самого. Знаменатель второго показателя равен числу всех ссылок, сделанных журналом. Он называется коэффициентом самоцитирования и показывает долю во всех цитированиях, сделанных журналом, ссылок, ведущих на него самого.

Высокий коэффициент самоцитируемости говорит о том, что журнал почти никто не цитирует, кроме него самого, и это, по заключению, например, Р. Руссо [Rousseau, 2002], свидетельствует о малой заметности журнала. Высокий коэффициент самоцитирования может означать несколько другое: журналу некого цитировать, кроме самого себя. То есть журнал относится к замкнутой, изолированной научной дисциплине. Но если при этом также высок коэффициент самоцитируемости, можно предположить наличие неэтичной установки в политике редакции на цитирование только собственных статей.

Для журналов обычно считается допустимым, еще не свидетельствующим о проблемности издания, коэффициент самоцитируемости в 30–35 %. При этом, как правило, он высок у слабых изданий, получающих мало ссылок, и низок у журналов-лидеров.

Например, для двух российских изданий, данные по которым есть на текущий момент в общественно-научной части JCR-2013, коэффициент самоцитируемости составляет 71% («Социологические исследования») и 75% («Вопросы психологии»). Эти журналы находятся внизу импакт-рейтингов в своих дисциплинах. Если взять журналы-лидеры в соответствующих рубриках JCR-2013, то для них коэффициент самоцитируемости будет 1,9% (American Sociological Review) и 2,9% (Educational Psychologist) соответственно. В частности этот эффект приводит к тому, что исключение самоцитирования при вычислении импакт-факторов журналов (т.е. учет только ссылок, полученных журналом из других изданий), как правило, слабо влияет на рейтинг ведущих журналов с высокими показателями, однако сильно меняет взаимные позиции «на дне» рейтинга, среди малоцитируемых изданий [ср. McVeigh, 2004].

Что касается самоцитирования авторов, для его оценки применяются аналогичные коэффициенты, в первую очередь коэффициент самоцитируемости. Однако в этом случае есть два различных подхода к тому, что именно считать самоцитированием. «Прямое самоцитирование» — это ссылка на публикацию, среди соавторов которой есть исследуемый ученый, появившаяся в статье, среди соавторов которой есть тот же самый ученый. Однако есть еще и самоцитирование соавторов — ссылка на публикацию, среди соавторов которой есть исследуемый ученый, появившаяся в статье, среди соавторов которой есть соавтор(ы) исходной статьи (при этом самого исследуемого ученого в авторах цитирующей публикации может не быть). Поясним подробнее: пусть ученый S написал (сам или в соавторстве) статью P1, которую цитирует статья P2. Это считается:

- прямым самоцитированием, если S — автор P2;
- самоцитированием соавторов, если среди авторов P2 есть хотя бы один из авторов P1.

Соответственно, коэффициент самоцитируемости ученого может считаться или как доля «прямого самоцитирования» во всех ссылках, которые получили его публикации, или как доля «самоцитирования соавторов» во всех ссылках, полученных его работами²⁵. Очевидно, что второй коэффициент будет всегда больше или равен первому. Вообще говоря, исключение при анализе работ ученого всех ссылок, являющихся «самоцитированием соавторов», предлагает весьма ригористичный подход: посчитаем только те цитирования, которые получены из «совсем чужих» работ — множество авторов которых никак не пересекается с множеством авторов цитируемой публикации.

В заключение заметим, что само по себе самоцитирование, конечно, не только не порочная практика, но даже неотъемлемая составляющая научной коммуникации. Журнал не может не ссылаться на свои публикации, это означало бы отсутствие всякой преемственности в его функционировании. Сложно представить себе автора, который бы не ссылался на свои предыдущие работы, — это был бы ученый либо постоянно меняющий область деятельности, либо стыдящийся, отрекающийся от своих прежних публикаций. Но нельзя и не помнить о возможных злоупотреблениях, когда самоцитирование гипертрофируется и искажает библиометрический анализ, если в нем не отслеживать данный аспект.

²⁵Заметим, что в случае с журналами такой вариативности нет по той причине, что одна и та же работа не может быть опубликована сразу в двух журналах, т.е. не бывает «журналов-соавторов».

2.3. Индикаторы относительной влиятельности статей (относительные, нормализованные импакт- индикаторы)

До сих пор мы рассматривали «абсолютные» индикаторы, которые зависят только от показателей самого исследуемого журнала/ученого/организации и не учитывают контекст, в котором те осуществляют свою исследовательскую/публикационную деятельность. Самый серьезный недостаток при этом то, что мы не учитывали, к какой научной области относятся оцениваемые статьи. Сравнение эффективности академической деятельности представителей различных наук невозможно осуществить с помощью абсолютных, «простых» индикаторов. Центральным фактом, который необходимо иметь в виду при оценке эффективности деятельности ученых или организаций, а также при оценке качества журналов является сильная зависимость абсолютных библиометрических показателей от научной дисциплины, в которой публикуются изучаемый автор/организация или выходит журнал.

Проиллюстрируем это на журналах. Цитируемость журналов серьезным образом зависит от их дисциплинарной принадлежности, поэтому медиана импакт-факторов сильно варьируется от рубрики к рубрике. Например, для клеточной биологии в JCR-2013 она составляет 3,33, а для математики — 0,58. Это означает, что половина журналов по клеточной биологии имеет импакт выше 3,33, в то время как половина журналов по математике получила импакт ниже 0,58. Более того, в категории «математика» вообще нет ни одного журнала с импакт-фактором выше 3,33 (три таких журнала появятся, если добавить категории «прикладная математика» и «математика, междисциплинарные приложения»).

Почему так происходит? Следует ли говорить о том, что математика имеет какие-то «недостатки» относительно биологических наук, какой-то внутренний изъян? Разумеется, нет. На самом деле можно указать как минимум три причины указанного эффекта:

- Различная практика цитирования, сложившаяся в журналах различных научных областей; можно сказать, варьирующаяся в зависимости от дисциплины «плотность ссылочного поля». Например, известно, что в одной статье по клеточной биологии в среднем 55 ссылок в пристатейном списке литературы, в то время как в математике — лишь 21 (JCR-2013). Поэтому «вероятность» получить ссылку у журнала по математике меньше⁶. Это индивидуальные особенности коммуникации, укоренившиеся в различных науках и никак не связанные с их «качеством» относительно друг друга.
- Разный средний «возраст» цитируемых источников. В одних дисциплинах основной массив цитирований ведет на свежую литературу (попадающую в двухлетний период, который учитывается при подсчете импакт-фактора), в других областях науки чаще цитируются источники, вышедшие давно (мы упоминали об этом в п. 2.2.2 и поговорим подробнее в разделе 2.7). Чем больше доля ссылок, ведущих на документы старше двух лет, тем меньше они поднимают импакт-факторы журналов соответствующей области науки.
- В некоторых дисциплинах цитируемая литература хорошо представлена в журнальных блоках Web of Science CC, а в некоторых встречается много ссылок на материалы, не охваченные базой данных.

⁶Разумеется, здесь мы используем приближенный подход, оставляя в стороне междисциплинарность и возможность получения ссылки «извне» дисциплины.

Например, в тех областях, где больше ссылок на книжные, а не журнальные источники; на труды конференций, диссертации, кодексы, интернет-страницы и т. д. [см. Larivière et al., 2006]. Все эти ссылки «пропадают» для импакт-фактора, даже если по своему возрасту цитируемая литература попадает в двухлетнее публикационное окно, — они также не увеличат импакт никакого журнала.

Эти дисциплинарные особенности не учитываются в абсолютном значении импакт-фактора, что делает невозможным его использование в комплексном библиометрическом исследовании. Поэтому для решения двух важных задач: а) сравнение между собой исследовательских единиц, работающих в разных областях науки; б) комплексная оценка деятельности организации, занимающейся исследованиями сразу в нескольких научных областях — вводятся относительные библиометрические индикаторы. Цель их — оценить научную деятельность изучаемых объектов в сравнении с другими однотипными объектами в той же научной области. Это сравнение «на фоне коллег» или, в обратной формулировке, «относительно конкурентов». Относительные индикаторы потребуются в том случае, если необходимо сравнить достижения химика и математика или получить интегральный показатель вуза, чьи сотрудники публикуют статьи в широком спектре научных дисциплин.

2.3.1. Относительный импакт-фактор, относительная цитируемость

Начнем с относительного показателя для журнала. Самый простой из них — отношение импакт-фактора издания к среднему импакт-фактору дисциплины, к которой журнал относится. При этом среднее по дисциплине обычно берется не как отношение суммы всех

импакт-факторов журналов данной дисциплины к числу таких журналов, а несколько иначе: используется т. н. «агрегированный импакт-фактор» (aggregate impact factor). Дисциплина в целом (т. е. совокупность статей, опубликованных во всех журналах, отнесенных к данной дисциплине) рассматривается как некий единый «метажурнал», для которого вычисляется традиционный импакт-фактор: отношение числа всех ссылок, полученных в году Y статьями данной дисциплины, вышедшими в годах $Y-1$ и $Y-2$, к числу статей данной дисциплины, вышедших в годах $Y-1$ и $Y-2$. Этот показатель для дисциплин как целого также публикуется в JCR. Именно он — «агрегированный» («собранный») импакт-фактор дисциплины — берется в качестве знаменателя, именно на него делится импакт каждого журнала, чтобы получился относительный импакт-фактор.

В целом можно сказать, что если относительный импакт-фактор журнала больше 1, это значит, что в среднем статьи журнала цитируются чаще, чем средняя статья, опубликованная в данной дисциплине, и наоборот. Следует также иметь в виду, что один и тот же журнал может быть приписан сразу к 2–3 дисциплинам и у него, соответственно, может быть посчитано несколько относительных импакт-факторов — относительно агрегированного значения в каждой из дисциплин. В таком случае с помощью описанного метода можно узнать, в какой из научных областей это издание занимает более престижную позицию: в той из дисциплин, у которой ниже агрегированный импакт-фактор, относительный импакт такого журнала, очевидно, будет выше.

Аналогичный показатель можно считать и для авторов. Назовем его для удобства «относительная цитируемость» (хотя это слишком общее название). Относительная

цитируемость ученого равна отношению средней цитируемости его статей (среднее число ссылок на одну статью) к средней цитируемости статей некоей «референтной группы», на фоне которой мы рассматриваем деятельность ученого.

Один из наиболее важных вопросов при библиометрическом анализе — что выбрать в качестве референтной группы? Это могут быть статьи организации, в которой работает ученый, или более узко — статьи его лаборатории/факультета. В первом случае, однако, следует помнить, что если организация публикуется сразу в нескольких областях науки (яркий пример: классический университет), то сопоставление скорее всего будет некорректным: если ученый работает в высокоцитируемой науке, то у него будет незаслуженное преимущество перед коллегами, если в малоцитируемой — он окажется в несправедливо невыгодном положении. Референтная группа должна, по крайней мере, публиковаться в той же научной области, что и оцениваемый автор. Кроме того, как всегда, необходимо четко задавать публикационное окно для оцениваемых статей и окно цитирования — для статей, чьи ссылки учитываются при подсчете индикатора⁴⁷.

⁴⁷При выборе публикационного окна больше одного года проявляется, однако, некоторая слабость предложенного метода. Предположим, что автор публиковал свои статьи не равномерно из года в год, а, например, по нарастающей — все больше и больше, — в то время как во всей его дисциплине выходило примерно одинаковое число статей ежегодно. Тогда автор получает незаслуженную относительную недооценку своих работ, поскольку шансы на получение ссылок у статей, выпущенных позже, меньше, чем у статей, вышедших ранее. То же самое верно для относительного импакт-фактора: если журнал ощутимо увеличил число публикуемых статей в тот год, который предшествовал отчетному, он будет проигрывать журналу, не принимавшему таких редакторских решений (ес-

Тот же метод легко обобщается на целую организацию при условии, что ученые в ней публикуются в более-менее одной тематической области. Можно определить среднюю цитируемость статьи данного вуза/института относительно средней цитируемости статьи, публикуемой в той же самой дисциплине всеми российскими учеными или всеми учеными мира. Случай, когда публикационный профиль организации более широк и затрагивает несколько тематических областей, будет рассмотрен в п. 2.3.2.

Отметим, что в качестве референтной группы также могут быть взяты все статьи того же журнала, где публикуется ученый, — в этом случае мы получим оценку, насколько именно его работы выше или ниже по цитируемости, чем средний уровень издания. Преимущество данного метода — в нем практически отсутствует необходимость следить за областью науки: все статьи в журнале чаще всего принадлежат к одной и той же дисциплине (исключение: мультидисциплинарные издания — *Nature*, *Science*, *Proceedings of the National Academy of Sciences*, «Вестник РАН» и т. д.). Если ученый публикуется сразу в нескольких журналах, что более чем естественно, необходим более комплексный подход, который мы также рассмотрим в п. 2.3.2.

Корректную референтную группу можно выбрать несколькими способами, в конце концов этот выбор зависит от задач, которые ставит перед собой специалист, проводящий библиометрическое исследование.

ли предположить, что научный уровень всех статей в обоих журналах одинаков). Эти недостатки преодолеваются при использовании более тщательного подхода, описанного в п. 2.3.2.

2.3.2. Относительная цитируемость разнородного потока публикаций («normalized citation impact», «crown indicator»)

Представим себе теперь ситуацию, максимально приближенную к реальной: необходимо оценить научный уровень статей, опубликованных организацией широкого научного профиля, например университетом.

Сначала, конечно, необходимо задать публикационное окно и окно цитирования, чтобы определить, какие именно публикации и на каком временном интервале мы оцениваем — так, как это делалось в п. 2.3.1. Здесь особых сложностей не возникает. Однако выделенный публикационный поток будет крайне неоднородным: он будет содержать публикации из разных журналов и разных тематических областей, кроме того это будут публикации разных типов (исследовательские статьи, научные обзоры, письма и т. д.), это приводит к дополнительным сложностям.

В данном случае необходимо исходить из двух основных положений, которые уже фигурировали в п. 2.3.1:

- определять «научный уровень» можно только на некотором фоне, в некотором контексте, относительно чего-то; необходимо задать референтную группу;

- сравнивать научный уровень публикаций относительно друг друга можно, только если они однотипны, т. е. все их характеристики, не имеющие прямого отношения к научному уровню, одинаковы.

Если мы анализируем позиции факультета внутри вуза, то референтной группой будет весь вуз. Если исследуется университет как целое, то в качестве референтной группы логично выбрать все российские вузовские публикации (заметим: включая публикации и самого анализируемого универ-

ситета) — если мы хотим оценить эффективность научной деятельности университета на фоне других вузов России. Можно также взять вообще все отечественные работы — так мы узнаем, как выглядит университет на фоне всей российской науки.

Однако для унификации процедуры оценки мы в дальнейшем будем считать референтной группой весь мир, т. е. все мировые публикации. Таким образом, будет оцениваться соответствие исследуемого научного коллектива общемировому уровню. Именно так построена работа в продукте InCites: можно найти показатели автора, университета или целой страны относительно среднемирового уровня⁵⁸.

На цитируемость публикации влияет ряд факторов, не имеющих непосредственного отношения к ее научному уровню:

- область науки (как мы видели, средняя цитируемость сильно зависит от научной дисциплины);

- год издания (чем раньше издана статья, тем больше ссылок она могла получить к моменту измерения);

- тип публикации (как уже говорилось в п. 2.1.1, в среднем, например, научные обзоры цитируются чаще, чем оригинальные исследовательские статьи⁶⁹).

⁵⁸Заметим, что этот подход не замещает собой оценку более компактных групп относительно друг друга. Например, если посчитаны показатели одной организации и целой страны относительно всего мира, то индикатор, показывающий эффективность деятельности этой организации на фоне этой страны, не равен отношению двух найденных показателей друг к другу.

⁶⁹Этот факт далек от чистой теории. Например, документы типа Article, опубликованные в журнале Tetrahedron в 2010 г., получили за 2010–2013 гг. (согласно Web of Science CC) в среднем по 9,0 ссылок. Тот же показатель для публикаций типа Review равен 35,2 — разница почти в 4 раза! Учет типа документа не менее важен, чем коррекция на область науки, к которой он относится.

Для учета всех этих факторов и корректного определения относительной цитируемости всего публикационного потока организации/автора вводится показатель «ожидаемое число ссылок» (expected citation rate, или baseline): для каждой отдельной статьи исследуемых организации/автора это среднее число ссылок, которое получили все мировые публикации, имеющие ту же дисциплинарную принадлежность, тот же год выхода и тот же тип документа, что и рассматриваемая статья. Если рассматриваемая статья организации/автора в действительности получила больше ссылок, чем ожидаемое для нее их число, то она превосходит средний мировой уровень «такой же, как она» публикации (т. е. имеющей те же характеристики дисциплина / год выхода / тип документа), если меньше — уступает ему.

Относительная цитируемость одной публикации, таким образом, равна отношению реального («наблюдаемого») числа полученных ею ссылок к их ожидаемому числу. Для того чтобы посчитать относительную цитируемость совокупности всех публикаций организации/автора, берется среднее арифметическое значений относительной цитируемости для каждой отдельной публикации, написанной организацией/автором. Следуя терминологии InCites2, обозначим этот показатель «Normalized Citation Impact» (NCI)¹⁰.

¹⁰В западной литературе данный или аналогичный показатель еще называют «mean normalized citation score», «normalized mean citation rate», «normalized impact» и даже «crown indicator» (подчеркивая его важность). Иногда используется альтернативный способ усреднения по всему массиву ссылок, полученных публикациями организации/автора. Сначала считается ожидаемое число ссылок сразу для всех статей (суммированием по каждой), потом на эту сумму делится суммарное наблюдаемое число ссылок, полу-

В методику расчета показателя NCI заложена чрезвычайно грамотная схема, позволяющая максимально корректно оценить относительную цитируемость любого набора статей. Однако и этот индикатор — как абсолютно любой библиометрический показатель, взятый сам по себе, — имеет свои ограничения. Поскольку он так или иначе считает средние величины, на него могут оказывать очень большое влияние отдельные высокочитируемые статьи организации/автора. Даже для целых стран, как показали Aksnes & Sivertsen [Aksnes, Sivertsen, 2004], отдельные работы, получившие очень много ссылок, могут значительно повлиять на средние страновые показатели. С не меньшей осторожностью следует применять NCI на коротких временных промежутках, например для статей 1–2-летней давности. Ввиду того что весь массив мировых публикаций получил еще недостаточно много ссылок к этому моменту, показатели ожидаемой цитируемости могут быть малопредсказуемы и не давать надежного, достоверного базиса нормализации.

Наконец, как уже говорилось ранее в сноске, NCI, полученный относительно мирового уровня, не всегда дает возможность сравнивать исследуемые единицы между собой. Как рельефный пример — если у организации $NCI = 0,5$, а у всей России $NCI = 0,67$ (последняя цифра — реальное значение из InCites2 для статей 2010–2014 гг. выхода, с тем же окном цитирования), это не обязательно означает, что организация работает слабо относительно всей российской науки. Возможно, это химический институт,

ценное публикациями организации. Этот способ используется в InCites1 (под названием «Category Actual/Expected Citations»), и именно так определяется показатель «цитируемость, нормализованная по областям науки» [Писляков, 2014].

публикующий статьи лишь в два раза слабее среднемирового уровня, в то время как позиция России в химии ощутимо ниже (0,40 от среднемирового уровня цитируемости). При этом значение $JNCI = 0,67$ для всей страны получается из-за более качественных публикаций в физике⁸¹¹.

2.3.3. Нормализация по журналам и комбинированный анализ относительной цитируемости

Подход, аналогичный представленному в п. 2.3.2, также возможен, если в качестве референтной группы брать публикации не в рамках одной и той же научной области, а в одних и тех же журналах. Тогда «ожидаемое число ссылок» для каждой публикации исследуемых организации/автора будет равно среднему числу цитирований, полученных публикациями данного журнала, которые вышли в том же году и относятся к тому же типу документа. Наблюдаемое число ссылок для статьи, естественно, остается без изменений. Необходимо взять его отношение к ожидаемому показателю и усреднить результат по всем статьям организации/автора. Полученный индикатор, вновь используя термины InCites2, назовем Journal Normalized Citation Impact (JNCI)⁹¹².

⁸¹¹Данный пример иллюстративен, но охватывает лишь часть указанной проблемы: так, найдя значение $JNCI$ для России с ограничением по химическим наукам, все равно некорректно сравнивать его с $JNCI$ отдельной организации, специализирующейся на химии: необходим комплексный подход с учетом и годов выхода статей, и типов документа. Такой показатель предусмотрен в плане развития платформы InCites.

⁹¹²Здесь вновь возможна иная нормализация, с подсчетом сразу всех ожидаемых ссылок, суммой по всем статьям, а затем деление на полученный результат суммы всех полученных в реальности ссылок. Это соответствует метрике «Journal Actual/Expected Citations» в интерфейсе InCites1 и показателю «цитируемость, нормализованная по журналам» [Писляков, 2014].

Этот индикатор характеризует то, как «выделялись» статьи данной организации на фоне всех статей тех журналов, где публиковались ее сотрудники. Были публикации ее ученых в среднем выше среднего уровня публикующих их журналов – или ниже него. Тот же самый смысл имеет $JNCI$, если мы исследуем публикации отдельного ученого или, наоборот, целой страны. Показатель, посчитанный для целых стран, называют также по-другому — «индекс Матфея». Его подсчет дает повод рассуждать о том, «дискриминируются» ли в плане цитируемости статьи ученых из разных стран в рамках одного и того же журнала [Писляков, Дьяченко, 2009; Pisyakov, Dyachenko, 2010].

Нормализация по журналам в известном смысле легче и точнее: как мы уже говорили, здесь почти не встает вопрос об определении тематической области, к которой относится публикация. В рамках одного журнала, за некоторыми исключениями (Science, Nature и т. д.), все статьи принадлежат к одной и той же дисциплине. Или, по крайней мере, эти отличия будут в среднем меньше, чем погрешность при любом другом автоматическом определении тематики публикации. Более того, если сами дисциплины приписываются статьям на основе тематической рубрикации журналов, в которых они опубликованы (как обычно делается в библиометрических базах данных), то получается, что нормализация по журналам дает более точный, детальный взгляд на сопоставление публикации, чем нормализация по областям науки.

Здесь, однако, есть ключевой момент. Предположим, что некоторый университет получает высокий показатель нормализованной по журналам цитируемости, $JNCI$. Но при этом, если мы посчитаем его цитируемость, нормализованную по областям

науки, NCI, может оказаться, что она меньше единицы. Это происходит в том случае, если авторы данного вуза выбирают для публикации своих статей слабые журналы. В рамках этих слабых журналов они могут превзойти уровень остальных статей, так что цитируемость их публикаций превысит средние показатели этих непрестижных изданий (и JNCI будет больше 1). Но из-за того что сами журналы мало цитируются на фоне других изданий в той же области науки, цитируемость статей авторов вуза не достигнет среднего уровня по дисциплине.

Конечно, возможна и обратная ситуация, когда авторы организации подают свои статьи в ведущие журналы, успешно проходят процедуру рецензирования и отбора, одна-

ко, попав туда, их публикации не достигают среднего уровня цитируемости этих изданий. В таком случае мы можем получить высокую цитируемость относительно областей науки и низкую — при нормализации по журналам.

Так или иначе очевидно, что для всестороннего исследования научной деятельности организации или автора (а также научных подразделений, целых регионов, стран и т. д.) следует использовать оба показателя — и NCI, и JNCI. Суммируем выводы, которые можно сделать при их совмещенном анализе, в таблице 3. В качестве примера будем рассматривать публикации организации, хотя аналогичные результаты можно получить, исследуя другие научные единицы.

Таблица 3

Анализ нормализованных показателей цитируемости для организации

	JNCI > 1	JNCI < 1
NCI > 1	В среднем публикации сотрудников таких организаций цитируются выше среднего уровня цитируемости как журналов, в которых они выходят, так и областей науки, к которым они относятся	В среднем публикации сотрудников таких организаций цитируются ниже среднего уровня цитируемости журналов, в которых они выходят, но при этом выше уровня областей науки, к которым они относятся. Это может быть связано с выбором сотрудниками сильных журналов для своих публикаций ¹³
NCI < 1	В среднем публикации сотрудников таких организаций цитируются выше среднего уровня цитируемости журналов, в которых они выходят, однако ниже уровня областей науки, к которым они относятся. Это может быть связано с выбором сотрудниками слабых журналов для своих публикаций	В среднем публикации сотрудников таких организаций цитируются ниже среднего уровня цитируемости как журналов, в которых они выходят, так и областей науки, к которым они относятся

¹³Осторожная формулировка «может быть связано» вызвана тем, что при нормализации, используемой в InCites2, сложно определить понятие «среднего уровня цитируемости совокупности журналов, в которых публикуются ученые организации, относительно среднего уровня цитируемости областей науки, к которым относятся эти журналы». Точнее, характеристика «силы» (научного уровня) совокупности журналов, в которых публикуются сотрудники, не определяется однозначным образом значениями NCI и JNCI для организации. При использовании другого способа нормализации [Писляков, 2014] понятие «среднего уровня совокупности журналов относительно дисциплин» может быть формализовано более четко.

В дополнение к характеристикам, представленным в таблице 3, отметим, что показатель $JNCI < 1$ можно трактовать и в положительном смысле для организации: ее сотрудники сумели опубликоваться в журналах, уровень которых выше, чем их конкретные статьи. В то же время $JNCI > 1$ может означать, что ученые регулярно «недооценивают» свои работы и, быть может, им стоит задуматься о подаче рукописей в журналы более высокого уровня (последнему наблюдению я обязан О. В. Москалевой (частная беседа).

Исследование NCI и $JNCI$ используется для анализа цитируемости статей и публикационной стратегии организаций как в фиксированный момент времени, так и в динамике. В последнем случае изменение показателей организаций отражает направление и темп эволюции научной деятельности ученых соответствующих вузов/институтов/лабораторий и т. д.

2.3.4. Ранговые индикаторы

Рассмотрим еще один подход, используемый при проведении кросс-дисциплинарных сопоставлений журналов. Можно принципиально отказаться от исследования абсолютных значений библиометрических показателей, а сравнивать лишь места, позиции в рейтинге, которые занимают издания в своих дисциплинах. После сортировки списков журналов в двух различных дисциплинах в порядке убывания какого-либо выбранного библиометрического индикатора (например, импакт-фактора, пятилетнего импакт-фактора; или индекса влияния, который будет введен в разделе 2.5) проводится анализ только занятых в этой иерархии изданий порядковых мест.

Суть рангового метода — разбиение упорядоченного списка-рейтинга

на n равных частей и определение, в какую из этих частей попадают журналы. Например, если n принимается равным 4, тогда говорят о квартилях — журналы первого квартиля (попадающие в верхнюю четверть списка по рассматриваемому параметру), второго квартиля, третьего квартиля, четвертого квартиля. В системе ранговых индикаторов считается, что журналы, попавшие в первый квартиль, выше журналов, попавших (не только в этой, но и в любой другой дисциплине) во второй квартиль и т. д. При этом журналы, попавшие в один и тот же квартиль в различных дисциплинах, считаются в рамках этого метода одинаковыми по научному уровню.

Как выбирать показатель n — непростой вопрос, и ответ на него зависит от конкретной задачи. Фактически, n задает «точность», «разрешающую способность» рангового индикатора. Журналы, оказавшиеся равными при выборе $n = 4$ (попавшие в один квартиль), могут получить разный уровень при $n = 5$ («квинтиль») и т. д. И хотя чем больше n , тем выше «разрешение» рангового метода, у более грубого подхода, у деления на те или иные более крупные «процентили» (части), есть свое преимущество: некоторые группы журналов попадают в один и тот же процентиль, что интуитивно соответствует представлению о том, что существуют «примерно одинаковые» издания.

Существует предельный случай «уточнения» рангового метода, когда весь ряд занятых журналами мест от 1 до N (N — число журналов в рассматриваемой дисциплине) проецируется на отрезок $[0; 1]$, и каждый журнал оказывается на своей точке этого отрезка. Несложно вычислить, что при этом показатель «относительной позиции» журнала, занимающего n -е место в своем дисциплинарном рейтинге, равен $(N - n) / (N - 1)$.

При кросс-дисциплинарном сравнении «выигрывает» тот журнал, у которого указанная «относительная позиция» выше. Это показано на рис. 10, где сравниваются журналы двух различных дисциплин. В одной из них (дисциплина А) 5 изданий, соответственно они занимают места с 1-го по 5-е в рейтинге по своей дисциплине, во второй рубрике (дисциплина В) 7 изданий с местами с 1-го по 7-е. После проецирования на отрезок $[0; 1]$ итоговое кросс-дисциплинарное ранжирование выглядит так (в угловые скобки заключены издания, занявшие одинаковые места): $\langle A_1, B_1 \rangle - B_2 - A_2 - B_3 - \langle A_3, B_4 \rangle - B_5 - A_4 - B_6 - \langle A_5, B_7 \rangle$.

Аналогичным образом та или иная модификация рангового метода может быть применена при сравнении как авторов, так и организаций между собой. При этом важно корректно и полно составить ту группу, объединяющую работающих в одной научной области, внутри которой будет определено взаимное место авторов/организаций (аналог научной дисциплины для журналов). В случае с авторами или узкоспециализированными организациями это сделать проще, в случае с организациями широкого профиля — сложнее, в этом случае может оказаться, что ранговые подходы неприменимы.

В заключение отметим, что одно из преимуществ ранговых методов — защита от сильной асимметричности распределений цитирования, когда небольшое число журналов (авторов, организаций) получают значительное число ссылок и создают такие средние показатели по своей группе, что все остальные участники не могут их достичь (ср. сноску п. 2.2.4, где ранговый подход, а именно использование медианы, предлагался для анализа импакт-факторов). В этом случае отказ от учета абсолют-

ных значений индикаторов и исследование лишь мест в рейтингах может оказаться более эффективным.

2.4. Индикаторы «экстремальной» цитируемости

Еще один подход, коренным образом отличающийся от разобранных выше, заключается в том, чтобы не рассчитывать среднюю цитируемость всего массива статей, а заниматься «крайностями», наблюдаемыми

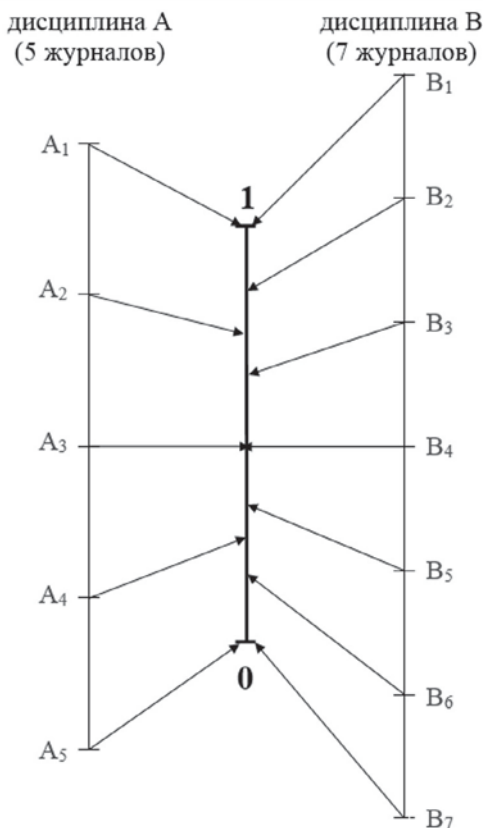


Рис. 10. Сравнение рангов журналов проецированием на отрезок $[0; 1]$.

A1... A5 — журналы дисциплины А.

B1... B7 — журналы дисциплины В

в нем, — теми статьями, которые получили аномально малое или аномально высокое число ссылок. Поэтому назовем показатели, основанные на такой методике, индикаторами «экстремальной» цитируемости, имея в виду «экстремум», предел возможной цитируемости как со стороны максимума, так и со стороны минимума. Со стороны минимума очевидным образом фигурируют публикации, не получившие ни одной ссылки. Со стороны максимума — т. н. «высокоцитируемые» статьи (*highly cited papers*), определение которых дается в п. 2.4.1.

Отличительной чертой «экстремально-го» подхода является то, что определяется лишь доля статей, попавших в две «предельные» категории, а число реально полученных ими ссылок не учитывается (впрочем, в случае минимального цитирования это число известно — 0). Важен вопрос, сколько в массиве высокоцитируемых и нецитируемых статей, каков их процент. Оба метода могут применяться по отдельности, они независимы друг от друга, необязательно использовать их в паре.

2.4.1. Высокоцитируемые статьи

В библиометрии использовались различные способы определения высокоцитируемых статей. Это могли быть и статьи, получившие определенное абсолютное число ссылок (данный метод, естественно, не может использоваться при анализе мультидисциплинарного множества публикаций), и статьи, занявшие высокие места по цитируемости в рамках определенной группы (например, «лучшие (по числу полученных ссылок) 10 публикаций университета»). Однако сейчас наиболее распространено определение, использующее «процентили» — верхний срез в n % наиболее цитируемых статей. В зависимости от ис-

следовательских задач показатель n может варьироваться, мы разберем подход, используемый в базе данных ESI.

В рамках ESI высокоцитируемой считается публикация, которая попадает в 1 % наиболее цитируемых работ в мире среди тех, которые вышли в том же, что и она, году и в той же области науки. Для каждого года выхода и каждой области науки в ESI высчитываются пороговые значения числа цитирований, которое должна получить к настоящему моменту статья, чтобы попасть в 1 % высокоцитируемых. Как и ранее в п. 2.3.2, год публикации фиксируется для того, чтобы убрать преимущество статей, вышедших давно, перед более свежими публикациями: у последних физически меньше времени набрать большое число цитирований. Ограничение по научной дисциплине, как мы знаем, «уравнивает» статьи, вышедшие в областях науки с высоким и низким средним уровнем цитируемости. Например, по данным на сентябрь 2014 г. высокоцитируемой статьей 2004 г. выхода по дисциплине «микробиология» была лишь та, которая получила за прошедшие 10 лет не менее 214 цитирований, в то время как статье по математике достаточно было получить 68 ссылок. Аналогичные пороги для публикаций 2009 г. выхода по этим дисциплинам составляли 116 и 34 ссылки соответственно¹⁰¹⁴.

Следует иметь в виду, что «высокоцитируемость» — динамическая категория, и статья, однажды ставшая высокоцитируемой, может в дальнейшем потерять этот статус. Это произойдет, если активность ее цитирования в дальнейшем сократится и ее

¹⁰¹⁴Заметим, что, в отличие от методов, разобранных в п. 2.3.2, нормализация не затрагивает тип документа. При этом в ESI обрабатываются только документы типа Article и Review.

обойдут другие публикации того же года выхода из той же области науки. Возросший с течением времени порог, необходимый для попадания статьи в высокоцитируемые, может превзойти число реальных ссылок, полученных данной статьей. Очевидно, что возможен и обратный процесс — статья, не являвшаяся высокоцитируемой, может начать получать много ссылок и попасть в престижную категорию. Поэтому анализ высокоцитируемых статей надо проводить в некоторый фиксированный момент, на множестве публикаций, являющихся в настоящее время высокоцитируемыми. Далее исследование можно периодически повторять, отслеживая динамику.

При анализе высокоцитируемых статей исследуется прежде всего абсолютный показатель — сколько таких работ опубликовал ученый, или лаборатория, или институт, или даже целая страна. Это важный результат, след участия не просто «в науке», но «в науке высочайшего уровня». Такие статьи можно называть научными шедеврами [Писляков, 2011].

Но еще более важным является относительный индикатор, показывающий, какова доля статей, написанных той или иной исследовательской единицей, которые становятся высокоцитируемыми. Может быть, каждая 20-я статья, написанная организацией/ученым, становится высокоцитируемой? Или, может быть, лишь каждая 200-я попадает в эту категорию? При этом здесь существует четкий ориентир, контрольный уровень, относительно которого можно сравнивать эффективность деятельности научного коллектива на мировом фоне. Поскольку, согласно определению, в мире каждая сотая статья является высокоцитируемой, то, если 1 % публикаций научного коллектива (каждая сотая статья) достига-

ет категории высокоцитируемости, это говорит о соответствии деятельности ученых мировому уровню¹¹⁵.

При реализации на практике следует помнить, что высокоцитируемыми становятся лишь 1 % статей, следовательно, анализ высокоцитируемости можно проводить только на больших публикационных массивах, охватывающих тысячи исходных документов. В противном случае выводы исследования будут слишком зависимы от случайных выбросов.

2.4.2. Коэффициент нецитируемости

Коэффициент нецитируемости, как следует из его названия, это доля научных работ, которые ни разу не были процитированы. Результаты научной деятельности научного коллектива тем эффективнее, чем меньше значение данного коэффициента. С некоторой степенью огрубления можно говорить, что чем ниже коэффициент нецитируемости, тем меньше «бесполезных», «бесплодных» публикаций выходит из-под пера сотрудников рассматриваемого коллектива¹²¹⁶.

Теперь предположим, что некоторая статья не получила на текущий момент

¹¹⁵Строго говоря, ситуация несколько сложнее. Высокоцитируемые статьи часто пишутся в больших соавторских коллективах, и для корректного соотнесения с мировым уровнем следует использовать дробный счет (fractional counting, см. п. 2.1.1). Точнее будет сказать, что если организация пишет статьи, становящиеся в менее 1 % случаев высокоцитируемыми, то она заведомо проигрывает на общемировом фоне. Это порог снизу, а вот обратное часто неверно.

¹²¹⁶Зеркальный взгляд на это явление реализован в InCites2, где считается показатель %Documents Cited, доля процитированных хотя бы один раз документов. Коэффициент нецитируемости, выраженный в процентах, равен 100 % минус %Documents Cited. Но следует иметь в виду, что в InCites2 используется не нормализованный по времени показатель (т. е. не такой, как вводится далее в данном разделе).

ни одной ссылки. Очевидно, что нельзя предугадать, будет ли она процитирована в неограниченной перспективе. Поэтому для грамотного анализа коэффициента нецитируемости выделяется некоторый временной интервал, величину которого фиксируют для публикаций любого года выхода, с тем чтобы более старые статьи не имели преимущества перед вышедшими недавно. Например, если в качестве такого интервала выбран пятилетний отрезок, то производится поиск статей, опубликованных в 2004 г. и не процитированных до 2008 г. включительно, опубликованных в 2005 г. и не процитированных до 2009 г. включительно и т. д. Это позволяет провести корректный анализ динамики коэффициента нецитируемости и определить, например, что доля ни разу не процитированных публикаций организации/автора с 2004 по 2009 г. увеличилась или уменьшилась. Для сравнения различных организаций или авторов между собой необходимо оставаться в рамках одного дисциплинарного направления, поскольку среднемировой коэффициент нецитируемости, как и остальные ненормализованные показатели цитируемости, сильно зависит от рассматриваемой области науки.

В отличие от анализа высокоцитируемости, исследование другого показателя «экстремальной» цитируемости не требует очень больших массивов данных. Например, как было показано [Писляков, 2010], пятилетний коэффициент нецитируемости всех российских статей, вышедших в 2004 г. (по SCIE и SSCI), составил около 40 %. Это означает, что исследуемому массиву публикаций достаточно содержать сотни или даже десятки статей.

2.5. «Взвешенные» индикаторы. Собственный фактор (Eigenfactor) и индекс влияния статьи (Article Influence)

Все изученные ранее индикаторы имели одну общую черту: ссылки, которые получали журнал/автор/организация, были для них «одинаковы», показатели «не замечали», откуда именно получено то или иное цитирование. Замена одного цитирующего журнала на другой, который дает столько же ссылок, не привело бы к изменению ни одного индикатора.

Следующие два индикатора называются «взвешенными», поскольку они учитывают полученные журналом цитирования с различным весом — в зависимости от того, насколько «влиятелен» тот источник, из которого получено цитирование. Эта влиятельность, в свою очередь, зависит от цитируемости самого источника ссылки. Таким образом, ссылка из высокоцитируемого журнала, например Nature, «ценится» выше, чем ссылка из малоизвестного регионального издания. Аналогичная идея лежит в основе расчета ранга «авторитетности» веб-страниц некоторыми поисковыми машинами, например Google (т. н. «PageRank»).

Взвешенный журнальный индикатор, получивший название «собственный фактор» (Eigenfactor), был предложен в 2007 г. специалистами из лаборатории Карла Бергстрёма (Университет Джорджа Вашингтона). Разработчики данного показателя иллюстрировали его с помощью процесса случайного перехода по ссылкам: некий читатель переходит от журнала по случайной выбранной в нем библиографической ссылке к другому журналу. Для соединения не связанных между собой ссылками изданий дополнительно вводится некоторая ненулевая вероятность того, что читатель выберет

следующий журнал случайным образом, а не путем следования по ссылке. Итоговый показатель «престижа» издания будет пропорционален числу заходов читателя в соответствующий журнал.

Данный процесс реализует описанную выше идею о разной ценности ссылок из журналов с разной цитируемостью. Ведь для того чтобы воображаемый читатель часто заходил в некий журнал, необходимо не только чтобы «путей» (ссылок) к этому журналу было много, но и чтобы приходили эти ссылки из часто посещаемых читателем журналов.

Еще один механизм, которым иллюстрируют взвешенные показатели, — это «задача престижа» журналом в результате итеративной процедуры. Сначала каждому журналу назначается одинаковый начальный «престиж», равный единице. На первом этапе итерации каждый журнал «делит» свой единичный престиж между всеми ссылками, которые он «посылает» другим журналам. Чем меньше в журнале сделано ссылок, тем больше «вес» каждой из них¹⁷. Таким образом рассчитывается «престиж», полученный каждым журналом из других журналов. Это значение престижа, полученного в результате первой итерации, принимается в качестве нового значения престижа журнала, которое на второй итерации вновь «делится» между всеми ссылками, исходящими из журнала. Далее процесс повторяется до достижения квазистабильного состояния, когда очередной шаг итерации уже практически не меняет значения «престижа», полученного каждым из журналов.

¹⁷Точно так же в случае случайного перехода по ссылкам: чем меньше ссылок в журнале, тем больше вероятность, что воображаемый читатель воспользуется какой-то конкретной из них, а значит, выше ценность каждой ссылки.

(Точный алгоритм устроен таким образом, что данное состояние будет обязательно достигнуто.)

Согласно строгому математическому определению, для вычисления значения собственного фактора решается матричное уравнение и находится собственный вектор матрицы вероятностей переходов (отсюда название «собственный фактор»). Собственный фактор — показатель журнала как целого. Если его нормировать на число статей в издании, получится индикатор, характеризующий среднюю статью в журнале и называемый «индексом влияния статьи» (Article Influence). Оба показателя публикуются с 2007 г. в JCR.

При вычислении собственного фактора журнала и индекса влияния статьи учитываются ссылки текущего года на публикации, вышедшие за пять предыдущих лет, — публикационное окно и окно цитирования совпадают с установленными для пятилетнего импакт-фактора. Отличительной особенностью этих показателей является то, что ссылки журнала на свои же статьи не учитываются, — самоцитирование при расчете исключается.

Использование взвешенных индикаторов в совокупности с традиционным импакт-фактором (двух- или пятилетним) позволяет взглянуть на цитируемость журналов более объемно. В частности, при их сопоставлении, как правило, рельефно выделяются журналы, которые получили высокие импакт-показатели благодаря самоцитированию или ссылкам (зачастую взаимным) из узкого, замкнутого круга изданий. При высоком значении импакт-фактора (импакт-факторов) эти журналы будут иметь низкие значения взвешенных индикаторов: такого рода «островки взаимного цитирования» не дадут высокие

значения собственному фактору журнала / индексу влияния статьи. В модели «случайного блуждания» читатель редко будет оказываться на этих островках, если они не связаны с остальными журналами полученными из них ссылками (а кроме того, из рассмотренных нами взвешенных индикаторов вообще исключено самоцитирование). Например, один из журналов, уличенный в практике «взаимного договорного цитирования», в JCR-2012 занимал 14-е место по импакт-фактору в дисциплине «экономика» — и при этом 197-е(!) место по индексу влияния статьи.

Наконец, отметим, что индекс влияния в какой-то степени сглаживает разницу уровней активности цитирования в различных научных дисциплинах (главная проблема при использовании абсолютных значений импакта), поскольку «престиж» журнала распределяется поровну между всеми исходящими из него ссылками. В дисциплинах, где среднее число ссылок в одной статье велико, вес одной ссылки будет меньше, чем в тех областях науки, где ссылок в расчете на статью меньше и «ссылочное поле» менее плотное. Кроме того, пятилетнее публикационное окно также способствует более равновесному учету цитирований в различных научных дисциплинах.

2.6. Индекс Хирша (h -index) и «хирш-подобные» показатели

Индекс Хирша был предложен физиком из Калифорнийского университета в Сан-Диего Х. Хиршем [Hirsch, 2005]. Этот показатель кардинальным образом отличается от всех рассмотренных ранее, поскольку он пытается дать комплексную оценку одновременно числу публикаций ученого и их цитируемости (т. е. «количеству и качеству»). Кроме того, методика подсчета

индекса Хирша намеренно уходит от определения средних величин, средней цитируемости статей (в этом он схож с методиками «экстремальной» цитируемости, разобранными в разделе 2.4). Это имеет свой смысл, поскольку усреднение зачастую не дает полной картины об исследуемом множестве статей и не позволяет производить корректное сравнение эффективности научной деятельности авторов и организаций. При усреднении сильные искажения могут быть вызваны отдельными публикациями, которые получают очень много ссылок, являясь фактически «выбросами», выделенными случаями.

Индекс Хирша имел невероятный успех в научном сообществе, вдохновив специалистов в библиометрии на создание многих десятков своих «модификаций» и «усовершенствований». Мы рассмотрим лишь два примера этих «хирш-подобных» метрик, однако начнем с рассмотрения классического индекса.

2.6.1. Индекс Хирша

Согласно определению, индекс Хирша массива публикаций ученого равен h , если есть h статей из этого массива, каждая из которых получила не менее h цитирований, а каждая из остальных статей при этом получила не более h цитирований. Эквивалентное определение: индекс Хирша массива публикаций ученого равен h , если есть h статей из этого массива, каждая из которых получила не менее h цитирований, но нет $h + 1$ статей, каждая из которых получила бы не менее $h + 1$ цитирований.

Например, если у ученого есть 5 статей, которые получили 10, 8, 2, 2 и 0 ссылок соответственно, то его индекс Хирша равен 2, поскольку есть 2 статьи, каждая из которых цитировалась не менее 2 раз (таких статей

даже 4), но уже нет 3 статей, каждая из которых цитировалась бы не менее 3 раз¹⁴¹⁸.

Графически это проиллюстрировано на рис. 11. Каждая точка на графике соответствует одной публикации ученого. Публикации упорядочены по цитируемости: от тех, которые получили больше всего ссылок, к наименее цитируемым статьям. По оси абсцисс на рис. 11 отложен порядковый номер публикации в этом упорядоченном списке. По оси ординат — сколько ссылок получила статья с соответствующим номером. Ввиду упорядочения по цитируемости, у точек постепенно уменьшается ордината при увеличении абсциссы. Если мысленно соединить точки для получения графика, то h -индекс будет находиться в районе пересечения этого графика с показанной на рисунке биссектрисой $y = x$. Точнее, индекс Хирша будет соответствовать максимальному номеру статьи, находящейся не ниже прямой $y = x$ (выше или на этой прямой). Как видно из геометрии графика, у рассматриваемого ученого есть h статей, каждая из которых цитируется не менее h раз, но при этом у него нет $h + 1$ статей, каждая из которых цитируется не менее $h + 1$ раз: статья с номером $h + 1$ находится ниже прямой $y = x$, а значит получила менее $h + 1$ ссылок.

Основное свойство индекса Хирша: увеличение публикационной активности без достаточной цитируемости публикуемых работ не приведет к его росту. И наоборот: получение большого числа цитирований одной-двумя статьями также не повысит индекс серьезным образом. Если автор на-

писал 100 статей, но при этом каждая из них получила 5 ссылок, индекс Хирша ученого будет равен всего лишь 5. С другой стороны, если автор написал лишь 5 статей, каждая из которых получила по 100 ссылок, его индекс все равно будет равен 5. Для достижения высокого значения индекса Хирша необходимо, чтобы автор писал много таких работ, каждая из которых получала бы много цитирований. Именно в этом смысле h -индекс пытается отразить сбалансированную оценку сразу и публикационной активности ученого, и цитируемости его работ.

Введенный изначально для авторов, индекс Хирша может быть распространен на организации, страны и вообще любой набор статей, для каждой из которых известно число полученных ссылок.

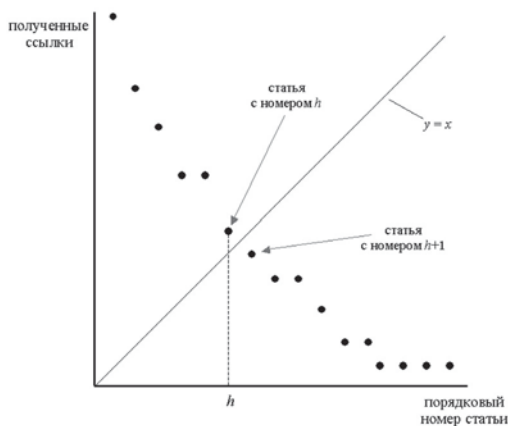


Рис. 11. Индекс Хирша (h -индекс)

Для ученых часто считают «пожизненный» индекс Хирша за всю научную карьеру (как было сделано Х. Хиршем в исходной статье), однако это приводит к очевидному неравенству пожилых и молодых специалистов. Дело в том, что такой показатель не может уменьшаться с течением времени,

¹⁸В строгом математическом виде определение индекса Хирша формулируется так: пусть у ученого есть N статей, упорядочим их от наиболее цитируемых к наименее цитируемым. Обозначим c_i число ссылок, полученных i -й статьей, $c_i \geq c_{i+1}$ для всех $i = 1 \dots N - 1$. Тогда индекс Хирша $h = \max\{i: c_i \geq i\}$.

только расти или оставаться неизменным¹⁵¹⁹, — и ученый может уже много лет находиться на пенсии и не писать научных трудов, однако h -индекс его будет не меньше, чем он был на пике карьеры.

Поэтому для получения более осмысленного показателя следует, как и для любой библиометрической величины, задать публикационное окно и окно цитирования. Например, можно рассматривать все статьи, вышедшие у ученого за пятилетний промежуток, и учитывать ссылки, полученные этими статьями в течение того же пятилетнего интервала.

Следует упомянуть еще о некоторых ограничениях, «недостатках» индекса Хирша:

- h -индекс не предполагает поправки на дисциплинарную область: сравнение абсолютных значений индекса у ученых, работающих в разных областях науки, невозможно; это не нормализованный по областям науки показатель;

- h -индекс является целым числом, следовательно, не обладает идеальной «разрешающей способностью» (были предложены способы обобщить h -индекс на дробные значения, здесь мы не будем на них останавливаться);

- в индексе не учитывается реальное число ссылок, полученное статьями выше и ниже точки h : важно только местонахождение этой точки. Так, например, если бы каждая из h первых статей на рис. 11 получила в 10 раз больше ссылок (а цитируемость остальных публикаций осталась прежней), на значении h -индекса это бы ни-

как не отразилось. $(h + 1)$ -я статья все равно осталась бы лежать ниже линии $y = x$ и индекс Хирша сохранил свое значение. И наоборот: допустим, что два гипотетических автора опубликовали по 50 статей каждый. Каждая статья первого из них получила по 10 ссылок, а у второго 10 статей получили по 10 ссылок, а остальные 40 статей вообще не цитировались. В этом случае, при очевидном неравенстве «вклада» авторов, их индекс Хирша будет иметь одинаковое значение — 10.

Успех индекса Хирша был во многом связан с тем, что ряд ученых, а главное, администраторов науки, посчитали его простым и универсальным средством для оценки ученых/организаций и т. д. То, что эффективность научной деятельности можно охарактеризовать одним числом, оказалось слишком сильным соблазном. В действительности, как мы видим, h -индекс — это не панацея, а лишь один из библиометрических индикаторов, который должен использоваться обязательно в комплексе с другими метриками для разносторонней оценки авторов и научных коллективов.

2.6.2. h_a -индекс

Вообще говоря, условие равенства числа статей и минимального числа полученных каждой из этих статей ссылок является в определении индекса Хирша произвольным. Почему это не требование, чтобы h статей получили каждая не менее 2 h ссылок? Или, наоборот, чтобы каждая из h статей получила не менее $h/2$ ссылок? Имея это в виду, Н. Ван Эйк и Л. Валтман [Van Eck, Waltman, 2008] «обобщили» определение индекса Хирша, назвав его h_a -индекс.

Определение: обобщенный h_a -индекс массива публикаций равен h_a , если каждая из h_a статей этого массива получила не ме-

¹⁵¹⁹Строго говоря, все-таки существует нечастая ситуация, при которой h -индекс ученого способен уменьшиться. Это может произойти в случае, если ученый отозвал свою уже опубликованную в журнале статью (retraction) — из-за найденных серьезных ошибок или обнаружения плагиата и т. д. На это обстоятельство обращал внимание В. Глэнцел (в частной беседе).

нее αh цитирований, а каждая из остальных — не более αh цитирований.

Графически смысл h_α -индекса отражен на рис. 12. По графику h_α -индекс можно определить таким же образом, как и обычный индекс Хирша, но следует искать пересечение не с прямой $y = x$, а с прямой $y = \alpha x$. На рис. 12 показаны примеры для $\alpha = 2$ и $\alpha = 1/2$.

Основной смысл нового индекса в следующем. Как мы говорили, индекс Хирша одновременно оценивает число публикаций и их цитируемость, делая это в каком-то смысле «равномерно» (в смысле условия равенства числа статей и минимального числа полученных каждой из этих статей ссылок). При помощи h_α -индекса мы можем управлять этой «равномерностью». При увеличении α на первое место выходит цитируемость, наличие высокочитируемых работ. При уменьшении — все важнее становится продуктивность, большое число публикаций. Это понятно из графического представления h_α -индекса, а мы продемонстрируем на упомянутом ранее примере: пусть один автор написал 5 статей, каждая из которых получила по 100 ссылок, а вто-

рой написал 100 статей, каждая из которых получила 5 ссылок. При $\alpha = 5$ у первого автора h_α -индекс равен 5 (есть 5 статей, каждая из которых процитирована не менее $\alpha \times 5 = 5 \times 5 = 25$ раз), а у второго — 1 (есть лишь 1 статья, которая процитирована не менее $\alpha \times 5 = 1 \times 5 = 5$ раз, а двух статей, получивших по 10 ссылок, нет). При $\alpha = 1/5$ у первого автора h_α -индекс остается равным 5, а у второго возрастает до 25 (есть 25 статей, каждая из которых процитирована не менее $\alpha \times 25 = (1/5) \times 25 = 5$ раз)¹⁶²⁰.

Таким образом, считая h_α -индекс при разных α , можно исследовать разные аспекты исследуемого множества статей, в зависимости от поставленных задач больше фокусируясь на публикационной активности, на числе публикаций, или делая основной акцент на качественном аспекте, на цитируемости изучаемых работ.

2.6.3. g-индекс

Еще одна попытка «улучшить» индекс Хирша — это введенный L. Egghe [Egghe, 2006] g-индекс. Он часто встречается в литературе (хотя и реже, чем h-индекс). По определению, g-индекс ученого равен такому максимальному g, что g наиболее цитируемых статей ученого получили вместе не менее g^2 ссылок. (Соответственно, любые $g + 1$ его статей совместно получили меньше $(g + 1)^2$ ссылок.) Обратите внимание, что здесь

²⁰Если рассмотреть предельные случаи, то для любого автора максимальное значение α , при котором его h_α -индекс все еще больше нуля (т. е. равен 1), совпадает с числом цитирований, которое получила его самая высокочитируемая статья. Число статей при этом не играет никакой роли. Наоборот, при уменьшении α h_α -индекс возрастает и достигает своего максимального значения, которое равно общему числу статей данного автора, получивших хотя бы одну ссылку. Здесь совершенно не важна цитируемость статей (за исключением триггера «была ли статьей получена хотя бы одна ссылка»).

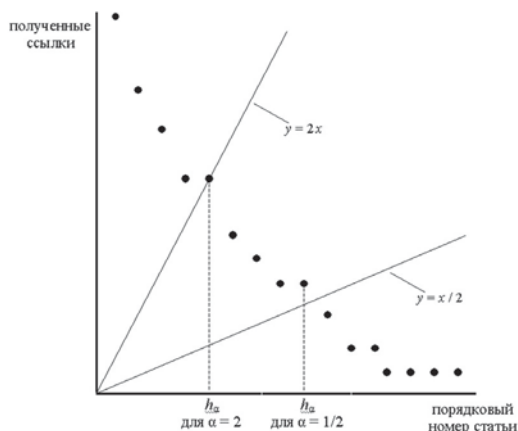


Рис. 12. h_α -индекс. Пример для $\alpha = 2$ и $\alpha = 1/2$

рассматривается совокупная цитируемость первых g статей, упорядоченных в порядке убывания цитируемости. В случае h -индекса требования по цитируемости предъявлялись к каждой отдельной статье, в этом ключевое отличие индексов друг от друга.

Вернемся к примеру из п. 2.6.1, где рассматривались пять работ ученого, которые получили 10, 8, 2, 2 и 0 ссылок соответственно. g -индекс этого автора равен 4 — поскольку есть 4 статьи, которые цитировались вместе не менее $4^2 = 16$ раз (на самом деле они цитировались $10 + 8 + 2 + 2 = 22$ раза), но нет 5 статей, которые бы в сумме цитировались $5^2 = 25$ раз. Напомним, h -индекс данного ученого оказался равным 2. В действительности g -индекс всегда больше или равен h -индексу. Это следует из того, что каждая из h статей получила, по определению индекса Хирша, не менее h ссылок, а значит, вместе они получили как минимум h^2 цитирований. Поэтому h статей заведомо удовлетворяют условию, которое предусмотрено g -индексом. Следовательно, $g \geq h$.

Можно переформулировать определение g -индекса так: g -индекс ученого равен такому максимальному g , что у ученого есть g статей со средней цитируемостью всей их совокупности (в расчете на одну статью) не менее g . Определения эквивалентны, поскольку наличие g статей со средней цитируемостью их совокупности не менее g означает, что в сумме эти статьи получили не менее g^2 ссылок.

В чем преимущество g -индекса? Он лишен одного из недостатков, отмеченных для индекса Хирша. Цитируемость наиболее цитируемых статей (получивших больше ссылок, чем статья с номером g) может влиять на значение g -индекса. Так, если в приведенном выше примере самая цитируемая статья со временем получит еще три ссыл-

ки (и их станет 13, цитируемость остальных статей пусть не меняется), g -индекс ученого достигнет пяти. Напомним, что в h -индексе не учитывается реальное число ссылок, полученное статьями выше и ниже точки h : важно только местонахождение этой точки. Любое увеличение цитируемости самой цитируемой статьи не повлияет на h -индекс.

2.7. Показатели хронологического распределения ссылок

В тексте данной главы неоднократно упоминалось о «старении» научного знания. Как быстро новое знание воспринимается, вводится в научный оборот и, напротив, как долго не теряет своей актуальности и остается востребованным. Эти характеристики можно оценить по хронологической структуре библиографических ссылок, которые содержатся в том или ином журнале, работах ученого, статьях института или, наоборот, которые цитируют данный журнал, работы данного ученого, статьи института. В данном случае речь уже не идет о попытке оценить качество, эффективность публикаций по числу полученных ими ссылок. Это другое исследование. И тем не менее журнал, цитирующий в основном классические источники многолетней давности, отличается от издания, стремящегося знакомить своих читателей с самой современной литературой по рассматриваемому вопросу. А если ученый опубликовал работу, в дальнейшем регулярно цитируемую из года в год, его вклад, скорее всего, будет иного характера, чем у исследователя, получившего много ссылок в первые 2–3 года после публикации статьи и впоследствии забытого.

Для хронологической характеристики распределения сделанных журналом/автором/

организацией цитирований чаще всего используют значение медианы этого распределения. Половина сделанных ссылок ведет на источники, опубликованные позже значения медианы, половина — на более свежую литературу. На рис. 13 показано хронологическое распределение ссылок, сделанных в статьях журнала *Annual Review of Biochemistry*, вышедших в 2013 г. В этих статьях была 21 ссылка на публикации 2013 г., 424 ссылки на источники, вышедшие в 2012 г., и т. д. В общей сложности статьи 2013 г. содержали 4883 ссылки, из них половина пришлась на документы, вышедшие в 2007–2013 гг., половина — на более поздние источники. Поэтому медиана соответствует документам семилетней давности. В JCR для этого показателя используют термин «время полужизни сделанных ссылок» (citing half-life). Таким образом, для данного журнала в 2013 г. время полужизни сделанных ссылок равно 7 (затенен-

ной областью на рисунке отмечены ссылки на более поздние источники)¹⁷²¹.

Время полужизни ссылок зависит от научной области, к которой относится журнал. Собственно, оно и характеризует эту область с точки зрения темпов устаревания производимого в ней научного знания. В JCR-2013 минимальное время полужизни сделанных ссылок у журналов дисциплины «нанонауки и нанотехнологии» (5,9 лет). Максимальный уровень этого показателя оценить сложно, поскольку JCR не указывает его точное значение для тех областей науки, где оно больше 10 лет. Таких дисциплин в JCR-2013 в общей сложности 48, среди них, в частности, история, социология, зоология, геология, лингвистика и др. В этих областях научное знание имеет максимально продолжительный цикл жизни, и больше половины всех публикуемых ссылок ведет на источники более чем десятилетней давности.

Аналогичный показатель также вводится не для сделанных ссылок, а, наоборот, для полученных журналом цитирований. Он называется *cited half-life*, время полужизни полученных ссылок, и характеризует то, насколько свежие статьи журнала цитируются в мировой науке. Если его величина мала, значит, в науке используются прежде всего свежие, недавние публикации издания. Если велика — полезными оказываются и достаточно старые материалы, публиковавшиеся в журнале. Время полужизни полученных ссылок для целых научных дисциплин варьируется в JCR-2013 от 4,1 года (снова нанонауки) до «более 10 лет» (31 дисциплина). Безусловно, *cited half-life* связан с *citing*

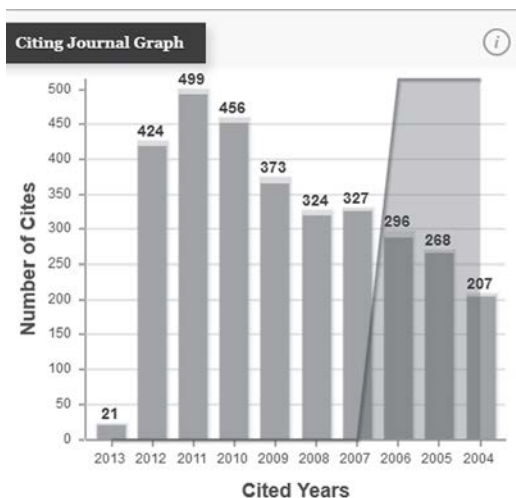


Рис. 13. Хронологическое распределение ссылок в статьях журнала *Annual Review of Biochemistry*, вышедших в 2013 г. Скриншот интерфейса базы данных JCR (Thomson Reuters)

²¹Если быть точным, в приведенном примере на публикации 2007–2013 гг. приходится 49,6 % всех ссылок (2424 из 4883). Если добавить еще один год (2006), то будет охвачено уже 55,7 % всех ссылок. Поэтому делается пропорциональная поправка и время полужизни для этого издания указывается в JCR равным 7,1 года.

half-life: напрямую они не зависят друг от друга, однако оба показателя отражают скорость эволюции знания в соответствующей научной области. Поэтому, например, среди 31 дисциплины со временем полужизни полученных ссылок более 10 лет в 28 дисциплинах время полужизни сделанных ссылок также превышает 10 лет.

Оба изученных индикатора хронологического распределения ссылок в журналах с легкостью могут быть распространены на статьи, написанные отдельным ученым или каким-либо научным коллективом.

ЗАКЛЮЧЕНИЕ

Мы рассмотрели основные библиометрические показатели, характеризующие цитируемость разных «научных единиц» — ученых, научных коллективов, организаций и целых стран. Были также разобраны разнообразны журнальные показатели. Заметим, что это далеко не исчерпывающий список индикаторов, существующих в библиометрии, и даже не исчерпывающий перечень индикаторов, доступных в инструментах Thomson Reuters. Не имеющие прямого отношения к цитируемости и оценке эффективности научной деятельности, нами были оставлены в стороне такие вопросы, как измерение международного сотрудничества в публикациях, характеристики сетей соавторства, индексы научной специализации, коэффициенты ко-цитирования и др.

Иные показатели не рассматривались ввиду их тривиальности — например, интегральная публикационная активность или суммарная цитируемость того или иного объекта исследования.

Неоднократно в течение настоящей главы встречалось напоминание о том, что идеальных индикаторов не существует, что они должны использоваться не по отдельности, а в комплексе. Цель, поставленная здесь нами, — научить грамотно и многосторонне изучать эффективность науки с помощью библиометрии. Но результаты даже самого подробного библиометрического исследования не отменяют необходимости участия экспертов, профессионалов-предметников при принятии решений из области science policy: при распределении грантов и финансирования, осуществлении кадровой политики, открытии/закрытии научных направлений, программ и лабораторий и т. д. На основании чистой библиометрии не могут выноситься решения на поле управления наукой — ее роль заключается в том, чтобы помогать в принятии таких решений, сообщать дополнительные сведения об изучаемых «научных единицах»: ученых, лабораториях, факультетах, организациях, регионах, странах, журналах. В западном наукометрическом сообществе в таких случаях говорят о концепции «informed decision», «решения при полной информации», в реализации которой и должен играть свою роль библиометрический анализ.

Писляков В. В. Библиометрические индикаторы: практикум. М.: НФПК; Инфра-М, 2014. 60 с.

Писляков В. В. Соавторство российских ученых с зарубежными коллегами: публикации и их цитируемость. Препринт WP6/2010/01. М.: ГУ-ВШЭ, 2010. 40 с.

Писляков В. В. Шедевры научного творчества: анализ высокочитруемых статей российских ученых // На-

учно-техническая информация. Сер. 2: Информационные процессы и системы. 2011. № 12. С. 1–8.

Писляков В. В., Дьяченко Е. Л. Эффект Матфея в цитировании статей российских ученых, опубликованных за рубежом // Научно-техническая информация.

Сер. 2: Информационные процессы и системы. 2009. № 3. С. 19–24.

Aksnes D. W., Sivertsen G. The effect of highly cited papers on national citation indicators. *Scientometrics*. 2004. Vol. 59. Iss. 2. P. 213–224.

Eck N. J. van, Waltman L. Generalizing the h- and g-indices // *Journal of Informetrics*. 2008. Vol. 2. Iss. 4. P. 263–271.

Egghe L. Theory and practise of the g-index // *Scientometrics*. 2006. Vol. 69. Iss. 1. P. 131–152.

Hirsch J. E. An index to quantify an individual's scientific research output // *Proceedings of the National Academy of Sciences*. 2005. Vol. 102. № 46. P. 16569–16572.

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. Vol. 57. Iss. 8. P. 997–1004.

McVeigh M. E. Journal self-citation in the Journal Citation Reports – Science Edition (2002). 2004. URL: <http://wokinfo.com/essays/journal-self-citation-jcr>.

Pislyakov V., Dyachenko E. Citation expectations: are they realized? Study of the Matthew index for Russian papers published abroad // *Scientometrics*. 2010. Vol. 83. № 3. P. 739–749.

Rousseau R. Journal evaluation: Technical and practical issues // *Library Trends*. 2002. Vol. 50. Iss. 3. P. 418–439.

Помимо справочного аппарата данной главы, для более глубокого знакомства с библиометрическими индикаторами читателю рекомендуются следующие отечественные и зарубежные источники:

Бредихин С. В., Кузнецов А. Ю., Щербакова Н. Г. Анализ цитирования в библиометрии. Новосибирск; М.: ИВМИМГ СО РАН, НЭИКОН, 2013.

Кузнецов А. Ю., Бредихин С. В. Методы библиометрии и рынок электронной научной периодики. Новосибирск; М.: ИВМИМГ СО РАН, НЭИКОН, 2012.

Писляков В. В. Методы оценки научного знания по показателям цитирования // *Социологический журнал*. 2007. № 1. С. 128–140.

Egghe L., Rousseau R. Introduction to Informetrics: Quantitative Methods in Library, Documentation and Information Science. Amsterdam e. a.: Elsevier Science Publishers, 1990.

Garfield E. Citation indexes for science. A new dimension in documentation through association of ideas // *Science*. 1955. Vol. 122. № 3159. P. 108–111.

Garfield E., Sher I. H. New factors in the evaluation of scientific literature through citation indexing // *American Documentation*. 1963. Vol. 14. № 3. P. 195–201.

Glänzel W. Bibliometrics as a Research Field: A course on theory and application of bibliometric indicators. Course Handouts. Leuven, 2003.

Glänzel W., Moed H. F. Journal impact measures in bibliometric research // *Scientometrics*. 2002. Vol. 53. Iss. 2. P. 171–193.

Gonzalez-Pereira B., Guerrero-Bote V., Moya-Anegón F. A new approach to the metric of journals scientific prestige: The SJR indicator // *Journal of Informetrics*. 2010. Vol. 4. Iss. 3. P. 379–391.

Handbook of Quantitative Science and Technology Research / ed. by H. F. Moed, W. Glänzel, U. Schmoch. Dordrecht: Kluwer Academic Publishers, 2004.

Moed H. F. Citation Analysis in Research Evaluation. Dordrecht: Springer, 2005.

Moed H. F. Measuring contextual citation impact of scientific journals // *Journal of Informetrics*. 2010. Vol. 4. Iss. 3. P. 265–277.

Pislyakov V. Comparing two «thermometers»: Impact factors of 20 leading economic journals according to Journal Citation Reports and Scopus // *Scientometrics*. 2009. Vol. 79. № 3. P. 541–550.

Pislyakov V., Shukshina E. Measuring excellence in Russia: Highly cited papers, leading institutions, patterns of national and international collaboration // *Journal of the Association for Information Science and Technology*. 2014. Vol. 65. Iss. 11. P. 2321–2330.

Rehn C., Gornitzki C., Larsson A., Wadskog D. Bibliometric Handbook for Karolinska Institutet. 2014. URL: https://kib.ki.se/sites/kib.ki.se/files/Bibliometric_Handbook_2014.pdf.

Thomson Reuters. InCites Indicators Handbook. 2014. URL: <http://researchanalytics.thomsonreuters.com/m/pdfs/indicators-handbook.pdf>.

НАУЧНЫЕ ПУБЛИКАЦИИ КАК СРЕДСТВО КОММУНИКАЦИИ, АНАЛИЗА И ОЦЕНКИ НАУЧНОЙ ДЕЯТЕЛЬНОСТИ



О. В. Москалева

Советник директора Научной библиотеки им. М. Горького Санкт-Петербургского государственного университета, кандидат биологических наук.

RESEARCH PUBLICATIONS AS MEANS OF COMMUNICATION,
ANALYSIS AND ASSESSMENT OF RESEARCH ACTIVITY

DOI 10.15826/B978-5-7996-1352-5.0006

This chapter gives a brief description of the evolution of research journals and creation of research publication databases. The author provides information about citation theories, as well as methods of evaluating journals and research output of individual scholars, organizations and countries, based on citation analysis. The peculiarities of research publications in different subject areas are described, as well as approaches to comparative citation analysis in various research fields. Special attention is paid to alternative methods of publication activity evaluation, and use of bibliometric indicators for university and research institution rankings.

В главе приводится краткое описание эволюции научного журнала, формирование баз данных научных публикаций. Приведена информация о существующих теориях цитирования и основанных на показателях цитирования способах оценки качества журналов, научной деятельности отдельных ученых, организаций и стран. Описываются отличительные особенности публикаций в разных научных областях и возможные подходы к сравнительному анализу цитирования публикаций в разных областях знаний. Отдельно рассматриваются альтернативные цитированию способы оценки публикаций и использование библиометрических показателей при построении рейтингов университетов и научных организаций.

Целью любой научной деятельности является в конечном итоге получение новых знаний и создание новых технологий, способствующих улучшению качества жизни. Для того чтобы полученные результаты могли быть использованы, необходимо донести информацию о них до реальных «потребителей», а также

до «производителей» новых научных результатов.

В древности и в Средние века большая часть обмена знаниями происходила в личном общении и личной переписке, а основные важнейшие результаты излагались в виде научных трактатов, являвшихся итогом многолетней деятельности ученых. В даль-

нейшем акцент начал смещаться в сторону формирования научного журнала, посредством которого полученные результаты исследований доносятся до заинтересованных сторон. Ту же цель преследует становление системы профессиональных научных встреч — конференций, симпозиумов, семинаров и т. д. В последнее время особую роль приобретают открытые источники информации в сети Интернет — интернет-журналы, интернет-конференции, научные блоги и т. п.

Особое место в системе научной коммуникации занимают патенты и иные охраняемые документы, фиксирующие приоритет исследователя в создании новой технологии, материала, полезной модели, штамма микроорганизма, сорта растений и т. д. От обычной научной публикации (статьи в журнале, докладе на конференции или монографии) такие публикации отличаются еще и тем, что являются полноправными объектами товарно-денежных отношений и не могут быть использованы безвозмездно для производства товаров и услуг. Тем не менее по отношению к патентам возможно применение тех же подходов, что и по отношению к иным научным публикациям в плане их рассмотрения как средства коммуникации, анализа и оценки научной деятельности.

3.1. История создания и характеристики научных журналов

Еще в 60-е гг. прошлого века в книге «Малая наука, большая наука» Дерек де Солла Прайса, считающегося одним из основоположников наукометрии, отмечался экспоненциальный рост количества научных журналов, научных организаций, ученых [Price, 1965]. Рост количества ученых затрудняет их личное общение, экспоненциальный рост числа источников информации

приводит к обострению проблем, связанных с поиском релевантной достоверной информации, обостряет проблему выбора издания для публикации полученных результатов и т. д.

Старейшие научные журналы — *Journal des sçavans*, *The Philosophical Transactions of the Royal Society*, *Giornale de' Letterati*, *Acta eruditorum* — появились в XVII в., при этом «Философские труды Королевского общества» (*The Philosophical Transactions of the Royal Society*) выходят с момента создания (март 1665 г.) до настоящего времени без перерывов. В самом начале эти издания больше походили на собрание новостей об изданных книгах, но постепенно в них стали появляться и оригинальные статьи и сообщения о полученных научных результатах. К середине XIX в. научные журналы приобрели практически современную привычную форму, с современной структурой статей, предполагающих наличие ссылок на использованные источники, рефератов и других привычных атрибутов научной статьи. По словам В. Г. Белинского, «журнал есть не наука и не ученость, но, так сказать, фактор науки и учености, посредник между наукою и учеными. Как бы ни велика была журнальная статья, но она никогда не изложит всей системы какого-либо знания. Она может представить результаты этой системы, чтобы обратить внимание ученых, как скорое известие, и публики, как рапорт о случившемся» [цит. по: Акопов, 2007], т. е. основной функцией журнала с момента его создания и до сих пор является коммуникация.

В настоящее время, по данным Ulrich's Periodicals, в мире насчитывается порядка 200 тыс. изданий, относящихся к периодическим (журналам), но далеко не все из них являются научными журналами в строгом

понимании этого термина. Условимся, что научным является журнал, все статьи в котором проходят предварительное рецензирование.

Точное количество научных журналов в настоящее время можно приблизительно оценить в 70–100 тыс. Только в России ежегодно создаются около 200 новых научных журналов. Динамику изменения количества журналов можно оценить приблизительно по количеству журналов, индексируемых Web of Science Core Collection, — за последние 15 лет количество изданий, представленных в Journal Citation Reports (Science & Social Science), увеличилось почти вдвое — с 6,6 тыс. в 1997 г. до 11,5 тыс. в 2013-м. При этом если рост количества научных журналов в естественных и точных науках (Science) происходил все это время, хотя с 2007 г. значительно большее количество журналов включалось в индекс ежегодно, то для социальных наук количество журналов до 2007 г. оставалось практически постоянным, и лишь потом начало происходить заметное увеличение количества индексируемых журналов (рис. 14).

Если считать, что научные журналы как явление в целом и их отбор для базы данных Web of Science Core Collection подчи-

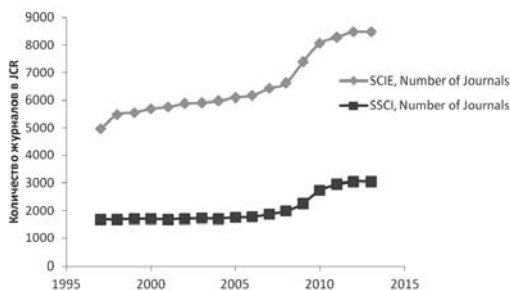


Рис. 14. Изменение количества журналов в Science Citation Index Expanded и Social Sciences Citation Index

няются закону Парето (который в наиболее общем виде формулируется как «20 % усилий дают 80 % результата, а остальные 80 % усилий — лишь 20 % результата»), то в настоящее время количество научных журналов можно оценить примерно в 70 тыс., включая журналы по гуманитарным наукам и искусству, которые индексируются в отдельной базе Arts & Humanities Citation Index, но для которых не рассчитываются наукометрические показатели в Journal Citation Reports.

Одновременно, естественно, возрастает и количество научных статей, опубликованных в этих журналах. На диаграмме (рис. 15) приведено изменение количества документов, проиндексированных в базах

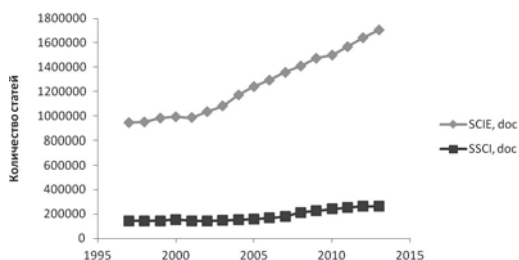


Рис. 15. Изменение количества статей в Science Citation Index Expanded и Social Sciences Citation Index

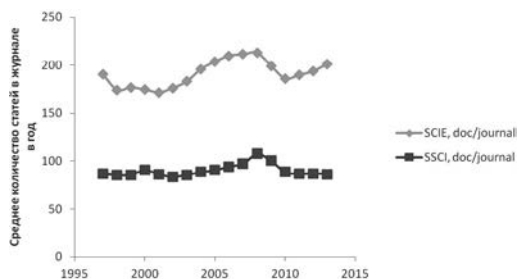


Рис. 16. Среднее количество статей в журнале в Science Citation Index-Expanded и Social Sciences Citation Index

данных Science Citation Index Expanded и Social Sciences Citation Index за этот же промежуток времени.

Интересно отметить, что динамика количества статей, опубликованных в среднем в одном журнале (рис. 16), по ключевым датам хорошо совпадает с изменением количества журналов: плавное увеличение их количества совпадает с постепенным нарастанием или стабильным количеством научных журналов, а резкое снижение количества статей в расчете на один журнал по времени совпадает с резким увеличением количества индексируемых в базе данных журналов (2008 г.). Таким образом, можно предположить, что новые журналы возникают тогда, когда предметный состав издаваемых журналов перестает соответствовать развитию науки и требуется создание новой дискуссионной площадки для обсуждения результатов научных исследований в новых возникающих научных областях.

Интересно отметить, что упомянутое выше распределение Парето абсолютно справедливо и к отбору журналов для ин-

дексирования в базах данных. Отмечается, что относительно небольшое количество журналов публикует значительную долю научных статей, получающих еще большую долю цитирований. Согласно анализу, проведенному по журналам, индексируемым Web of Science, получается соотношение, приведенное на рис. 17.

В книге Д. Прайса (которую можно посоветовать прочитать всем интересующимся историей науки и наукометрией) указаны следующие временные периоды удвоения показателей, характерные для различных сторон деятельности, так или иначе связанных с наукой:

«100 лет – число лиц, упоминаемых в национальных биографических справочниках;

50 лет – занятость (рабочая сила), население, число университетов;

20 лет – совокупный национальный продукт, важные научные открытия, число выдающихся физиков, число известных химических элементов, точность инструментов, число поступающих в колледжи на 1000 населения;

15 лет – бакалавры гуманитарных и точных наук, научные журналы, число членов научных обществ, число известных химических соединений, совокупное по всем отраслям число реферативных журналов;

10 лет – число известных астероидов, литература по неевклидовой геометрии, литература по рентгеновым лучам, литература по экспериментальной психологии....» и т. д. [Price, 1965].

Показательно, что срок в 15 лет, за который произошло удвоение количества журналов, полностью совпадает с указанным для журналов в работе Д. Прайса.

Данные OECD по количеству исследователей на 1000 работающих хорошо укладываются в приведенную схему (рис. 18).



Рис. 17. Иллюстрация неравномерного распределения статей и их цитирований по научным журналам

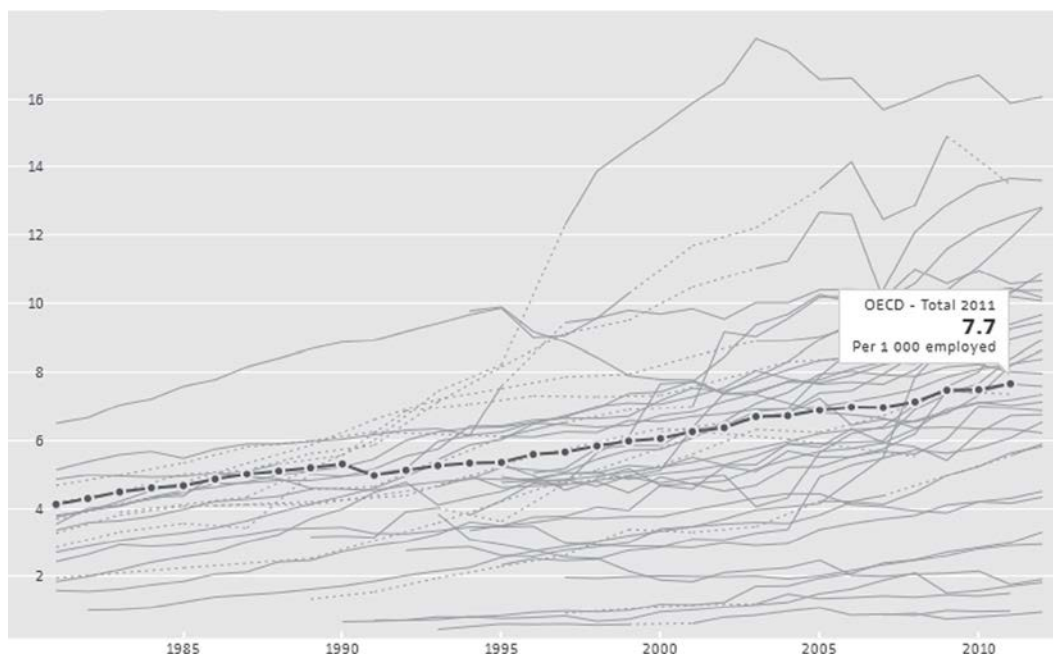


Рис. 18. Изменение количества научных работников по странам (количество занятых в науке на 1000 работающих). Жирным выделено среднее значение по странам OECD

Анализ данных OECD по абсолютному количеству исследователей в разных странах и сопоставление этих данных с количеством публикаций в Web of Science Core Collection показывает явную тенденцию к возрастанию количества публикаций в расчете на исследователя (рис. 19).

Тем не менее это не обязательно означает повышение «производительности» ученых, скорее это отражение тенденции к увеличению количества соавторов в статьях, отмеченной в статье [Börner et al., 2004] (рис. 20).

Постепенное возрастание среднего количества соавторов публикаций отмечалось и ранее. В среднем число работ с большим, чем 1, количеством авторов, обратно пропорционально $(n - 1)$, если n — количество соавторов [Прайс, Бивер, 1976].

Говоря о становлении научных журналов, нельзя не упомянуть о совершенно новом типе научных журналов, возникших с развитием информационных технологий и всемирной сети Интернет — о журналах открытого доступа. В целом их отличие от традиционных журналов заключается в том, что эти журналы существуют преимущественно в электронном виде и обеспечивают полный доступ к опубликованным статьям для всех пользователей сети Интернет. Первыми такими журналами стали журналы Public Library of Science (PLOS), однако количество журналов Open Access растет ежегодно, дополняясь возможностями публикации отдельных статей открытого доступа в традиционных журналах, а также публикациями в открытых

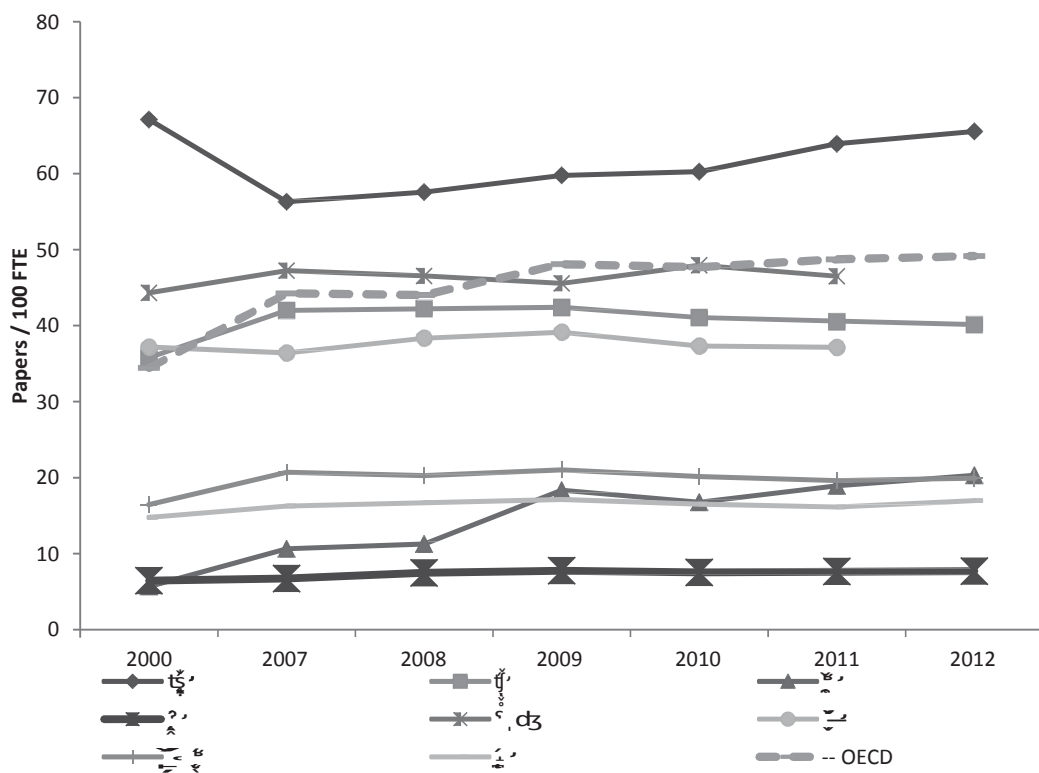


Рис. 19. Изменение количества статей в Web of Science на 100 научных работников по странам

архивах — репозиториях научных публикаций, наиболее известными из которых являются PubMed или arXiv.

В связи с этим встает вопрос «контроля качества» журналов открытого доступа, их соответствия общепринятым нормам публикационной этики и, соответственно, доверия к опубликованным в них материалам. Весьма интересным с этой точки зрения является подход библиотекаря из Университета Колорадо Джеффри Белла, анализирующего журналы открытого доступа и издательства по признакам соответствия публикационной этике [Beall, 2012].

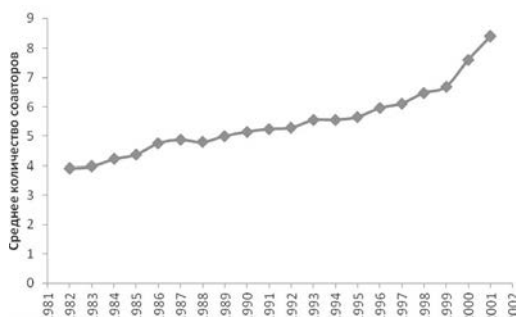


Рис. 20. Изменение количества соавторов в среднем в одной научной статье [Börner et al., 2004]

По результатам анализа цитирования журналов (в том числе самоцитирования, взаимного цитирования журналов, издаваемых одним издателем, нарушений правил рецензирования, недостоверных сведений о редакционной коллегии журнала и т. д.) составляется постоянно обновляемый «список Белла»¹, при этом попадание журнала в данный список достаточно часто является одним из оснований исключения журнала из международного индекса цитирования. Так, например, в 2013 г. из Journal Citation Report был исключен Life Science Journal, издаваемый находящимся в списке Белла издательством Marsland Press. В том же 2013 г. был поставлен своеобразный рекорд — за манипуляции с цитированиями из Journal Citation Report были исключены 66 журналов².

Как бы то ни было, количество научных публикаций в настоящее время уже так выросло, что ученый просто не в состоянии прочитать даже все публикации по своей узкой специальности, не то что отслеживать публикации в смежных областях знаний, что крайне необходимо при постоянно повышающейся междисциплинарности исследований.

3.2. Базы данных публикаций как инструменты поиска и анализа

Когда количество научных журналов в мире превысило 300, начали появляться первые реферативные журналы, облегчающие ученым поиск необходимой информации. В конце XIX в. появились первые указатели — индекс юридических документов Shepard's Citations и Index Medicus,

однако настоящая революция в этом плане произошла в 1960 г., когда Ю. Гарфилд организовал Институт научной информации, впервые начавший составлять машиночитаемый индекс научных публикаций с возможностью учета цитирований. На основе созданных баз данных научных публикаций стало возможным проводить масштабные исследования, связанные с анализом цитируемости.

Существующие библиотечные базы данных можно разделить на несколько типов:

- библиографические, содержащие только библиографические записи;
- реферативные, содержащие рефераты (аннотации) публикаций и иную дополнительную информацию о документе;
- полнотекстовые, содержащие, как следует из названия, полные тексты включенных в базу данных документов;
- комбинированные.

Указатели (индексы) цитирования научных статей представляют собой, как правило, реферативные базы данных, содержащие дополнительно количественную информацию о цитировании статей, извлеченную из приставных списков литературы.

База данных Science Citation Index, разработанная в Институте научной информации Ю. Гарфилдом, была запущена в 1964 г., чуть позже появились Social Sciences Citation Index (1973) и Arts & Humanities Citation Index (1978). В этих базах индексируются все без исключения типы публикаций в основных научных журналах, строго отбираемых по правилам, с которыми можно ознакомиться на сайте³. В 1990 г. началось составление Conference proceedings Citation Index — Science и Social Sciences and

¹<http://scholarlyoa.com>

²<http://blogs.nature.com/news/2013/06/new-record-66-journals-banned-for-boosting-impact-factor-with-self-citations.html>

³<http://wokinfo.com/russian/editors/>

Humanities. С 1997 г. эти базы данных стали доступны в сети Интернет на платформе ISI Web of Knowledge (с января 2014 г. — платформа Web of Science). С 2010 г. началось индексирование книг (Book Citation Index — Science и Social Science, с архивом начиная с 2005 г.). Эти 7 указателей, а также 2 химических индекса (Current Chemical Reactions и Index Chemicus) объединены в базу данных Web of Science Core Collection (до 2014 г. — Web of Science)^{4, 5}.

Эти базы данных являются реферативными, и структура их дает возможность проводить анализ и поиск по названиям публикаций, авторам, ключевым словам, аффилиации авторов, названиям журналов и конференций, предметным категориям и исследовательским областям. Списки процитированной литературы индексируются в отдельных базах данных, что позволяет по ним проводить независимый поиск и обнаруживать ссылки даже на те источники, индексация которых в Web of Science Core Collection не проводится.

Индексация как самих статей, так и при-статейных списков литературы позволяет осуществлять автоматический учет цитирований для всех проиндексированных документов. На основании этих данных рассчитываются показатели журналов, рассчитываются основные индикаторы по научным областям, становится возможным анализ показателей цитирования для любых наборов публикаций за любой период времени — для отдельных авторов, организаций, стран и т. д.

⁴См. подробное описание Web of Science Core Collection в главе 2. Библиометрические индикаторы в ресурсах Thomson Reuters.

⁵Подробнее об особенностях всех баз данных, размещенных на платформе Web of Science, см. далее в разделе 3.8. Важность источника информации о публикациях для оценки научной деятельности.

Кроме Web of Science Core Collection на платформе Web of Science расположены и другие ресурсы, доступ к которым зависит от подписки организации:

- Biological Abstracts;
- Biosis Citation Index;
- CAB Abstracts;
- Cab Global Health;
- Current Contents Connect;
- Data Citation Index;
- Derwent Innovations Index;
- Food Science Technology Abstracts;
- Inspec;
- Medline;
- Zoological Record.

Новым этапом в развитии платформы Web of Science стало размещение на ней региональных баз данных (указателей) научного цитирования. Региональные указатели научного цитирования индексируют наиболее значимые источники научных публикаций в своих странах или регионах. В отличие от Web of Science Core Collection, большинство публикаций, индексируемых в региональных указателях, опубликованы не на английском языке. В 2009 г. на платформе Web of Science был размещен указатель китайских научных публикаций Chinese Science Citation Database, в 2013 г. — указатель научной литературы на испанском и португальском языках SciELO Citation Index и в 2014 г. — корейский указатель KCI — Korean Citation Index. В сентябре 2014 г. компания Thomson Reuters объявила о начале работы по созданию, совместно с Научной электронной библиотекой eLIBRARY.RU, российского указателя научного цитирования Russian Science Citation Index, который также будет размещен на платформе Web of Science. В российском указателе научного цитирования

будут индексироваться лучшие научные журналы из России и стран СНГ.

По всем базам, имеющимся в доступе пользователей, можно осуществлять поиск информации как одновременно, так и по каждой базе в отдельности. При этом в некоторых из баз данных, размещенных на платформе Web of Science, учитывается цитирование (например, в BIOSIS или Data Citation Index), а в некоторых – нет (например, в Inspec или Medline).

All Times Cited Counts

23 in All Databases

19 in Web of Science Core Collection

19 in BIOSIS Citation Index

0 in Chinese Science Citation Database

4 in Data Citation Index

0 in SciELO Citation Index

Осуществив только один поиск информации по ключевым словам, например, исследователь получит не только список проиндексированных документов, но еще и огромное количество полезной для себя информации:

1. В какие годы наиболее активно проводились исследования по интересующей тематике?

2. В каких странах и исследовательских организациях исследуется проблема?

3. Какие исследователи наиболее активно работают в интересующей области?

4. Какие журналы публикуют работы, посвященные интересующей теме, и в каких журналах публикуют самые читаемые (цитируемые) работы?

5. Какие фонды поддерживают исследования по данной тематике?

6. На каких конференциях можно представить результаты своей работы?

Получив ответы на данные вопросы, исследователь сможет оптимальным образом спланировать свою работу, выстроить стратегию сотрудничества и стратегию публикаций.

Первичной единицей измерения во всех случаях является отдельно взятая статья, для которой можно измерить как количество ссылок в приведенном списке использованной литературы, так и впоследствии количество цитирований. Исследуя наборы статей для автора, журнала, научной области, организации, страны и т. д., можно вывести различные статистические параметры, характерные для исследуемых субъектов и позволяющие делать выводы и предсказания с использованием известных статистических законов. Как отмечал Д. Прайс, множество научных статей подобно газу, где каждая молекула ведет себя самостоятельно, но поведение газа в целом подчиняется вполне определенным законам и может быть легко предсказуемо [Pрайс, 1965].

3.3. Природа цитирования

В работах по истории и социологии науки ведется много споров о природе цитирования, о мотивации ученых при выборе статей, на которые они ссылаются в своих работах, о значимости тех или иных ссылок и их классификации. Подробный обзор работ, посвященных этой теме, можно найти в книге «Анализ цитирования в библиометрии» [Бредихин, Кузнецов, Щербакова, 2013]. С точки зрения научной этики наиболее взвешенной является нормативная теория цитирования, построенная на сформулированных в 1973 г. Р. Мертоном принципах научной этики [Merton, 1973]. Они заключаются в том, что 1) ученый путем

публикаций передает плоды своих трудов в общее пользование, 2) оценка научного значения не зависит от того, кто сделал научный вклад, т. е. независимость результатов научных исследований от национальности, классовой принадлежности или личных качеств ученого, 3) бескорыстность научной деятельности, т. е. недопустимость влияния на представленные результаты или оценку или умалчивание чужих результатов, способствующих собственной выгоде, 4) организованный скептицизм, т. е. критичность как к собственным, так и чужим результатам. Если основываться на этих постулатах, то ссылки в научных работах делаются для того, чтобы обозначить работы, являющиеся основой для излагаемого исследования, описывающие использованные методы исследования, связанные тематически и необходимые для обсуждения полученных результатов. Таким образом, исследователи цитируют те материалы, которые подтвердили для них свою ценность.

Кроме нормативной теории цитирования следует отметить также существование социальной конструктивистской теории цитирования [Latour, Woolgar, 1986], суть которой состоит в том, что цитирования являются способом убеждения читателя, для чего авторы часто ссылаются на работы авторитетных ученых вместо выбора наиболее релевантных ссылок для придания веса своей работе. Такой подход слабо соотносится с нормами научной этики и, по сути, априори обвиняет авторов в намеренной фальсификации, некорректном цитировании и умалчивании существующих данных.

Концепция гандикапа, на основе которой разработано обоснование цитирований Николайсеном [Nicolaisen, 2004], во многом перекликается с конструктивистской теорией цитирования, предполагая, что уве-

личивая количество ссылок в статье, автор придает своему тексту дополнительный вес, независимо от того, насколько ссылки по смыслу связаны с излагаемым материалом.

Рефлексивная теория цитирования [Wouters, 1998; Wouters, 1999] основывается на различной интерпретации терминов «библиографическая ссылка» и «цитирование». Цитаты рассматриваются в качестве индикаторов, создающих формализованное представление науки, и в первую очередь следует учитывать не причины, заставившие автора процитировать ту или иную работу, а то, как цитаты отражают характеристики науки. Воутерс предлагает определение науки как информационного цикла, состоящего из цикла экспертной оценки и цикла цитирований. Появление библиометрических индикаторов снижает роль экспертизы в оценке публикаций, поскольку научная публикация может измеряться с помощью анализа цитирования или семантического анализа (по совпадающим словам, словосочетаниям). Такой подход позволяет создать основу для формирования связей между публикациями и дальнейшего создания карт науки.

При рассмотрении различных теорий цитирования нельзя не упомянуть эффект Матфея, проанализированный социологом Р. Мертоном [Merton, 1968]. В общем случае это феномен неравномерного распределения преимуществ, в котором сторона, уже ими обладающая, продолжает их накапливать и приумножать, в то время как другая, изначально ограниченная, оказывается обделена еще сильнее и, следовательно, имеет меньше шансов на дальнейший успех. В применении к науке эффект Матфея заключается, во-первых, в том, что известные ученые получают высокое признание за исследования,

которые не всегда могут считаться значимыми, тогда как неизвестные специалисты за аналогичные результаты получают гораздо меньшее признание, во-вторых, в том, что работа, получившая признание, превращается в «прецедентный текст», воспринимающийся не с точки зрения его содержания, а с точки зрения его конституированного «значения».

В 90-х гг. группа под руководством М. Боница обнаружила явление, которое они назвали эффектом Матфея для стран [Bonitz et al., 1997; Bonitz et al., 1999; Bonitz, Scharnhorst, 2001], который заключается в том, что если ученые страны публикуются в престижных изданиях, то, скорее всего, они получают в среднем больше цитирований, чем предполагает средняя цитируемость журнала. Исследователями было введено понятие «индекса Матфея», представляющего собой частное от деления средней цитируемости авторов конкретной страны в журнале за вычетом средней цитируемости по журналу в целом на среднюю цитируемость журнала. Если индекс оказывается больше нуля, то страну цитируют выше нормы, и наоборот. В. В. Писляков в работе 2009 г. исследовал проявление эффекта Матфея для российских авторов на примере журналов, индексируемых в базе данных Web of Science в области физики и химии. Было показано значительное положительное влияние международного сотрудничества на цитируемость отечественных статей в зарубежных журналах [Писляков, Дьяченко, 2009].

Возможность сравнительно легкого учета цитирований статей после появления компьютерных баз данных приводит к соблазну для администраторов науки использовать их как показатели качества статей или эффективности работы ученых, научных кол-

лективов, организаций и т. д. Однако если вдуматься, то количество цитирований скорее является не показателем качества, а показателем полезности или влияния. Так, большое количество ссылок на методическую работу часто обусловлено большим распространением предложенного автором метода исследования, хотя саму статью большинство из цитирующих, возможно, и не читали – такие ссылки чаще всего являются формальными.

В связи с разными подходами возникает проблема о неравнозначности веса различных ссылок для оценки публикаций. В общем виде различные типы цитирований можно представить следующей схемой [Chubin, Moitra, 1975] (рис. 21):

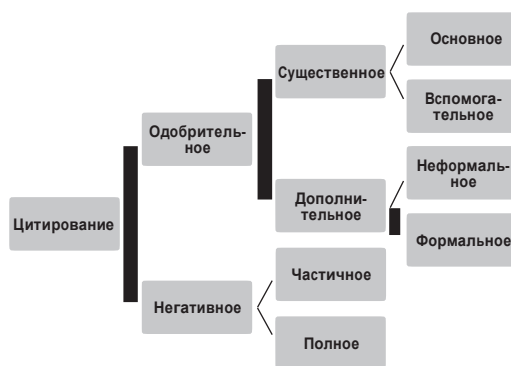


Рис. 21. Типы цитирований

Проведенный в соответствии с этой схемой анализ библиографических ссылок из статей по физике, опубликованных в период 1968–1969 гг., показал, что только 20 % ссылок были формальными и около 5 % — негативными. В любом случае наличие даже негативной ссылки означает определенный интерес к предмету исследования, т. е. ссылка любого типа ценна для анализа трендов развития научных направлений.

Поскольку характер цитирования можно определить только при проведении тщательного анализа статьи и цитируемых в ней материалов, то становится практически невозможным автоматический учет цитирований с приданием им, в зависимости от типа ссылки, какого-то веса. Поэтому при автоматическом определении индикаторов все ссылки и цитирования в настоящее время учитываются как равнозначные.

Отдельно следует упомянуть такой тип цитирования, как «самоцитирование». Это понятие имеет совершенно разный смысл и значение в применении к разным объектам. Если речь идет о самоцитировании автора, то причины этого могут быть либо в упоминании необходимых для понимания работы ранее опубликованных материалов автора, либо в желании «накрутить» таким образом свои показатели цитирования (что, по сути, является нарушением принципа бескорыстности, описанного выше). Еще более сложным является определение самоцитирования в статьях с несколькими соавторами. В отношении журнала «самоцитированием» является цитирование статей, опубликованных в том же журнале, что также может быть вполне оправданным и добросовестным, а может быть следствием попыток «накрутки» показателей журнала или сигналом того, что данный журнал читают только те, кто публикует там статьи (о различии между коэффициентами «самоцитируемости» и «самоцитирования» см. раздел 2.4). Примерно так же обстоит дело и в случае самоцитирований по отношению к научной организации. Нормальным считается во всех случаях самоцитирование в пределах 30 %, значительное превышение этой величины требует уже отдельного рассмотрения причин.

3.4. Анализ цитирования

Анализ цитирования не сводится к простому подсчету количества цитирований отдельной статьи или статей автора, организации и т. д., хотя даже это может дать достаточное количество информации для определения актуальности или важности исследований. Анализ цитирований позволяет установить связи между документами, на основании которых могут быть выделены кластеры работ, связанных по тематике, что является основой для составления карт науки. Связи между документами могут отслеживаться как ретроспективно (назад), так и проспективно (вперед). Каждая статья содержит список процитированной литературы и, в свою очередь, в дальнейшем может быть процитирована в других работах.

В основу метода «библиографического сочетания», предложенного Кесслером [Kessler, 1963], положен принцип выделения взаимосвязи между статьями на основании совпадений в списках цитируемой литературы. Чем больше таких совпадений, тем более тесно связаны между собой данные статьи. Подобный ретроспективный анализ возможно проводить независимо от языка публикаций, для него не требуется экспертная оценка, выявленные сочетания изменяются в зависимости от времени с появлением новых документов, доступных для анализа.

Похожий метод практически одновременно был предложен в СССР и США [Маршкова, 1973; Small, 1973]. Это метод анализа ко-цитирования, основанный на том, что между документами имеется связь, если эти документы цитируются в опубликованных позднее документах. Это уже проспективная связь, которая может быть прослежена не сразу после опубликования документов, а по мере появления цитирующих

эти документы публикаций. Если в случае ретроспективного анализа один раз установленная связь остается постоянной и могут только добавляться документы, принадлежащие группе (кластеру) публикаций, то в случае проспективного анализа могут появляться новые связи между ранее опубликованными документами. Как и в случае ретроспективного анализа цитирований, сила связи между документами зависит от количества работ, одновременно цитирующих рассматриваемые.

Описанные кратко методы анализа цитирований широко применяются при построении карт науки. Например, исследовательские фронты (research fronts), представленные в разделе базы данных Essential Science Indicators на платформе InCites, построены с использованием метода ко-цитирования.

Однако значительно более часто исследователи сталкиваются с более простыми методами учета и анализа цитирований, основанными, например, на подсчете полученных цитирований и дальнейшем их нормировании на количество процитированных документов.

Здесь следует обратить внимание на то, что цитирования различных документов и в разных областях знаний сильно различаются, поэтому некорректно проводить просто сравнение основанных на цитировании показателей без учета особенностей.

В журналах в настоящее время выделяют разные типы документов — статьи (Articles), обзоры (Reviews), короткие сообщения, письма, редакционные материалы, материалы конференций и т. д. Эти документы различаются как по длине списков цитируемой литературы (или их отсутствию, как в большинстве редакционных или новостных материалов), так и по частоте и длительности их цитирования.

Значительно различаются данные параметры и в различных областях знаний, причем здесь появляется еще и дополнительный фактор — отличия в типе цитируемых документов. Все это необходимо учитывать при сравнительном анализе показателей цитирования разных ученых, журналов, организаций и т. д.

Показатели цитирования можно рассматривать в применении к разным субъектам, и в каждом случае они будут иметь свои особенности, хотя методы определения могут быть одинаковыми или похожими.

3.5. Показатели научных журналов

По отношению к журналам впервые анализ цитирования был применен в начале XX в. для решения проблем комплектования библиотек периодическими журналами. Гросс и Гросс [Gross, Gross, 1927], исследуя библиографические ссылки в ведущем на тот момент издании по химии The Journal of the American Chemical Society, предложили использовать частоту цитируемости различных журналов для их ранжирования. Данное исследование стало отправной точкой для серии работ, которые привели к появлению метрики, получившей название «импакт-фактор» (т. е. фактор влияния). Эта метрика была разработана с целью выявления наиболее значимых журналов в конкретной области исследования. В 1955 г. вышла работа Юджина Гарфилда [Garfield, 1955], в которой изложен подход к вычислению метрики «импакт-фактор научных журналов», используемый до настоящего времени и представляемый ежегодно начиная с 1975 г. в отчетах по цитированию научных журналов (Journal Citation Reports), вышедших сначала в бумажном виде, а в настоящее время доступных на информационно-аналитической платформе InCites.

Подробное описание различных библиометрических индикаторов, характеризующих научные журналы, дано в главе «Библиометрические индикаторы в ресурсах Thomson Reuters», здесь же приведем информацию о других возможных способах оценки качества журналов, разработанных с целью избежать недостатков импакт-фактора и позволяющих сравнивать уровень журналов в разных научных дисциплинах.

SJR (Scimago journal rank) учитывает, как и Eigenfactor, влияние ссылок, т. е. ссылка из престижного журнала учитывается с большим весом, чем ссылка из менее престижного. Кроме того, самоцитирования журнала в этом расчете учитываются только на 33 % [González-Pereira, Guerrero-Bote, Moya-Anegón, 2010; Guerrero-Bote, Moya-Anegón, 2012]. Для расчета используется процедура, сходная с алгоритмом Google PageRank™ [Page et al., 1998]. Исходно каждому журналу присваивается ранг, равный 1, который в дальнейшем распределяется между журналами, в которых цитируются статьи данного журнала, пропорционально количеству полученных цитирований. Соответственно после определенного количества циклов таких перераспределений рангов максимальное значение показателя для дальнейшего распределения между журналами оказывается у самого влиятельного журнала, получающего максимальное количество цитирований. Соответственно одно цитирование из такого журнала оказывается значительно весомее, чем цитирование из менее престижного журнала, который только раздает цитирования, но не получает их взамен. При расчете этого показателя устанавливается окно цитирования 3 года. Полученный «престиж» для получения окончательного значения SJR нормируется на количество

опубликованных в журнале документов. Таким образом, данный показатель оказывается независим от области знаний и учитывает уровень журналов, из которых поступают цитирования.

Еще одним показателем, призванным компенсировать недостатки классического импакт-фактора, является SNIP (Source Normalized Impact per paper) [Moed, 2010]. Ключевым в расчете данного показателя является определение потенциала цитирования журнала. Это осуществляется путем определения индивидуального окружения журнала, т. е. для исследуемого журнала определяются все журналы, в которых за период в последние 10 лет встречаются цитирования статей, опубликованных в исследуемом журнале, и вычисляется средняя длина списков литературы в цитирующих статьях. При обработке данных по спискам литературы из них исключаются все ссылки, которые ведут за пределы базы данных, по которой рассчитываются показатели. Таким образом, учитываются различия в традициях цитирования в разных областях знаний. Вычисленный стандартным образом 3-летний импакт-фактор делится на потенциал цитирования, и таким образом получаем независимый от области знаний показатель рейтинговости журнала. Корректность определения потенциала цитирования зависит от полноты базы данных, на основании которой он рассчитывается: если статьи из исследуемого журнала цитируются в источниках, не попадающих в индексируемые, то потенциал цитирования окажется недооцененным, поэтому уровень покрытия научных областей в базе данных оказывается решающим фактором для расчета данного показателя. Для учета возможного недостаточного покрытия области для расчета

потенциала цитирования берутся ссылки только на журналы, присутствующие в базе данных, а ссылки на отсутствующие источники игнорируются.

Следует отметить, что для расчета показателей SJR и SNIP в числителе и знаменателе учитываются одинаковые типы документов (articles, conference proceedings papers и reviews), а расчет осуществляется по базе данных Scopus.

Знание различных наукометрических показателей журнала позволяет исследователю разработать для себя оптимальную стратегию публикаций, выбирая наиболее подходящие журналы как по научной специализации, так и по их престижности.

3.6. Показатели для ученых и организаций

Для оценки результативности научной деятельности ученых и организаций чаще всего используют такие простые показатели, как количество опубликованных работ, общее количество их цитирований и среднее количество цитирований на одну публикацию. Если сравнивать данные показатели за конкретный период времени и в одной и той же области, то можно более или менее адекватно построить рейтинг ученых, а дополнительное нормирование для организаций на количество ученых позволит построить и рейтинг организаций по их совокупной научной продуктивности. Однако в данном случае также имеется довольно сильная зависимость от области знаний, что не позволяет таким простым образом сравнивать ученых и организации из разных научных областей [Price, 1963].

Еще в работе Д. Прайса была отмечена общая закономерность для распределения авторов по количеству написанных статей, свидетельствующая о том, что 25 %

авторов производят 75 % статей. На рис. 22 представлено число публикаций для четырех групп высокопродуктивных авторов: 1) члены Национальной академии наук, 2) 19 выдающихся ученых XIX в., 3) наиболее продуктивные авторы-химики, 4) авторы по указателю «Ученых записок Королевского общества».

То же самое распределение мы получим, анализируя авторов, объединенных любым признаком. На рис. 33–35 представлено в аналогичном формате распределение российских авторов по научным областям (рис. 33, см. с. 185) и по городам (рис. 34, см. с. 186) по данным РИНЦ, а также распределение авторов по количеству публикаций по данным Core Collection Web of Science (рис. 35, см. с. 186) за 2000–2014 гг. (всего, по США, по предметной категории Biochemistry molecular biology — в мире, США и Калифорнийском университете).

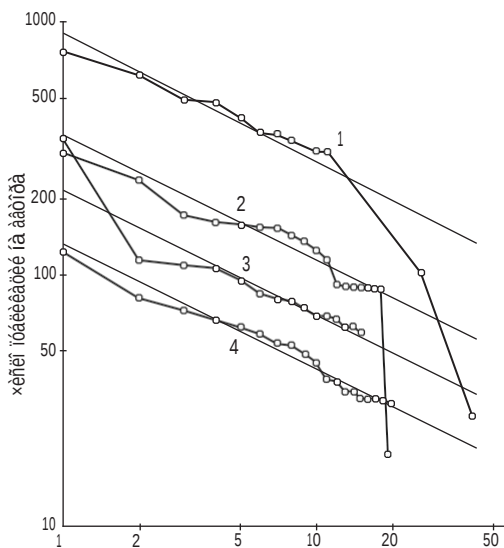


Рис. 22. Распределение авторов по количеству публикаций в разных научных областях [Прайс, 1963]

Таким образом, статистически распределение является общим для любой совокупности авторов и очень похоже на уже упоминавшийся закон Парето.

В работе В. В. Налимова и З. М. Мульченко [Налимов, Мульченко, 1969] — основополагающем труде по наукометрии — отмечается, что оценка деятельности ученых по суммарному числу публикаций приносит большой вред науке, поскольку из того, что талантливые ученые публиковали много работ, совершенно не следует обратное: что любой, опубликовавший много работ, является талантливым ученым. Следовательно, просто количество публикаций нельзя считать критерием эффективности ученого. Принимая то, что мерой полезности публикации является ее цитируемость, становится возможным производить сравнительную оценку работ разных ученых.

Широко распространенное возражение против оценки работ ученых по цитируемости их публикаций, связанное с учетом отрицательных цитирований, В. В. Налимов считает несущественным. Он отмечает, что «в науках, не имеющих острой политической направленности, не нужно проводить разграничения между положительным и отрицательным цитированием. Если работа цитируется, значит, высказанные в ней идеи послужили толчком для развития новых работ. И с этой точки зрения неважно, продолжают ли исходные идеи развиваться или они радикально переосмысливаются». Второе возражение — о возможном запаздывании признания некоторых работ — остается неснятым, однако и учесть это каким-либо образом оказывается невозможно.

В подтверждение возможности использовать критерий цитируемости в качестве оценки эффективности работы ученого

В. В. Налимов и З. М. Мульченко приводят работу американских исследователей, показавшую, что уровень цитируемости ученых-физиков хорошо согласуется с экспертной оценкой, проведенной представителями академического сообщества, на основании которой исследуемые ученые были разбиты на 4 группы — «плодотворно работающие ученые» (много печатаются и часто цитируются), «плодовитые физики» (много пишут, но редко цитируются), «физики-новаторы» (сравнительно мало печатаются, но каждая из работ много цитируется) и «молчаливые физики» (мало печатаются и редко цитируются).

Похожее исследование, позволяющее выделить группы ученых по их цитируемости, по анализу цитируемости академиков АН СССР по разным научным дисциплинам было проведено в МГУ. Российским же ученым Р. Ф. Васильевым было введено понятие средней частоты цитирования C/P (C — полное число ссылок на работы ученого, P — полное число его работ) и показано, что эта величина не определяется возрастом или известностью ученого.

Следует, однако, отметить, что все сравнительные исследования цитируемости ученых или средней частоты цитирования проводились в пределах одной и той же области знаний. Если обратиться к приведенной в разделе в базе данных Essential Science Indicators таблице среднего цитирования статей по областям знаний (Average Citation Rates for papers published by field, 2004–2014), то видно, насколько сильно различаются средние показатели цитируемости по разным областям.

В связи с этим для сравнения ученых разных специальностей необходимо применять более сложные подходы и алгоритмы, учитывающие существующие отличия

в среднем уровне цитируемости, точно так же, как это делается и для сравнения журналов из разных областей знаний.

Для выявления особо высокоцитируемых статей в разделе Essential Science Indicators есть таблица, позволяющая определить, что статья, опубликованная в конкретном году, относится к 0,01 % или 1 % наиболее цитируемых в конкретной области знаний (таблица процентиля). Соответственно по этому параметру, а именно по параметру попадания в определенный интервал процентиля, возможно осуществлять сравнение статей из разных областей знаний. Однако для сравнения наборов публикаций

(ученого, группы, организации, страны) такой подход является крайне трудоемким, поскольку, во-первых, напрямую применим к отдельным статьям, и, во-вторых, данные по процентилям представлены в данном разделе за последние 10 лет, т. е. ранее опубликованные статьи самостоятельно проанализировать нельзя.

При сравнении показателей цитируемости ученых даже в одной и той же области также возникает масса проблем. Предположим, что двое ученых имеют одинаковое количество публикаций и совокупно одинаковое количество цитирований, т. е. показатель средней частоты цитирования у них

InCites™ Essential Science Indicators™



THOMSON REUTERS

BASELINES-CITATION RATES

RESEARCH FIELDS	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	ALL YEARS
ALL FIELDS	22.11	20.36	18.25	16.24	13.84	11.56	8.99	6.19	3.47	1.03	0.09	10.84
AGRICULTURAL SCIENCES	16.36	15.10	13.59	11.50	9.15	7.53	5.82	3.89	2.14	0.63	0.07	7.30
BIOLOGY & BIOCHEMISTRY	32.36	29.35	26.10	22.72	19.59	16.52	12.50	8.51	4.71	1.37	0.12	15.90
CHEMISTRY	21.33	20.49	18.57	16.56	15.31	12.93	10.62	7.65	4.68	1.42	0.10	11.82
CLINICAL MEDICINE	25.88	24.13	21.34	18.01	15.23	12.64	9.60	6.53	3.59	1.03	0.10	12.10
COMPUTER SCIENCE	7.08	6.60	5.98	5.25	7.14	6.18	4.59	3.05	1.52	0.43	0.05	4.80
ECONOMICS & BUSINESS	17.14	15.58	13.38	10.79	8.27	6.62	4.76	2.92	1.38	0.40	0.07	6.63
ENGINEERING	10.26	9.21	8.98	8.38	7.17	6.52	5.11	3.59	1.94	0.59	0.07	5.37
ENVIRONMENT/ECOTOLOGY	25.92	24.23	21.03	18.45	15.69	12.50	9.74	6.74	3.67	0.99	0.11	11.58
GEOSCIENCES	20.97	19.31	17.87	14.97	13.22	11.27	8.44	6.02	3.22	0.98	0.12	10.31
IMMUNOLOGY	37.76	35.83	31.31	27.97	23.62	19.68	14.81	10.14	5.42	1.53	0.13	18.70
MATERIALS SCIENCE	14.69	14.44	13.37	13.39	11.55	10.23	8.58	6.18	3.67	1.12	0.10	8.62
MATHEMATICS	7.96	7.13	6.43	5.72	4.97	4.07	3.16	2.08	1.13	0.34	0.05	3.73
MICROBIOLOGY	31.12	28.98	25.33	22.04	18.48	15.59	12.20	7.76	4.26	1.25	0.09	14.64
MOLECULAR BIOLOGY & GENETICS	51.52	46.19	41.70	37.52	30.49	25.38	19.00	13.05	6.72	1.85	0.12	24.15
MULTIDISCIPLINARY	72.04	67.16	63.29	54.36	45.43	40.98	32.37	19.41	10.18	2.74	0.18	33.56
NEUROSCIENCE & BEHAVIOR	34.70	32.14	28.39	25.48	21.17	17.54	13.40	9.08	4.93	1.41	0.13	16.90
PHARMACOLOGY & TOXICOLOGY	24.48	22.38	21.04	18.18	15.60	12.64	9.76	6.55	3.68	1.05	0.09	11.85
PHYSICS	17.39	16.32	14.71	13.48	12.57	10.83	8.71	6.22	3.86	1.20	0.09	9.88
PLANT & ANIMAL SCIENCE	17.31	15.63	14.12	11.90	10.10	8.48	6.57	4.43	2.40	0.73	0.09	8.25
PSYCHIATRY/PSYCHOLOGY	26.78	23.82	20.66	17.85	14.51	11.28	8.44	5.44	2.72	0.78	0.10	11.23
SOCIAL SCIENCES, GENERAL	14.19	13.06	11.57	9.73	7.53	5.97	4.45	2.85	1.49	0.45	0.07	5.82
SPACE SCIENCE	26.39	26.01	25.29	22.73	18.99	17.84	14.53	10.64	6.38	2.29	0.19	15.98

также будет совпадать. Однако распределение статей по количеству цитирований различается — у одного ученого из 50 публикаций с общим количеством цитирований 500 есть несколько очень высокоцитируемых публикаций при низком цитировании (или их отсутствии) остальных публикаций, а у другого цитирования распределены между статьями более равномерно. Как определить, какой из авторов работает эффективнее?

3.6.1. Индекс Хирша

В 2005 г., решая проблему выбора кандидатов на должность профессора физики, американский физик Хорхе Хирш из Университета Сан-Диего предложил ввести новый индикатор, который призван учитывать не только количество статей и цитирований, но и распределение статей по количеству полученных цитирований [Hirsch, 2005]. Этот индекс получил название индекса Хирша (*h*-индекс) и представляет собой количественную характеристику продуктивности ученого, основанную на распределении цитирований его работ. Подробное описание индекса Хирша и некоторых «хирш-подобных» показателей уже приведено в главе «Библиометрические индикаторы в ресурсах Thomson Reuters», поэтому в данном разделе сосредоточимся на особенностях использования этого индикатора, а также не упоминавшихся ранее его вариациях.

Первые по порядку *h* публикации, т. е. наиболее цитируемые публикации, составляют ядро Хирша. В своей работе Хирш исследовал публикации американских физиков и установил, что индекс Хирша порядка 10–12 соответствует постоянной исследовательской позиции в крупном университете, 15–20 — членству в Амери-

канском физическом обществе, а индекс 45 и выше — членству в Национальной академии наук США.

Следует отметить, что индекс Хирша, хоть и достаточно хорошо отражает уровень ученого, в сильной степени зависит от области знаний и других факторов, как, впрочем, и практически все наукометрические показатели. Если посмотреть индексы Хирша ученых разных специальностей, входящих в перечень наиболее цитируемых ученых по версии ресурса [Highlycited.com](http://highlycited.com), то получим следующую картину (см. табл. 4).

Таким образом, сравнивать эффективность работы ученых по индексу Хирша можно только в пределах одной и той же области знаний. Кроме того, априори понятно, что величина индекса Хирша ограничена сверху общим количеством публикаций ученого, т. е. при наличии *N* публикаций, пусть даже и очень высокоцитируемых, индекс Хирша никак не может оказаться выше *N*.

Интуитивно понятно, что этот индекс зависит и от длительности работы ученого, т. е. от его «академического возраста». На следующем графике, составленном по данным популярного российского ресурса «Кто есть кто в российской науке»⁶, хорошо видна тенденция к снижению индекса Хирша у ученых с меньшим «академическим возрастом», который здесь определяется по дате первой публикации в Web of Science CC (рис. 23).

Индекс Хирша может быть определен для любого набора статей — для отдельного ученого, для организации, для журнала и т. д. Во всех этих случаях не учитывается вклад в общее цитирование самых высокоцитируемых публикаций, т. е. если в хирш-ядро у разных авторов входят одинаковые

⁶<http://www.expertcorps.ru/science/whoiswho/>.

Таблица 4

Индексы Хирша ученых разных специальностей

Предметная область	Индекс Хирша наиболее цитируемого ученого
Biology & Biochemistry	114
Neuroscience	93
Plant & Animal Science	69
Physics	63
Geoscience	56
Mathematics	47
Computer Science	26

количество публикаций, то их суммарное цитирование не оказывает никакого влияния на собственно индекс Хирша. К существенным недостаткам индекса Хирша относят также то, что он не учитывает количество соавторов в публикациях (что весьма существенно в связи с постоянным ростом количества соавторов и появлением статей, количество авторов в которых измеряется тысячами), не может убывать во времени (т. е. фактическое прекращение активной научной деятельности никак не влияет на изменение уже достигнутого показателя), является по определению целым числом (снижается точность измерения), не учитывает тип документа, особен-

ности цитирования в различных научных областях и т. д.

Для компенсации указанных недостатков к настоящему времени разработано уже около сотни различных вариаций индекса Хирша. Приведем здесь только некоторые из наиболее часто используемых.

Первые вариации были предложены самим Хиршем в исходной статье. С целью учета «выбросов», которые дают публикации с большим количеством цитирований, а также публикации с меньшим числом цитирований, им был предложен *a*-индекс, являющийся отношением количества всех цитирований автора ($N_{c,tot}$) к квадрату индекса Хирша:

$$a = N_{c,tot}/h^2.$$

Предложенный в той же статье *m*-индекс дает возможность учесть длительность карьеры ученого и представляет собой частное от деления индекса Хирша на количество лет научной карьеры, например, от начала публикаций. Это оказывается более или менее верно при условии достаточно равномерной научной карьеры со стабильным количеством публикаций примерно одинакового качества.

Недооценку цитирований в *h*-ядре предлагается устранить путем введения комби-

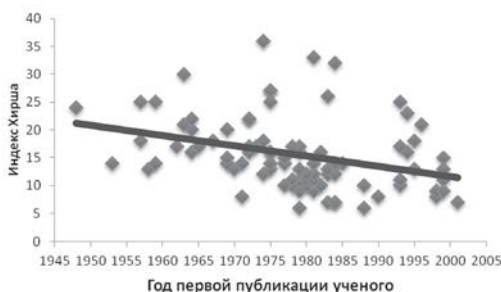


Рис. 23. Зависимость индекса Хирша от «академического возраста» ученого

нированного hg-индекса [Alonso et al., 2010], представляющего собой геометрическое среднее h-индекса и g-индекса: $hg = \sqrt{hg}$. К основным достоинствам hg-индекса можно отнести то, что он, во-первых, обеспечивает большую степень детализации, чем h и g, позволяя эффективнее оценивать ученых, например в случае с одинаковым значением h; во-вторых, смягчает влияние часто цитируемых работ и обеспечивает лучший баланс между влиянием большинства лучших работ и небольшого количества высокоцитируемых работ.

Для учета цитируемости статей, входящих в хирш-ядро, предлагается также дополнительно рассчитывать среднее число ссылок на статьи, входящие в хирш-ядро, или определять медиану числа цитирований h статей, входящих в хирш-ядро публикаций автора.

Количество соавторов публикаций возможно учитывать путем расчета следующих показателей:

- Individual h-index (original) — результат деления стандартного h-индекса на среднее число авторов в статьях, которые входят в хирш-ядро публикаций. Этот показатель призван уменьшить влияние на h-индекс числа соавторов публикаций, которое, по статистике, существенно отличается в различных областях знаний.

- Individual h-index (PoP variation) — вычисление h-индекса, когда вместо полного числа цитирований каждой статьи используется отношение числа цитирований к числу авторов публикации.

Существует также очень интересная версия иерархического индекса Хирша, который используется для оценки публикационной продуктивности на разных уровнях, т. е. для оценки совокупности исследователей, организации, страны, журнала и т. д.

Так, I-индекс [Prathap, 2006] определяется путем упорядочивания по убыванию индексов Хирша ученых одной организации, и по тому же принципу определяется показатель для организации, т. е. I-индекс показывает, что научная организация имеет индекс I, если не менее I ученых из этой организации имеют индекс Хирша не менее I. Аналогичным образом на следующем уровне можно отранжировать все научные организации страны и т. д.

3.7. Классификаторы, используемые в науке

Поскольку многочисленные исследования показывают значительную зависимость различных наукометрических показателей от конкретной научной области, необходимо определить, что же это такое и как провести границу между различными научными областями. Например, как отличить физическую химию от химической физики и что такое математическая биология — больше биология или все-таки математика?

Эта проблема теснейшим образом связана с созданием рубрикаторов, в первую очередь рубрикаторов научно-технической информации. В России эта проблема встала довольно остро при расширении издания реферативных журналов. В 1950-е гг. в информатике рубрикаторами стали называть перечни рубрик реферативных журналов (РЖ) и других периодических информационных изданий. В данном случае рубрика выполняет более специальную функцию, выступая как содержательный фрагмент РЖ или периодического библиографического указателя, и состоит из индекса и заголовка раздела, а также библиографических записей (с аннотациями или рефератами) тех произведений печати, которые по своему содержанию относятся к данной рубрике. Подробно история

создания рубрикатора, теория и практические применения описаны в книге «Рубрикатор как инструмент информационной навигации» [Гиляревский и др., 2008].

Основная задача рубрикатов состоит в удовлетворении потребностям двух разных взаимосвязанных информационных потоков — потока научных публикаций и потока информационных запросов. Эти два потока развиваются параллельно, но не синхронно, причем изменение структуры потока запросов почти всегда опережает изменение структуры документального потока. В основном используются различного вида систематические каталоги, построенные либо по иерархическому принципу, либо по фасетному типу.

Все основные библиотечные классификации строятся на основе иерархического «древа знаний» с выделением специальных разделов и подразделов для систематизации особых видов книг. Этот принцип применяется еще с античных времен. В Средние века классификация основывалась на системе «семи свободных искусств» и состояла из двух комплексов наук, изучавшихся в тогдашней школе: «тривиума» (грамматика, диалектика и риторика) и «квадриума» (арифметика, геометрия, музыка и астрономия). К концу XV в. в университетских библиотеках начали применять группировку книг в соответствии с существовавшими факультетами — философским, медицинским, юридическим и богословским, что стало началом возникновения так называемых факультетских систем классификации. Среди разрабатываемых позднее следует упомянуть одну из лучших классификаций первой половины XIX в. — классификацию К. Э. Бэра, в которой была предпринята одна из первых попыток расположить науки в последовательности, отражающей исто-

рию развития мира: науки о неорганической природе, науки об органической природе, науки о человеке и обществе.

В 1876 г. американским библиотечным деятелем М. Дьюи была предложена «десятичная классификация», которая создала базис для дальнейшей разработки различных систем автоматизации при классификации и поиске необходимой информации. Он теоретически обосновал и практически внедрил стандартизацию типовых делений (литературная форма, вид издания и т. п.) в различных разделах схемы и частично лингвистических, этнических и географических делений, используя прием факультативного превращения в постоянные подразделения окончаний индексов разделов «Филология» и «История». В 1905–1907 гг. на базе идей Дьюи была создана «Универсальная десятичная классификация» (УДК). Это пример классической иерархической классификации, обладающей следующими характерными чертами:

- универсальность, заключающаяся в охвате всех отраслей знания;
- логическая ступенчатая индексация, позволяющая неограниченно делить подклассы без нарушения основной структуры классификации;
- международная применимость благодаря использованию только цифровых десятичных индексов, всем понятных и легко запоминаемых;
- развитая система определителей и комбинационного построения индексов, обеспечивающих относительную гибкость при отражении достаточно узких и сложных понятий;
- устойчивый и четко организованный международный механизм поддержания классификации на уровне новых достижений науки.

Фасетная классификация была предложена индийским библиотековедом Ш. Р. Ранганатаном в 1933 г. Создание фасетных классификаций в первую очередь было связано с потребностью читателей отыскивать документ по специальной теме «не только тогда, когда именно она является непосредственным объектом поиска, но также тогда, когда поиск ведется по любому термину или группе терминов, входящих в сложное понятие. Для удовлетворения этих требований необходимо, чтобы не только понятия могли входить в неограниченное количество сочетаний, но также, чтобы в структуре системы были отражены родовые связи понятий и связи между разделами» [Ранганатан, 1970]. В качестве примера можно привести построенные по системе Ш. Р. Ранганатана фасеты и фокусы из области медицины (табл. 5).

По этому фрагменту фасетной классификации индекс документа по диагностике инфекционных заболеваний кишечника — 25:42:3, по лечению туберкулеза легких — 45:421:4.

Предметные рубрикаторы распределяют документы по предметам или понятиям, не

соотнося их с какими-либо областями знания. Предметный и классификационный принципы организации документов, таким образом, дополняют друг друга и облегчают поиски документов по разным типам запросов.

Наиболее распространенный в России Государственный рубрикатор научно-технической информации (ГРНТИ) представляет собой совмещение иерархического и фасетного классификаторов. Первые два уровня ГРНТИ представляют собой типичную иерархическую классификацию, например:

- 02.00.00 Философия
- 02.01.00 Общие вопросы философии
- 02.11.00 Общие проблемы современной философии и т. д.

На третьем уровне классификации появляются черты фасетной классификации: например, во всех рубриках код типа **.01.45 означает «Преподавание», а **.01.79 — «Кадры» по соответствующей области.

Классификация документов по ГРНТИ осуществляется, в частности, в Российском индексе научного цитирования.

Еще один рубрикатор, построенный по иерархическому принципу, часто используемый

Таблица 5

Фрагмент фасетной классификации по медицине

ФАСЕТЫ		
Органы тела	Проблемы медицины	Уход и лечение
1. Органы в целом	1. Общие проблемы	1. Питание
2. Органы пищеварения		
2.3. Пищевод	2. Морфология	2. Этиология
2.4. Желудок		
2.5. Кишечник		
3. Кровеносная система	3. Физиология	3. Диагностика
4. Органы дыхания	4. Болезни	4. Лечение
4.5. Легкие	4.2. Инфекционные	
	4.2.1. Туберкулез	

для анализа научной деятельности, — это номенклатура специальностей научных работников. В основе этой номенклатуры лежит дисциплинарная классификация наук по их предметному содержанию, и именно эта классификация представлена в иерархическом виде. При этом вид присуждаемой степени, являющийся второй системой классификации в данной номенклатуре, не имеет своего кодового обозначения, а характеризуется только словесно. Так, например, в области «математических наук» исследование характеризуется приоритетом точных математических методов. В «технических науках» специфика заключается в практическом приложении и внедрении полученных результатов в хозяйственную практику. В области «педагогических наук» специфика заключается в главной задаче — в распространении знаний в социальной среде.

Классификация наук в номенклатуре специальностей научных работников двух- или трехуровневая, причем трехуровневая система есть не во всех специальностях. В действующей в настоящее время номенклатуре есть еще и отсутствовавшая ранее надуровневая классификация, выделяющая группы специальностей — «Гуманитарные науки», «Искусствоведение и культурология», «Социально-экономические и общественные науки».

Примерная схема классификации приведена в табл. 6. Подробно до третьего уровня в таблице приведены только специальности в области математики.

Номенклатура ВАК через ГРНТИ сопрягается с международной классификацией знания УДК, однако установить полное соответствие оказывается весьма затруднительно.

Из наиболее часто употребляемых международных классификаций необходимо упо-

мянуть классификацию научных областей Организации экономического сотрудничества и развития, которая в последнее время используется все чаще для сравнительного анализа развития науки в разных странах (Frascati Fields of Science). В качестве верхнего уровня классификации в ней выделено шесть основных областей:

1. Естественные науки (Natural sciences).
2. Техника и технология (Engineering and technology).
3. Медицинские науки (Medical sciences).
4. Сельскохозяйственные науки (Agricultural sciences).
5. Общественные науки (Social sciences).
6. Гуманитарные знания (Humanities).

В последней версии этого рубрикатора, утвержденной в 2006 г., на втором уровне указанные области делятся еще на 42 уровня, которые, в свою очередь, детализированы еще больше. Основной принцип данного рубрикатора состоит в попытке сделать классификацию, в максимальной степени совместимую с имеющимися международными и национальными системами классификации знаний для гарантии сопоставимости при сравнении данных, касающихся Research & Development.

Сопоставимость различных рубрикаторов является одной из основных проблем при проведении наукометрических исследований, поскольку затрудняет сравнение данных, полученных из разных источников. Каждая из существующих баз данных пользуется своим рубрикатором, а нередко использует и несколько рубрикаторов для разных целей.

Для примера рассмотрим используемые для разных целей классификаторы различных ресурсов, размещенных на платформе Web of Science.

Таблица 6

Схема классификации

Шифр	Отрасль науки, группа специальностей, специальность		
01.00.00	Физико-математические науки	Математика	
01.01.00			
01.01.01			Вещественный, комплексный и функциональный анализ
01.01.02			Дифференциальные уравнения, динамические системы и оптимальное управление
01.01.03			Математическая физика
01.01.04			Геометрия и топология
01.01.05			Теория вероятностей и математическая статистика
01.01.06			Математическая логика, алгебра и теория чисел
01.01.07			Вычислительная математика
01.01.09			Дискретная математика и математическая кибернетика
01.02.00		Механика	
01.03.00		Астрономия	
01.04.00		Физика	
02.00.00	Химические науки		
03.00.00	Биологические науки		
05.00.00	Технические науки		
06.00.00	Сельскохозяйственные науки		
14.00.00	Медицинские науки		
25.00.00	Науки о земле		
	Гуманитарные науки		
07.00.00	Исторические науки и археология		
10.00.00	Филологические науки		
09.00.00	Философские науки		
	Искусствоведение и культурология		
17.00.00	Искусствоведение		
24.00.00	Культурология		

Таблица 6 (продолжение)

Шифр	Отрасль науки, группа специальностей, специальность		
	Социально-экономические и общественные науки		
19.00.00	Психологические науки		
08.00.00	Экономические науки		
13.00.00	Педагогические науки		
22.00.00	Социологические науки		
12.00.00	Юридические науки		
23.00.00	Политология		

Практически во всех базах, размещенных на платформе, используется классификация Research Areas (Web of Science CC, BIOSIS, Current Contents, Medline, National Citation Indices, Zoological Records). Во многих используются Web of Science Categories (Web of Science CC, Data Citation Index). Кроме этого, в разных базах используются и свои собственные рубрикаторы – Major Concepts в BIOSIS, Subject Areas в Data Citation Index и Derwent Innovation Index, MeSH Headings и MeSH Qualifiers в Medline, Disciplines в Current Contents и т. д.

Наиболее часто используемыми для наукометрического анализа оказываются, таким образом, Research Areas и Web of Science Categories, поскольку один из этих рубрикаторов (или оба) присутствуют во всех ресурсах, расположенных на платформе Web of Science. Еще один классификатор – Research Fields – используется в аналитическом ресурсе Essential Science Indicators.

В связи с этим рассмотрим три варианта классификации:

- Research Areas (152);
- Web of Science Categories (251);
- Research Fields (22).

Эти классификации используются для разных целей, и полного соответствия между ними установить практически невозможно.

Для классификации Research Areas используется иерархический принцип, в котором все 151 исследовательские области разбиты на 5 крупных направлений:

- Arts & Humanities;
- Life Sciences & Biomedicine;
- Physical Sciences;
- Social Sciences;
- Technology.

Классификация Web of Science Categories (предметной категории) в первую очередь используется для определения тематической принадлежности журнала в Journal Citation Reports. 28 категорий, относящихся к журналам по гуманитарным наукам и искусству указателя Arts & Humanities, в Journal Citation Reports, отсутствуют, т. к. для них не производится расчет импакт-факторов. Если осуществить поиск всех статей конкретного журнала в журнальных индексах Web of Science CC, то получим полное совпадение указанных в описании журнала предметной категории (или нескольких, если указана не одна) и предметной категории, к которой относятся все опубликованные в журнале статьи. Однако обратный поиск по предметной категории дает совершенно другие результаты – поиск по всем журнальным индексам по определенной предмет-

ной категории даст в результате перечень журналов, сильно отличающийся от того, который мы увидим при аналогичном поиске по предметной категории в Journal Citation Reports.

Поясним на примере.

Журнал PLoS ONE (один из наиболее престижных журналов открытого доступа) в Journal Citation Report отнесен к Multidisciplinary Sciences, где представлено всего 55 журналов, включая знаменитые Science и Nature. Все статьи этого журнала, представленные в Web of Science CC, относятся именно к Multidisciplinary Sciences, как и статьи из Science и Nature

и всех остальных 55 журналов, относящихся в соответствии с Journal Citation Reports к мультидисциплинарным. Однако поиск публикаций по признаку предметной категории Multidisciplinary Sciences приведет к обнаружению публикаций из почти сотни журналов. Однако это, скорее всего, связано с тем, что далеко не все журналы, индексируемые в Web of Science, присутствуют в Journal Citation Reports. Поиск по названиям журналов, отсутствующих в JCR, но появляющихся в списке источников статей категории Multidisciplinary Sciences, дает результат, аналогичный поиску по мультидисциплинарным журналам из JCR.

[Home](#)
[Journal Rankings](#)
[Journal Profile](#)
[Journal Profile](#)

CELL

ISSN: 0092-8674

CELL PRESS
600 TECHNOLOGY SQUARE, 5TH FLOOR, CAMBRIDGE, MA 02139
USA

[Go to Journal Table of Contents](#)
[Go to Ulrich's](#)

Titles

ISO: Cell
JCR Abbrev: CELL

Categories

BIOCHEMISTRY & MOLECULAR
BIOLOGY - SCIE;
CELL BIOLOGY - SCIE;

Languages

ENGLISH

26 Issues/Year;

Results: 17,834
(from Web of Science Core Collection)

You searched for: PUBLICATION
NAME: (Cell) ...More

Create Alert

Refine Results

Search within results for...

Web of Science Categories

☐ CELL BIOLOGY (17,813)
☐ BIOCHEMISTRY MOLECULAR BIOLOGY (17,813)

more options / values...

Sort by: **Publication Date -- newest to oldest**

☐ Select Page

Cit

- ☐ **SnapShot: Nuclear RNAPII Transcript Modification**
 By: Schmid, Manfred; Jensen, Torben Heick
 CELL Volume: 157 Issue: 5 Article Number: 1244.e1 Published: MAY 22 2014
- ☐ **Robin Holliday 1932-2014 Obituary**
 By: Holloman, Bill
 CELL Volume: 157 Issue: 5 Pages: 1001-1003 Published: MAY 22 2014
- ☐ **Die Another Day: A Painless Path to Longevity**
 By: Steculorum, Sophie M.; Bruening, Jens C.
 CELL Volume: 157 Issue: 5 Pages: 1004-1006 Published: MAY 22 2014

Для журналов, отнесенных в JCR к другим предметным категориям, наблюдается точно такая же зависимость, причем если журнал там отнесен к нескольким категориям, то те же категории мы увидим и в поисковом меню в Web of Science.

Тематические рубрикаторы Web of Science Categories и Research Areas являются журнальными рубрикаторами. То есть если индексируемый журнал (или иной источник публикаций, например материалы конференции) отнесен к одной или нескольким рубрикам Web of Science Categories и Research Areas, то все статьи и прочие публикации данного журнала (или иного источника) будут также отнесены к этим рубрикам. Рубрикаторы Web of Science Categories являются более дробными и используются только в Web of Science Core Collection. Research Areas являются более широкими рубрикаторами и могут как совпадать с Web of Science Categories, так и объединять несколько рубрик. Research Areas используются во всех базах данных на платформе Web of Science. Их наличие позволяет вести единый поиск по тематическим рубрикам.

В качестве примера сопоставления рубрикаторов можно рассмотреть результаты анализа статей 2012 г., найденных по разделу Evolutionary Biology из перечня Research Areas (см. табл. 7).

Аналогичный поиск по разделу Research Area Mathematics (su = mathematics) приведен в табл. 8. Отличие в данном случае в том, что названия раздела Evolutionary Biology в Web of Science Categories и Research Areas полностью совпадают, а в области математики такое совпадение отсутствует.

В аналитическом ресурсе Essential Science Indicators все материалы представлены в соответствии с рубрикацией по Research Fields:

- AGRICULTURAL SCIENCES
- BIOLOGY & BIOCHEMISTRY

- CHEMISTRY
- CLINICAL MEDICINE
- COMPUTER SCIENCE
- ECONOMICS & BUSINESS
- ENGINEERING
- ENVIRONMENT/ECOLOGY
- GEOSCIENCES
- IMMUNOLOGY
- MATERIALS SCIENCE
- MATHEMATICS
- MICROBIOLOGY
- MOLECULAR BIOLOGY & GENETICS
- MULTIDISCIPLINARY
- NEUROSCIENCE & BEHAVIOR
- PHARMACOLOGY & TOXICOLOGY
- PHYSICS
- PLANT & ANIMAL SCIENCE
- PSYCHIATRY/PSYCHOLOGY
- SOCIAL SCIENCES, GENERAL
- SPACE SCIENCE

На архивной версии сайта ScienceWatch.com⁷ представлено определение исследовательских областей для Essential Science Indicators как обусловленных дискретным набором журналов. Каждая из областей (за исключением Arts & Humanities, для которых основные научные индикаторы не рассчитываются) представлена набором журналов по узким научным направлениям. Двадцать две широкие области являются, таким образом, «верхим» уровнем для Web of Science Categories, по которым классифицируются журналы. Те журналы, которые попадают в Journal Citation Report в разные категории, при распределении по широким областям могут либо оказаться в нескольких категориях, либо попасть в ту область, из которой получают максимальное количество цитирований. В первую очередь это касается журналов, которые относятся

⁷<http://archive.sciencewatch.com/about/met/fielddef/>.

Таблица 7

**Результаты анализа статей 2012 г., найденные по разделу Evolutionary Biology
из перечня Research Areas**

Web of Science Categories	records	% of 7021	Research Areas	records	% of 7021
ANTHROPOLOGY	1381	19.670	ANTHROPOLOGY	1381	19.670
BIOCHEMISTRY MOLECULAR BIOLOGY	1578	22.475	BIOCHEMISTRY MOLECULAR BIOLOGY	1578	22.475
BIODIVERSITY CONSERVATION	39	0.555	BIODIVERSITY CONSERVATION	39	0.555
BIOLOGY	930	13.246	LIFE SCIENCES BIOMEDICINE OTHER TOPICS	930	13.246
CELL BIOLOGY	33	0.470	CELL BIOLOGY	33	0.470
DEVELOPMENTAL BIOLOGY	165	2.350	DEVELOPMENTAL BIOLOGY	165	2.350
ENTOMOLOGY	61	0.869	ENTOMOLOGY	61	0.869
ECOLOGY	3050	43.441	ENVIRONMENTAL SCIENCES ECOLOGY	3050	43.441
EVOLUTIONARY BIOLOGY	7021	100.000	EVOLUTIONARY BIOLOGY	7021	100.000
GENETICS HEREDITY	2147	30.580	GENETICS HEREDITY	2147	30.580
MATHEMATICAL COMPUTATIONAL BIOLOGY	40	0.570	MATHEMATICAL COMPUTATIONAL BIOLOGY	40	0.570
MEDICINE RESEARCH EXPERIMENTAL	21	0.299	RESEARCH EXPERIMENTAL MEDICINE	21	0.299
MULTIDISCIPLINARY SCIENCES	7	0.100	SCIENCE TECHNOLOGY OTHER TOPICS	7	0.100
NEUROSCIENCES	23	0.328	NEUROSCIENCES NEUROLOGY	23	0.328
PALEONTOLOGY	69	0.983	PALEONTOLOGY	69	0.983
PHYSIOLOGY	66	0.940	PHYSIOLOGY	66	0.940
PLANT SCIENCES	472	6.723	PLANT SCIENCES	472	6.723
ZOOLOGY	195	2.777	ZOOLOGY	195	2.777

Таблица 8

**Результаты анализа статей 2012 г., найденные по разделу Mathematics
из перечня Research Areas**

Web of Science Categories	records	% of 55671	Research Areas	records	% of 55671
ACOUSTICS	29	0.052	ACOUSTICS	29	0.052
AGRICULTURE MULTIDISCIPLINARY	80	0.144	AGRICULTURE	80	0.144
ASTRONOMY ASTROPHYSICS	67	0.120	ASTRONOMY ASTROPHYSICS	67	0.120
AUTOMATION CONTROL SYSTEMS	979	1.759	AUTOMATION CONTROL SYSTEMS	979	1.759
BIOCHEMICAL RESEARCH METHODS	1017	1.827	BIOCHEMISTRY MOLECULAR BIOLOGY	1077	1.935
BIOCHEMISTRY MOLECULAR BIOLOGY	60	0.108			
BIOLOGY	269	0.483	LIFE SCIENCES BIOMEDICINE OTHER TOPICS	269	0.483
BIOTECHNOLOGY APPLIED MICROBIOLOGY	842	1.512	BIOTECHNOLOGY APPLIED MICROBIOLOGY	842	1.512
BUSINESS FINANCE	241	0.433	BUSINESS ECONOMICS	1134	2.037
ECONOMICS	1051	1.888			
CHEMISTRY ANALYTICAL	225	0.404	CHEMISTRY	957	1.719
CHEMISTRY MULTIDISCIPLINARY	296	0.532			
CHEMISTRY PHYSICAL	436	0.783			
COMPUTER SCIENCE ARTIFICIAL INTELLIGENCE	490	0.880	COMPUTER SCIENCE	5465	9.817
COMPUTER SCIENCE INFORMATION SYSTEMS	62	0.111			
COMPUTER SCIENCE INTERDISCIPLINARY APPLICATIONS	2908	5.224			
COMPUTER SCIENCE SOFTWARE ENGINEERING	1231	2.211			
COMPUTER SCIENCE THEORY METHODS	1306	2.346			
DEMOGRAPHY	16	0.029	DEMOGRAPHY	16	0.029

Таблица 8 (продолжение)

Web of Science Categories	records	% of 55671	Research Areas	records	% of 55671
ENERGY FUELS	50	0.090	ENERGY FUELS	50	0.090
ENGINEERING BIOMEDICAL	76	0.137	ENGINEERING	3684	6.617
ENGINEERING CHEMICAL	50	0.090			
ENGINEERING CIVIL	84	0.151			
ENGINEERING ELECTRICAL ELECTRONIC	552	0.992			
ENGINEERING ENVIRONMENTAL	84	0.151			
ENGINEERING INDUSTRIAL	141	0.253			
ENGINEERING MECHANICAL	66	0.119			
ENGINEERING MULTIDISCIPLINARY	2912	5.231			
ENVIRONMENTAL SCIENCES	204	0.366	ENVIRONMENTAL SCIENCES ECOLOGY	204	0.366
GEOSCIENCES MULTIDISCIPLINARY	51	0.092	GEOLOGY	51	0.092
HEALTH CARE SCIENCES SERVICES	44	0.079	HEALTH CARE SCIENCES SERVICES	44	0.079
HISTORY PHILOSOPHY OF SCIENCE	76	0.137	HISTORY PHILOSOPHY OF SCIENCE	76	0.137
IMAGING SCIENCE PHOTOGRAPHIC TECHNOLOGY	63	0.113	IMAGING SCIENCE PHOTOGRAPHIC TECHNOLOGY	63	0.113
INSTRUMENTS INSTRUMENTATION	225	0.404	INSTRUMENTS INSTRUMENTATION	225	0.404
LOGIC	601	1.080			
MANAGEMENT	69	0.124			
MATERIALS SCIENCE MULTIDISCIPLINARY	134	0.241	MATERIALS SCIENCE	134	0.241
MATHEMATICAL COMPUTATIONAL BIOLOGY	1883	3.382	MATHEMATICAL COMPUTATIONAL BIOLOGY	1883	3.382
MATHEMATICS	24555	44.107	MATHEMATICS	55671	100.000
MATHEMATICS APPLIED	24900	44.727			

Таблица 8 (продолжение)

Web of Science Categories	records	% of 55671	Research Areas	records	% of 55671
MATHEMATICS INTERDISCIPLINARY APPLICATIONS	8994	16.156			
MECHANICS	2774	4.983	MECHANICS	2774	4.983
MEDICAL INFORMATICS	402	0.722	MEDICAL INFORMATICS	402	0.722
MEDICINE RESEARCH EXPERIMENTAL	358	0.643	RESEARCH EXPERIMENTAL MEDICINE	358	0.643
MULTIDISCIPLINARY SCIENCES	857	1.539	SCIENCE TECHNOLOGY OTHER TOPICS	1458	2.619
MUSIC	22	0.040	MUSIC	22	0.040
OPERATIONS RESEARCH MANAGEMENT SCIENCE	1391	2.499	OPERATIONS RESEARCH MANAGEMENT SCIENCE	1391	2.499
PHARMACOLOGY PHARMACY	158	0.284	PHARMACOLOGY PHARMACY	158	0.284
PHILOSOPHY	130	0.234	PHILOSOPHY	130	0.234
PHYSICS APPLIED	65	0.117	PHYSICS	3543	6.364
PHYSICS ATOMIC MOLECULAR CHEMICAL	436	0.783			
PHYSICS FLUIDS PLASMAS	831	1.493			
PHYSICS MATHEMATICAL	2665	4.787			
PHYSICS MULTIDISCIPLINARY	432	0.776			
PSYCHOLOGY EDUCATIONAL	55	0.099	PSYCHOLOGY	227	0.408
PSYCHOLOGY EXPERIMENTAL	62	0.111			
PSYCHOLOGY MATHEMATICAL	192	0.345			
PUBLIC ENVIRONMENTAL OCCUPATIONAL HEALTH	560	1.006	PUBLIC ENVIRONMENTAL OCCUPATIONAL HEALTH	560	1.006
ROBOTICS	12	0.022	ROBOTICS	12	0.022
SOCIAL SCIENCES INTERDISCIPLINARY	126	0.226	SOCIAL SCIENCES OTHER TOPICS	126	0.226

Таблица 8 (продолжение)

Web of Science Categories	records	% of 55671	Research Areas	records	% of 55671
SOCIAL SCIENCES MATHEMATICAL METHODS	1634	2.935	MATHEMATICAL METHODS IN SOCIAL SCIENCES	1634	2.935
SOCIOLOGY	11	0.020	SOCIOLOGY	11	0.020
STATISTICS PROBABILITY	9369	16.829			
THERMODYNAMICS	144	0.259	THERMODYNAMICS	144	0.259
WATER RESOURCES	84	0.151	WATER RESOURCES	84	0.151

к категории мультидисциплинарных. В случае этих журналов анализ осуществляется по отдельным статьям на основании списков процитированной литературы и профиля журналов, из которых статьи получили цитирования, поэтому такие журналы в Essential Science Indicators могут попадать в несколько из выделенных 22 широких областей, при этом их ранжирование определяется количеством статей и цитирований по соответствующей области. Для примера приведем данные из Essential Science Indicators по уже упоминавшемуся журналу PLoS ONE (табл. 9).

Таким образом, из-за большого разнообразия вариантов рубрикаторов, даже в пределах одного и того же ресурса, при определении любых наукометрических показателей (количество статей, количество цитирований) необходимо четко указывать, какие именно данные используются, на основании какой базы данных, за какой конкретно период и с использованием какого конкретно рубрикатора они получены.

3.8. Важность источника информации о публикациях для оценки научной деятельности

Огромное значение при проведении любых наукометрических исследований имеет четкое указание источника получения све-

дений о публикациях и цитированиях. Это связано в первую очередь с тем, что различные базы данных формируются по совершенно различным принципам. Google Scholar или Microsoft Academic Search формируются на основании данных из сети Интернет, т. е. любых данных о публикациях, которые можно найти в сети: журналов открытого доступа, списков процитированной в найденных статьях литературы, репозитории публикаций, личных страниц ученых в Интернете, материалов конференций, открытых баз публикаций и т. д. С одной стороны, это обеспечивает максимальную широту представленных материалов, а с другой стороны, ограничивает тем, что в эти системы не попадают сведения из закрытых источников и публикации, отсутствующие в электронном виде (или упоминания о таковых). Кроме того, Google не может полностью гарантировать достоверность информации, полученной, например, из размещенного на личной странице ученого перечня публикаций, или достоверность вручную внесенных ученым сведений о публикациях при создании личного профиля в Google Scholar.

Международные индексы цитирования ставят перед собой задачу проиндексировать наиболее значимые журналы в каждой области знаний (см., например, принципы

Таблица 9

Данные из Essential Science Indicators

Research Areas	Rank in Research Area	Journals	Web of Science Documents	Cites	Cites/Paper
MULTIDISCIPLINARY	4	PLOS ONE	5888	39861	6,769871
MICROBIOLOGY	16	PLOS ONE	4909	33427	6,80933
IMMUNOLOGY	21	PLOS ONE	706	46299	65,57932
MOLECULAR BIOLOGY & GENETICS	29	PLOS ONE	12899	95157	7,377083
NEUROSCIENCE & BEHAVIOR	30	PLOS ONE	9942	60119	6,046972
PLANT & ANIMAL SCIENCE	30	PLOS ONE	7123	3446	0,483785
CLINICAL MEDICINE	36	PLOS ONE	24375	118063	4,84361
BIOLOGY & BIOCHEMISTRY	43	PLOS ONE	9825	5328	0,54229
ENVIRONMENT/ECOLOGY	62	PLOS ONE	3621	16147	4,459265
SOCIAL SCIENCES, GENERAL	88	PLOS ONE	1817	8235	4,532196
PSYCHIATRY/ PSYCHOLOGY	110	PLOS ONE	2543	8591	3,378293
PHARMACOLOGY & TOXICOLOGY	152	PLOS ONE	951	4431	4,659306
AGRICULTURAL SCIENCES	192	PLOS ONE	477	1256	2,633124
GEOSCIENCES	200	PLOS ONE	541	2388	4,414048
COMPUTER SCIENCE	214	PLOS ONE	372	976	2,623656

отбора журналов для Web of Science CC⁸), а многие национальные индексы цитирования стараются максимально охватить весь сегмент национальных научных публикаций, не ставя перед собой задачу исходного отбора лучших журналов – это становится уже следующим этапом развития национальных индексов, как, например, формирование рейтинга журналов Science Index в Российском индексе научного цитирования.

Профильные базы данных, такие, например, как AGRIS, не индексируют некие наборы журналов целиком, а формируются путем от-

бора конкретных статей, в данном случае по сельскохозяйственной тематике. Попадание статьи в эту базу означает лишь то, что ее тематика совпадает с профилем базы данных, но не является критерием качества проведенного исследования или написанной статьи.

Даже размещенные на одной и той же платформе (например, на платформе Web of Science) различные базы данных формируются по разным принципам. Для Web of Science Core Collection осуществляется очень строгий отбор журналов, Medline на платформе Web of Science получает сведения о публикациях из крупнейшей библиографической базы статей по медицинским наукам, соз-

⁸<http://wokinfo.com/essays/journal-selection-process/>

данной Национальной медицинской библиотекой США (U.S. National Library of Medicine, NLM), охватывающей около 75 % мировых медицинских изданий. База данных BIOSIS Previews/RN — самая большая и наиболее исчерпывающая база данных по биологическим наукам в мире — охватывает оригинальные исследовательские отчеты, обзоры и избранные патенты США в области биологии и биомедицины, от аэрокосмической биологии до зоологии. Источниками BIOSIS являются периодические издания, журналы, труды конференций, обзоры, отчеты, патенты и краткие сообщения. Отбираются для включения в базу данных почти 6000 биологических журналов, 1500 международных конференций, а также книги и монографии. Zoological Record — крупнейшая база данных по названиям животных, которая берет свое начало от основанного в 1864 г. международного зоологического журнала. Сведения в эту базу попадают, кроме периодических журналов, из книг, научных отчетов, материалов конференций, и основной принцип отбора — достоверная информация, касающаяся вопросов биоразнообразия, систематики и иных зоологических данных.

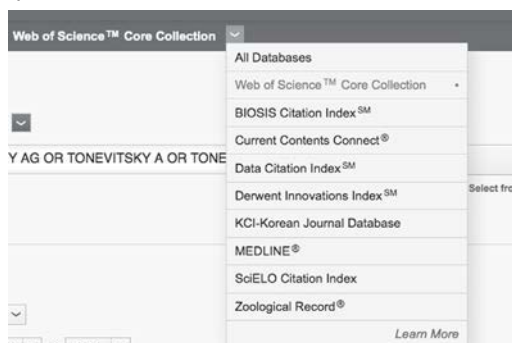
После переименования в 2014 г. платформы Web of Knowledge в Web of Science вопрос корректного указания источника данных встал особенно остро, поскольку в большинстве российских нормативных документов, в формах заявок на гранты и т. д. предлагается указывать количество публикаций, цитирований, индекс Хирша и т. д. в Web of Science, без уточнения, что имеется в виду — база данных Web of Science Core Collection или платформа Web of Science.

Рассмотрим на конкретном примере, какие возникают проблемы при игнорировании точного указания источника данных и почему это происходит.

В связи с тем что среди индексов цитирования, размещенных на платформе Web of Science, три относятся к области биологии и медицины (BIOSIS, Medline и Zoological Record), выберем из списка авторов в области биологии в РИНЦ ученого со значительным количеством публикаций и нераспространенной фамилией (для избежания лишних ошибок в поиске) и попробуем определить его наукометрические показатели в Web of Science при выборе различного набора баз данных.

34	☐	Тоневский Александр Григорьевич*	202	1297	18
		Московский научно-исследовательский онкологический институт им. П.А. Герцена (Москва)			

Сразу отметим, что приведенный пример не следует рассматривать как строгую инструкцию для поиска публикаций по фамилии автора, а лишь как иллюстрацию того, что одинаковый алгоритм может дать разные результаты при поиске, если не обращать внимания на выбор конкретной базы данных при поиске на платформе. Однозначное определение всех публикаций конкретного автора возможно лишь самим автором, поэтому ученым настоятельно рекомендуется зарегистрироваться в системе ResearcherID⁹ или ORCID¹⁰ и регулярно обновлять свой профиль в данных системах.



⁹<http://www.researcherid.com>.

¹⁰<http://www.orcid.org>.

Поиск по Web of Science Core Collection при выборе нескольких возможных написаний фамилии автора (TONEVITSKY AG OR TONEVITSKII AG OR TONEVITSKI AG) дает следующие результаты (рис. 24).

Поиск по тому же набору вариантов фамилии, но уже по всей платформе, а не только по Web of Science CC, даст совершенно другие результаты (рис. 25).

Данные различия связаны с тем, что при втором варианте поиска набор публикаций включает в себя также публикации из BIOSIS и Medline (рис. 26).

Для ученых-зоологов большую роль будут играть публикации в базе данных Zoological Record, для ученых, занимающихся преимущественно прикладными исследованиями и разработками, при поиске по всей платформе будут попадаться результаты, проиндексированные в Derwent Innovations Index.

Следует отметить, что при появлении, как планируется, в 2015 г. на платформе Web of Science нового индекса Russian Science Citation Index эта проблема станет актуальной не только для ученых медико-биологических или технических специальностей, но и для представителей всех других областей наук.

Таким образом, при поиске по всем базам данных платформы Web of Science к статьям, обнаруженным в Web of Science CC, добавляются данные из других баз, что дает изменение наукометрических показателей ученого.

Если мы попытаемся определить аналогичные параметры в других базах данных, например в Google Scholar и Microsoft Academic Search (в том случае, если у автора имеется персональный профиль в этих ресурсах или отсутствуют однофамильцы), то получим опять другие данные как по количеству статей и цитирований, так и для индекса Хирша.

Регистрация авторов в таких свободных системах значительно улучшает их видимость и, соответственно, дает более корректные результаты поиска, однако как Google Scholar, так и Microsoft Academic Search осуществляют поиск публикаций по всему пространству Интернета, а также позволяют авторам самостоятельно добавлять публикации в свои профили, поэтому наукометрические показатели в таких системах, как правило, оказываются значительно выше, чем в Web of Science CC или даже по всей платформе Web of Science.

Таким образом, невозможно говорить о количестве публикаций, цитировании, индексе Хирша и иных показателях вообще, необходимо указывать, по какой конкретно базе данных получены данные. Кроме этого, перед расчетом показателей по Web of Science Core Collection необходимо также проверить, достаточна ли глубина архива базы данных Web of Science Core Collection, зависящая от условий подписки организации, для охвата всех публикаций данного автора.

Размещение на платформе Web of Science Российского индекса цитирования, планируемое в недалеком будущем, может усугубить неопределенность, если не будет точного указания, какие конкретно сведения о публикациях необходимо приводить в заявках на конкурсы или отчетных документах.

3.9. Особенности публикаций и цитирования в разных научных областях

Прежде чем рассматривать возможные способы сравнения результативности научной деятельности на основании наукометрических показателей, необходимо рассмотреть, в чем же конкретно заключаются различия в публикациях представителей

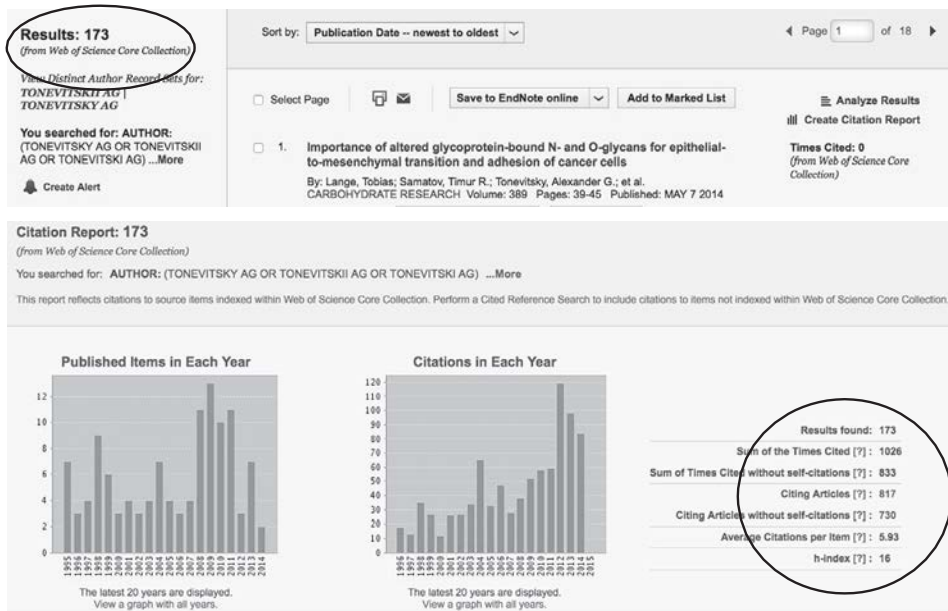


Рис. 24. Поиск по Web of Science Core Collection при выборе нескольких возможных написаний фамилии автора

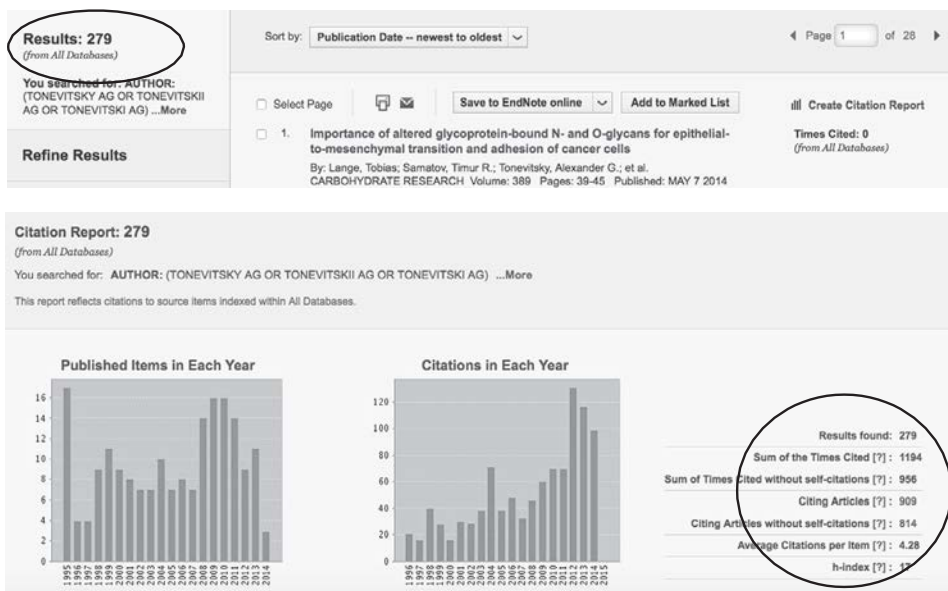


Рис. 25. Поиск по тому же набору вариантов фамилии, но уже по всей платформе

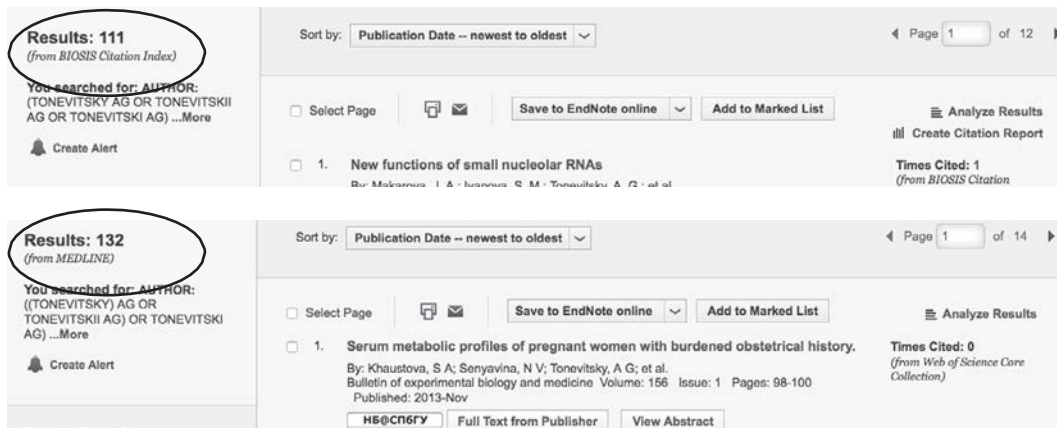
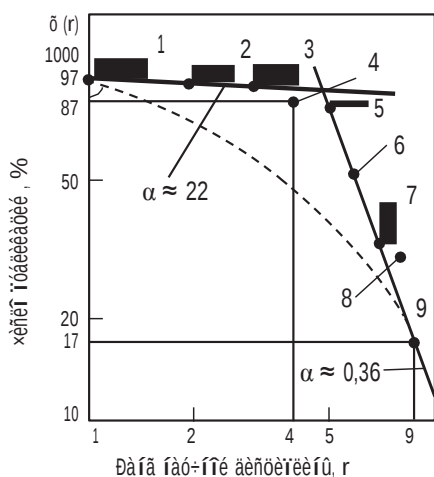


Рис. 26. Поиск по тому же набору вариантов фамилии по BIOSIS и Medline



Ранговое распределение научных дисциплин по числу публикаций с одним автором:

1 — английский язык, 2 — история, 3 — философия, 4 — математика, 5 — педагогика, 6 — психология, 7 — физика, 8 — биология, 9 — химия.

Массив не указан. Гипотеза, что данное распределение является распределением Ципфа с $\alpha \approx 0,36$ и $B \approx (8/((97/17)^{0,36} - 1)) - 1 \approx 8,2$, недостоверна, так как соответствующая кривая (пунктирная на графике) плохо описывает данные. Поэтому фиксируем $\alpha \approx 22$, соответствующее цифровому распределению, с $B \approx (3/((97/85)^{22} - 1)) - 1 \approx -0,83$

Рис. 27. Различия в научных дисциплинах по количеству публикаций с одним автором [по: Хайтун, 1983]

естественных наук и социально-гуманитарных направлений.

В чем же реально проявляются особенности публикаций и их цитирования в разных научных областях?

Еще в книге С. Д. Хайтуна «Наукометрия. Состояние и перспективы» (1983) приводятся данные о различиях в количестве соавторов в разных дисциплинах. По данным Берельсона [Berelson, 1960] приводится график распределения научных дисциплин по числу публикаций с одним автором (рис. 27).

Очень похожее распределение получается при анализе фактических данных о количестве соавторов по базе данных РИНЦ. На рис. 28 приведены средние значения доли статей с одним автором, рассчитанные по сведениям о 100 авторах — лидерах по количеству публикаций в РИНЦ по соответствующему научному направлению.

Аналогичный анализ статей, у которых количество авторов больше 10, дает практически прямо противоположный результат (рис. 29). В статьях по физике количество авторов может достигать до 3 тысяч

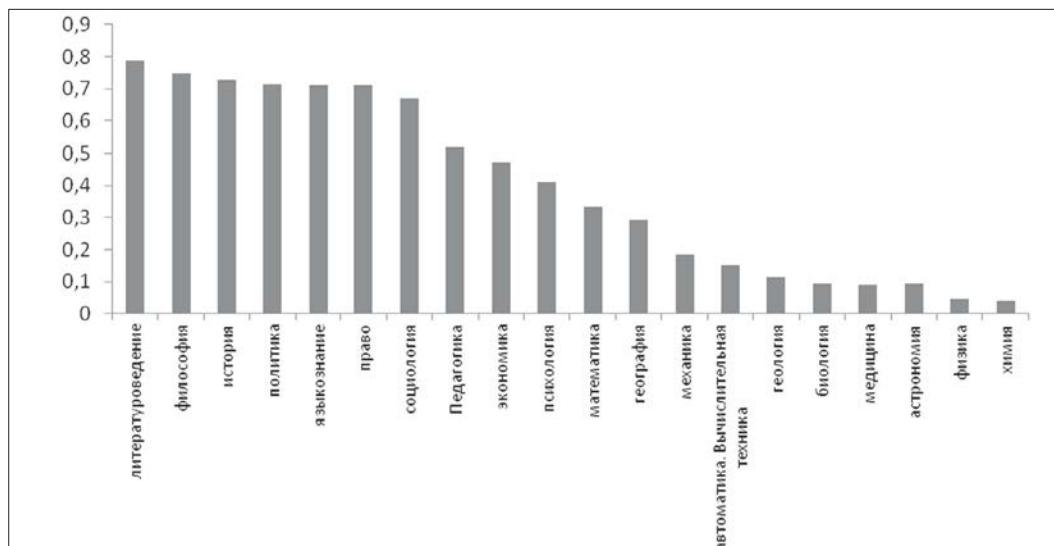


Рис. 28. Различия в научных дисциплинах по количеству публикаций с одним автором по данным РИНЦ

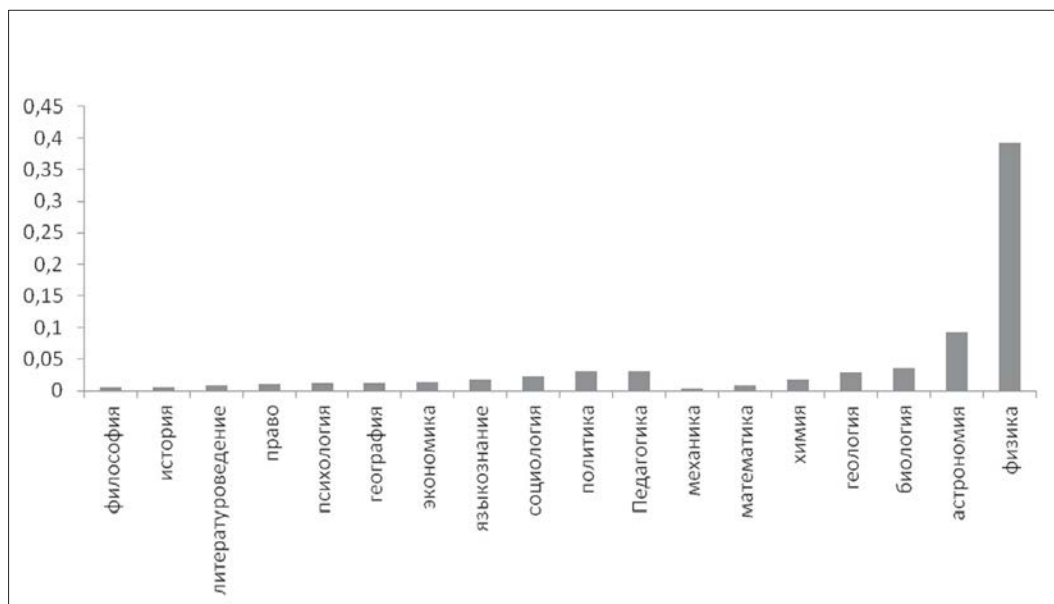


Рис. 29. Различия в научных дисциплинах по количеству публикаций с более чем 10 авторами по данным РИНЦ

(например, большие коллаборации ЦЕРНа), в то время как в статьях по математике и механике чаще всего встречаются статьи с 2–5 авторами, а максимальное количество статей с 6–10 авторами — у химиков, биологов и геологов.

Таким образом, первое отличие заключается уже в распространении соавторства в различных научных дисциплинах. Очевидно, что внутри каждой дисциплины также имеются различия — в физике наибольшее количество соавторов у экспериментаторов в области ядерной физики, физики элементарных частиц, а число соавторов в теоретических работах или, например, в области математической физики значительно меньше — примерно как у математиков. Однако разница по этому параметру между естественно-научными дисциплинами и социально-гуманитарными направлениями в целом видна невооруженным глазом уже из распределения статей с одним автором.

Значительно различается и распределение по типам публикаций. В базе данных Web of Science Core Collection при анализе, например, публикаций 2010–2012 гг., проиндексированных в SSCI, A&HCI, CPCI-SSH, BKCI-SSH, получаем следующее распределение по типам документов (рис. 30).

Аналогичный поиск по SCI-EXPANDED, CPCI-S, BKCI-S дает следующий результат (рис. 31).

Заметно, что книжные публикации играют в социальных науках и гуманитарных направлениях значительно большую роль, чем в естественно-научных дисциплинах, где подавляющее большинство публикаций составляют статьи в журналах и материалы конференций.

Следующее отличие заключается в том, что именно цитируют авторы в различных

View Records				
X Exclude Records				
	Field: Document Types	Record Count	% of 426160	Bar Chart
☐	ARTICLE	230916	54.185 %	
☐	BOOK REVIEW	74721	17.534 %	
☐	BOOK CHAPTER	52258	12.263 %	
☐	EDITORIAL MATERIAL	32361	7.594 %	
☐	PROCEEDINGS PAPER	30665	7.196 %	
☐	MEETING ABSTRACT	23700	5.561 %	
☐	REVIEW	6830	1.603 %	
☐	LETTER	5217	1.224 %	
☐	POETRY	4537	1.065 %	
☐	BOOK	4389	1.030 %	
☐	ART EXHIBIT REVIEW	2546	0.597 %	
☐	NEWS ITEM	2488	0.584 %	
☐	BIOGRAPHICAL ITEM	2484	0.583 %	
☐	RECORD REVIEW	1804	0.423 %	
☐	FILM REVIEW	1716	0.403 %	
☐	CORRECTION	1371	0.322 %	
☐	MUSIC PERFORMANCE REVIEW	1197	0.281 %	
☐	FICTION CREATIVE PROSE	592	0.139 %	
☐	DANCE PERFORMANCE REVIEW	462	0.108 %	
☐	THEATER REVIEW	347	0.081 %	
☐	TV REVIEW RADIO REVIEW	323	0.076 %	
☐	BIBLIOGRAPHY	190	0.045 %	
☐	REPRINT	118	0.028 %	
☐	MUSIC SCORE REVIEW	66	0.015 %	
☐	EXCERPT	51	0.012 %	
☐	SCRIPT	31	0.007 %	
☐	MUSIC SCORE	19	0.004 %	
☐	SOFTWARE REVIEW	16	0.004 %	
☐	DATABASE REVIEW	9	0.002 %	
☐	HARDWARE REVIEW	5	0.001 %	
View Records				
X Exclude Records				
	Field: Document Types	Record Count	% of 426160	Bar Chart

Рис. 30. Распределение публикаций по типам 2010–2012 гг. Web of Science CC SSCI, A&HCI, CPCI-SSH, BKCI-SSH

View Records				
X Exclude Records				
	Field: Document Types	Record Count	% of 1765314	Bar Chart
☐	ARTICLE	1110125	62.885 %	
☐	PROCEEDINGS PAPER	251441	14.243 %	
☐	MEETING ABSTRACT	228567	12.948 %	
☐	EDITORIAL MATERIAL	78410	4.442 %	
☐	REVIEW	60725	3.440 %	
☐	BOOK CHAPTER	57068	3.233 %	
☐	LETTER	40602	2.300 %	
☐	NEWS ITEM	18994	1.076 %	
☐	CORRECTION	10909	0.618 %	
☐	BOOK	3989	0.226 %	
☐	BIOGRAPHICAL ITEM	3940	0.223 %	
☐	BOOK REVIEW	3360	0.190 %	
☐	REPRINT	372	0.021 %	
☐	SOFTWARE REVIEW	64	0.004 %	
☐	BIBLIOGRAPHY	43	0.002 %	
☐	MUSIC SCORE	10	0.001 %	
☐	DATABASE REVIEW	8	0.000 %	
☐	HARDWARE REVIEW	6	0.000 %	
View Records				
X Exclude Records				
	Field: Document Types	Record Count	% of 1765314	Bar Chart

Рис. 31. Распределение публикаций по типам 2010–2012 гг. Web of Science CC SCI-EXPANDED, CPCI-S, BKCI-S

научных областях. В гуманитарных и общественных науках в списках литературы встречается гораздо больше монографий, а «естественники» ссылаются в основном

на статьи в журналах. Во многих научных областях, таких как социология, образование, политология, антропология, гуманитарные направления, значительно большую роль играют национальные публикации, соответственно и ссылок на них оказывается значительно больше, чем ссылок на индексируемые в Web of Science CC издания. В публикациях по гуманитарным наукам и искусству также большую роль играют ссылки на первоисточники (например, архивные материалы и произведения искусства), которые не индексируются в реферативных базах данных научных публикаций. В связи с этим отсутствие публикаций ученого в Web of Science не является свидетельством отсутствия к ним международного интереса.

Очень подробный анализ различий в принятых устоявшихся нормах цитирования в журнальных публикациях в социально-гуманитарных науках представлен в книге Хэнка Муда [Moed, 2005]. Отмечаются не только отличие в стиле цитирования в естественных науках и значительные различия не только между разными направлениями социально-гуманитарных наук, но даже между разными направлениями в одной и той же научной области. Так, в социологических журналах, публикующих преимущественно статьи, в которых используются количественные методы социологии, гораздо больше цитируются журнальные статьи, чем монографии, а в журналах с «качественными» исследованиями — наоборот. Таким образом, «количественные» исследования в социологии по манере цитирования ближе к Science, а «качественные» — к Humanities. Кроме того, в «количественных» статьях цитируются в основном «количественные» журналы, тогда как в статьях «качественного» характера присутствуют ссылки на оба типа журналов. Таким образом, простая

интерпретация результатов анализа цитирования в Web of Science CC может привести к выводу о том, что исследования в «количественной» социологии цитируются значительно больше, чем в «качественной».

Таким образом, показатели цитирования в социально-гуманитарных науках в значительно большей степени, чем в естественных, находятся в зависимости от уровня представления национальных журналов по соответствующим областям в международных индексах цитирования, от языка публикации и от конкретного направления науки. Следовательно, подходы к оценке публикационной активности в этих областях должны отличаться от применяемых для естественных наук. Делать выводы о состоянии исследований в социально-гуманитарной сфере только на основании цитирования в Web of Science CC надо весьма осторожно.

3.10. Способы оценки публикаций, отличные от анализа цитирования

С развитием интернет-технологий начали развиваться способы оценки публикаций, отличные от привычных библиометрических методов. Это способы оценки публикаций не по числу их цитирований в базах данных, а по их реальному использованию или по уровню интереса к публикациям, который также может быть измерен и выражен в совершенно конкретных значениях.

Особенно это актуально в отношении тех областей науки, в которых цитирования, как правило, могут не быть реальным отражением их ценности. В 2010 г. был принят манифест альтметрики¹¹, в котором говорится о том, что целью этого проекта является предоставление ученым инструментов быстрого анализа популярности научной

¹¹<http://altmetrics.org/manifesto/>

литературы, помогающей сделать выбор в пользу тех или иных из огромного количества опубликованных статей. При поиске публикаций через базы цитирований читатель ориентируется не в последнюю очередь на цитирование статей или импакт-факторы журналов, однако от момента публикации до появления цитирований проходит достаточно продолжительное время, поэтому показатели, рассматриваемые в альтметрике, могут оказаться весьма полезным инструментом. Кроме того, что очень существенно, такие показатели никак не зависят от научной области, в отличие от стандартных показателей цитирования.

В альтметрике учитываются такие показатели, как загрузка статей в менеджерах цитирования, их упоминания в социальных сетях, научных блогах, обсуждения статей читателями на сайтах журналов и т. д. Большая скорость представления данных альтметрики предоставляет исследователям возможность оперативно получать сведения о рекомендованных коллегами публикациях, самим делиться прочитанными статьями, настраивать оповещения о новых поступлениях и т. д. Особенно актуально это становится с расширением сети научных блогов, системы препринтов, что сокращает коммуникационный цикл с нескольких лет, как в случае ознакомления с бумажными научными изданиями, до нескольких недель и даже дней. Развитие альтметрики создает также предпосылки создания мощной системы общественной экспертизы публикаций, что уже реализуется как дополнительная альтернативная экспертиза в таких журналах открытого доступа, как PLoS One, BMC Research Notes или BMJ Open. В отличие от импакт-фактора журнала, альтметрики отражают влияние самой статьи, включая показатели престижности журнала.

Такой подход позволяет отслеживать влияние статей за пределами замкнутого академического сообщества, причем независимо от того, процитирована данная статья или нет в рецензируемых журналах. В этом плане развитые системы альтметрик могут оказаться более надежными показателями использования журналов, чем стандартно вычисляемые импакт-факторы, которые также подвержены манипулированию.

На основе регулярно обновляемой статистики использования журналов по компьютерным наукам, издаваемых ACM (например, <http://tods.acm.org/Statistics.html>), видно, что количество загрузок, а следовательно, прочтений статей из журналов значительно превышает количество цитирований статей из того же журнала, что может косвенно служить свидетельством в пользу использования. Так, для журнала ACM Transactions on Database Systems (на 25.05.2014 г.) приводятся следующие данные:

Publication years	1976–2014
Publication count	851
Citation Count	27250
Available for download	851
Downloads (6 Weeks)	7535
Downloads (12 Months)	45355
Downloads (cumulative)	703227
Average downloads per article	826,35
Average citations per article	32,02

Показатели альтметрики для научных публикаций можно сравнить в какой-то мере с показателями продаж для художественной и популярной литературы, помогающими читателям быстрее ориентироваться в море издаваемой литературы. А такие показатели, как закупка монографий и учебников в крупные научные и вузовские би-

блиотеки, предлагаются некоторыми исследователями в качестве критериев ценности учебных и научных изданий как альтернатива показателям цитируемости и импакт-факторам журналов.

Появившийся в 2014 г. проект Altmetrics for Institutions¹² помогает университетам и научным организациям быстро отслеживать влияние опубликованных сотрудниками работ по появляющимся в Интернете откликам, количеству скачиваний и отзывов и другим параметрам. Альтметрика сопоставляет упоминания о научных статьях во всех традиционных и социальных медиа, блогах, менеджерах цитирования, а также отражает цитирование или упоминание статей в документах государственной политики, обеспечивая сведения о применении исследования в реальной жизни. Соответственно на уровне организаций становится возможным определять реальное воздействие исследований конкретного института или университета на экономические или общественно-политические изменения в государстве и мире.

В то же время исследований, посвященных изучению альтметрик, пока еще немного, и говорить о том, насколько достоверно показатели альтметрики сегодня отражают реальное влияние научных статей, еще рано. В статье, опубликованной в журнале PLoS ONE в мае 2013 г. [Thelwall et al., 2013], проведено исследование, демонстрирующее, как работает альтметрика, на примере ряда социальных сетей. На примере статей из PubMed было проведено сравнение их реального цитирования в Web of Science CC и 11 альтметрик из социальных сетей. Обнаружены статистически значимые зависимости значительного цитирования и высоких альтметрик для таких социальных

сетей, как Twitter, Facebook, RH (research highlights, Nature Publishing Group journals), научных блогов (Nature.com Blogs, Research Blogging, ScienceSeeker), средств массовой информации, статистически достоверных свидетельств в пользу наличия положительной корреляции для LinkedIn, Reddits и некоторых других альтметрик найдено не было. На примере статей из других ресурсов (arXiv и некоторые отдельные журналы) ранее была показана также положительная связь между цитированиями статей и загрузками в Mendeley и упоминаниями в Twitter. Ограничение таких исследований пока состоит в отсутствии большой статистики, поскольку для статей прошлых лет с достоверными цитированиями отсутствуют показатели альтметрики, а для свежих статей с доступной альтметрикой еще недостаточно цитирований, так что необходимо проводить еще довольно много исследований и совершенствовать всю систему альтметрик.

3.11. Возможные подходы к сравнительному анализу цитирования публикаций в разных областях знаний

Из всего вышесказанного следует, что по таким стандартным показателям цитирования, как среднее цитирование публикации или индекс Хирша и его вариации, проводить сравнение результативности научной деятельности научных коллективов или отдельных авторов, работающих в разных научных направлениях, представляется весьма затруднительным. Это представляет собой довольно серьезную проблему, поскольку в настоящее время для принятия административных, кадровых или финансовых решений все чаще используются публикационные показатели, которые

¹²<http://www.altmetric.com/institutions.php>

называют «объективными», т. е. количественные показатели, не зависящие от мнения экспертов.

Какие возможны подходы для решения этой проблемы?

В настоящее время присутствуют все варианты мнений о возможности использования библиометрических показателей для оценки научной деятельности — от категорического отказа от их использования, по крайней мере для ряда дисциплин (гуманитарные науки, математика), до использования исключительно показателей цитируемости и отказа от экспертной оценки. Истина, конечно, где-то посередине, но для этого необходим выбор адекватных показателей и критериев и подходов к оценке.

Более или менее универсальным по отношению к разным областям знаний можно считать такой показатель, как доля публикаций в топовых журналах (или распределение публикаций по журналам, относящимся к разным квартилям). Аналогичным можно считать распределение публикаций по процентиллям (на основании сводных данных, представленных в Essential Science Indicators, например — Percentiles for papers published by field). Однако такой подход не годится для гуманитарных специальностей, поскольку, во-первых, ни импакт-факторы журналов, ни проценты для них не рассчитываются, и, во-вторых, (по таблице процентилей из ESI) данные приводятся по укрупненным областям, а как показано выше, даже внутри одной области существуют большие различия в цитируемости между узкими направлениями науки (теоретическая/экспериментальная физика, например, или систематика и генетика в биологических дисциплинах). Тем не менее для сравнительной оценки крупных организаций, где ведутся исследования по многим

направлениям науки, такой подход вполне приемлем, как и для сравнения узкопрофильных организаций, работающих в сходных научных направлениях, но сильно различающихся по своим размерам (количеству ученых).

Абсолютно очевидно, что более или менее корректного сравнения библиометрических показателей для наборов публикаций, относящихся к разным областям знаний, можно добиться только путем введения показателей, в которых можно было бы компенсировать тем или иным путем различия в уровне цитирования в разных научных областях и для разных типов документов в целом, в динамике цитирования, а в идеале и различия в распределении публикаций и цитирований в международных и национальных изданиях.

На сегодняшний день наилучшим путем для такого сравнения является расчет нормализованных показателей цитирования, который заключается в следующем. Как уже отмечалось в главе 2 «Библиометрические индикаторы в ресурсах Thomson Reuters», нормализованные показатели цитируемости широко используются в аналитическом инструменте InCites, размещенном на платформе Web of Science.

Для каждой узкой научной области и конкретного типа документов вычисляется среднее цитирование одной статьи, опубликованной в конкретном году, и далее цитирование конкретной статьи, опубликованной в том же году и относящейся к выбранной категории, сравнивается со средним показателем.

Например, выбираются все обзоры по молекулярной биологии, опубликованные в 2010 г., и определяется, сколько цитирований в среднем каждый из этих обзоров получил к 2013 г. Если разделить количество

цитирований к 2013 г. одного конкретного обзора 2010 г. (или средний показатель для всех обзоров, опубликованных ученым или организацией в 2010 г.) на полученный среднемировой показатель, то значение, равное 1, будет означать, что рассматриваемый обзор(ы) цитируется точно также, как в среднем по миру, больше 1 — лучше, а меньше 1 — хуже, чем в среднем по миру. Степень отклонения от единицы в ту или иную сторону покажет уровень «превосходства» статей ученого или организации (или наоборот).

Так, в базе данных Web of Science Core Collection в 2010 г. представлено 5653 обзора (Review), относящихся к категории Biochemistry & Molecular Biology. По состоянию на август 2014 г. эти обзоры процитированы в целом 159 331 раз, что составляет 28,19 цитирований на 1 статью. Это среднемировой показатель. При этом статьи ученых из США (2261 обзор) процитированы 75 016 раз (среднее значение 33,18), Германии (498 статей) — 16 708 раз (среднее значение 33,55), Китая (256 статей) — 5373 раза (среднее значение 20,99), России (116 статей) — 1186 раз (среднее значение 10,22). Аналогичные подсчеты по данным для обзоров 2012 г. приведены в табл. 10.

Таким образом, обзоры американских и немецких ученых в области биохимии и молекулярной биологии цитируются лучше, чем в среднем по миру, а китайских

и российских — хуже, причем понятия «лучше» и «хуже» имеют конкретное числовое значение.

Если сделать такие расчеты для каждого типа публикаций (оригинальные статьи, обзоры, материалы конференций и т. д.) для всех областей знаний и по всем годам, то усреднение полученных значений нормализованной цитируемости достаточно адекватно отражает «качество» всего набора публикаций ученого, научной группы, организации, страны. Аналогичным образом можно провести сравнение, например, российских организаций по отношению к среднероссийскому уровню или сравнение разных подразделений внутри одной организации по отношению к средним показателям по организации в целом. Естественно, чем обширнее база данных, из которой получают сведения о цитировании, тем более корректными будут такие расчеты.

Однако и при таком способе анализа результаты могут оказаться весьма спорными. В случае, если статей у ученого или организации мало, существует вероятность того, что единичные статьи окажутся высокоцитируемыми, что может исказить картину. Так, например, всего 3 статьи МИФИ, опубликованные по результатам исследований в ЦЕРНе в составе больших коллабораций, из 1,5 тысяч, опубликованных за 5 лет, привели к тому, что нормализованный показатель

Таблица 10

Значения нормализованного цитирования

Страны	2010 г.	2012 г.
США	33,18/28,19 = 1,18	13,58/11,03 = 1,23
Германия	33,55/28,19 = 1,19	13,15/11,03 = 1,19
Китай	20,99/28,19 = 0,74	8,52/11,03 = 0,77
Россия	10,22/28,19 = 0,36	4,32/11,03 = 0,39

цитируемости МИФИ оказался настолько высок, что этот университет опередил по этому показателю все вузы мира в рейтинге THE World University Ranking 2010 г. Эти 3 статьи составили более четверти от всех цитирований статей организации за данный период (по состоянию на август 2014 г. — 6669 цитирований из 21 841). Всего же за этот период в коллаборациях ЦЕРНа сотрудниками МИФИ опубликовано около сотни статей с более чем 10 тыс. цитирований, т. е. около 6 % всех публикаций принесли половину цитирований. Авторами самой высокоцитируемой статьи (около 4 тыс. цитирований) являются 170 человек из 131 организации, из которых только 1 — сотрудник МИФИ.

Таким образом, для корректной оценки организации и автора по публикациям желательно учитывать также и вклад автора (или организации) в той или иной статье. Поскольку в разных областях знаний бывают абсолютно разные традиции в плане указания порядка авторов (главный автор может стоять первым, последним, указан как *gerprint author*, авторы могут быть расположены просто в алфавитном порядке), то в самом простом случае вклад определяется как доля от общего количества авторов, в случае организации — доля авторов с определенной аффилиацией от общего количества авторов. Такой принцип подсчета публикаций и цитирований называется фракционированием. Так, если в статье имеются два автора из двух разных организаций, то каждой организации в зачет идет 0,5 статьи и соответственно 0,5 от всех цитирований. Если организаций больше, то статья вместе со всеми полученными цитированиями распределяется по организациям пропорционально их вкладу. Таким образом удается избежать артефактов, связанных с ростом числа авторов в статьях, что особенно актуально при

усилении международного сотрудничества и формировании больших распределенных исследовательских коллективов.

В 2010 г. консорциум британских университетов разработал систему метрик, независимых от источника данных, позволяющих проводить объективную оценку научной деятельности университетов на основании объективной информации, которые были названы *Snowball Metrics*¹³. Целью данного проекта была разработка прозрачной методологии учета показателей научной деятельности, которые могут применяться в разных странах, обеспечивают прозрачное и корректное сравнение сходных по структуре и направлениям деятельности организаций, которые могут быть использованы для различных целей самими организациями, грантодателями, вышестоящими организациями и т. д.

Большая часть предлагаемых метрик связана с оценкой публикаций университетов. В качестве основных метрик, характеризующих результативность исследовательской деятельности, предлагаются следующие:

- *Scholarly output* — учитываются журнальные публикации, серийные издания, книги, включая монографии и учебники, в том числе если сотрудники университета являются не только авторами, но и редакторами, произведения искусства, компьютерные программы и т. д. В данном показателе не учитываются патенты и диссертации. Эти показатели учитываются за определенный промежуток времени и для компенсации различных размеров вуза нормируются на количество персонала (в полных ставках).
- *Citation Count* — учитывается количество цитирований на единицу персонала (в полных ставках) за определенный период.

¹³<http://www.snowballmetrics.com/>

- Citations per Output — среднее цитирование одной публикации за определенный период.

- h-index — используется только для сравнения исследователей в одной дисциплинарной области.

- Field-Weighted Citation Impact — нормализованный по научным областям показатель цитирования. Учитывается за определенный временной промежуток (например, за три полных календарных года) с учетом типа публикации.

- Outputs in Top Percentiles — количество публикаций или доля публикаций, попадающих в определенный процент самых цитируемых по конкретным научным областям (1, 5, 10 и 25 %) за определенный промежуток времени. Если учитывается именно количество таких публикаций, то это значение нормируется на количество персонала (в полных ставках).

- Publications in Top Journal Percentiles — аналогично предыдущему показателю, но в отношении публикаций в журналах, попадающих соответственно в 1–25 % по журнальным показателям (импакт-фактор, SNIP, SJR).

- Показатели сотрудничества (Collaboration) — рассчитываются по количеству или доле публикаций, подготовленных в сотрудничестве с другими организациями, определяемыми по аффилиации авторов, при этом отдельно учитывается международное и национальное сотрудничество. Если учитывается количество, то, как и в предыдущих показателях, происходит нормирование на количество сотрудников организации.

- Collaboration Impact (влияние сотрудничества) — рассчитывается среднее цитирование публикаций, подготовленных в международном или национальном сотрудничестве, за определенный промежуток времени.

- Academic-Corporate Collaboration (сотрудничество университетов с бизнесом) и Academic-Corporate Collaboration Impact — аналогично предыдущим показателям, но в отношении совместных с промышленными предприятиями публикаций, что характеризует потенциал реального применения результатов исследовательской деятельности.

- Altmetrics — оценивается:

- Scholarly Activity — наличие публикаций организации в различных менеджерах цитирования, таких как Mendeley, CiteULike, Google Scholar Library, QUOSA, Papers, ScienceScape, MyScienceWork, colwiz, zotero, Academia.edu, ResearchGate и т. д.;

- Scholarly Commentary — ссылки на публикации в научных блогах, на сайтах, рецензии и комментарии к публикациям в PubMed, Faculty of 1000, Wikipedia и др.;
- Social Activity — упоминание публикаций организации в различных социальных сетях;

- Mass Media — упоминание публикаций организации в средствах массовой информации.

Эти показатели рассчитываются за определенный период времени как на количество сотрудников организации, так и на количество публикаций.

К показателям, используемым в Snowball Metrics, относятся также различные финансовые показатели, характеризующие научную деятельность (полученные гранты, например), участие в конференциях, количество объектов интеллектуальной собственности и доходы от внедрения патентов, количество созданных компаний, внедряющих научные достижения университета, их финансовое состояние.

Таким образом, для адекватной оценки научной деятельности предлагается

использовать весь спектр показателей, учитывающих как различия по областям знаний, так и размеры организаций, дополняя их показателями, альтернативными библиометрическим.

3.12. Использование библиометрических данных при построении рейтингов вузов и научных организаций

В международных рейтингах университетов, таких как Academic Ranking of World Universities (ARWU, или Шанхайский рейтинг), Times Higher Education World University Ranking (THE WUR) и QS World University Ranking, показатели, связанные с публикационной активностью, составляют от 20 до 60 % от окончательного балла, что говорит о чрезвычайной важности научных публикаций для оценки всех сторон деятельности университета — образовательной и научной, а также международной составляющей.

В Шанхайском рейтинге используются как «валовые» показатели продуктивности (количество тех или иных публикаций за определенные периоды времени), так и своего рода «качественные» показатели, характеризующие качество преподавательского состава и научную производительность университета. Эти показатели составляют в общей сложности 60 % от общего балла:

- количество высокоцитируемых ученых из числа работников университета (20 %);
- количество статей, опубликованных в журналах Nature и Science за последние пять лет (20 %);
- общее число статей, вошедших в индексы научной цитируемости Science Citation Index Expanded и Social Science Citation Index в предыдущем году (20 %).

Эти показатели рассчитываются по базе данных Web of Science Core Collection.

Отнесение ученого к категории высокоцитируемых означает, что по количеству цитирований своих статей он попадает в группу 250 самых цитируемых ученых по 22 предметным областям по классификатору, применяемому в Essential Science Indicators. На обновленной в 2014 г. версии сайта¹⁴ представлены данные с указанием аффилиации авторов (включая дополнительную).

Для расчета показателя, связанного со статьями в журналах Science & Nature, используется довольно сложная система расчета весов для соавторов с введением различных весов для первого автора, второго и для остальных авторов, т. е. появляются элементы фракционированного подсчета. В данном случае такой расчет можно рассматривать как своеобразный вариант «дробного счета», fractional counting, описанного в предыдущей главе. При этом в инструкции для авторов отсутствует четкое указание на предмет порядка указания авторов (см., например, <http://www.sciencemag.org/site/feature/contribinfo/prep/index.xhtml>), т. е. данный способ расчета «авторской доли» является в какой-то мере искусственным, поскольку может не учитывать реальный вклад авторов в написание статьи, который в разных областях науки и в разных журналах устанавливается по-разному в зависимости от сложившейся практики и традиций. В ряде областей «главный» автор указывается первым, в других — последним, при большом количестве авторов часто используется просто алфавитный порядок.

При подсчете общего количества статей учитываются только типы публикаций articles и proceeding papers с удвоением веса статей из Social Science Citation Index.

¹⁴<http://highlycited.com/>

Это в какой-то мере может компенсировать отсутствие в методике учета публикаций, индексируемых в третьем журнальном индексе Web of Science CC — Arts & Humanities Citation Index.

Таким образом, если количество публикаций в данном случае можно рассматривать как чисто количественный показатель, то сведения по высокоцитируемым авторам и публикации в Science & Nature, скорее, являются показателями качества публикаций университета.

Кроме основного рейтинга составляются также предметные рейтинги, учитывающие особенности отдельных научных или предметных областей. Так, для рейтингов ARWU-FIELD, которые рассчитываются для Natural Sciences and Mathematics, Engineering/Technology and Computer Sciences, Life and Agriculture Sciences, совсем не учитываются показатели по статьям в Science & Nature, Clinical Medicine and Pharmacy и Social Sciences, кроме количества публикаций по конкретной области знаний, учитывается количество публикаций в 20 % топовых журналов соответствующей предметной области. Для рейтингов ARWU-SUBJECT (Mathematics, Physics, Chemistry, Computer Sciences and Economics/Business), как и в рейтинге ARWU-FIELD, учитывается количество публикаций в 20 % топовых журналов соответствующей предметной области. Таким образом, предметные рейтинги ARWU в значительной степени учитывают как количественные, так и качественные показатели.

С 2004 по 2009 г. приложение к газете Times — еженедельник Times Higher Education — в сотрудничестве с компанией Quacquarelli Symonds (QS) представлял свой собственный рейтинг ведущих университетов мира. Данный рейтинг в каче-

стве источника наукометрических данных использовал базу данных Scopus, однако с 2009 по 2014 г. Times Higher Education составлял рейтинг в сотрудничестве с Thomson Reuters уже на основе данных Web of Science, а QS продолжает составлять собственные рейтинги, используя по-прежнему данные из Scopus.

Методика составления рейтинга QS предусматривает оценку научной деятельности университета исходя из количества цитирований статей вуза в базе данных Scopus, что рассматривается как показатель качества преподавательского состава. Учитывается количество цитирований, полученных статьями университета за 5 лет по данным базы Scopus в расчете на 1 преподавателя. При этом с 2011 г. количество цитирований учитывается без самоцитирований организации, т. е. показывает реальный внешний интерес к работам, выполненным в вузе. Исключение самоцитирований стало самым существенным изменением в методике составления рейтинга после 2009 г. Этот показатель составляет 20 % в общем балле.

Одновременно с общим рейтингом университетов (в сентябре-октябре каждого года) публикуются рейтинги по научным областям — QS World University Rankings by Faculty. Эти рейтинги составляются по 5 крупным областям по следующей методике, претерпевшей значительные изменения с 2013 г. (табл. 11).

Существенным отличием от основного рейтинга является то, что цитирования рассчитываются не на количество преподавателей, а на статью, а также учитывается индекс Хирша статей организации по соответствующей области за определенный период.

С 2011 г. QS представляет также рейтинг университетов по более узким предметным областям — QS World University Rankings by

Subject. Рейтинги составляются по 30 отдельным предметным областям и публикуются топ-200 университетов по каждой области. Методика составления данных рейтингов практически аналогична методике QS World University Rankings by Faculty с более ярко выраженными различиями в весах показателей для каждой из 30 оцениваемых областей.

QS составляет также и региональные рейтинги — по странам Азии, Латинской Америки, БРИКС, а также довольно интересный рейтинг молодых университетов — Top 50 under 50. Для России в данном аспекте интересен рейтинг университетов БРИКС, созданный QS в 2013 г. Для этого рейтинга базовой методикой для составления являлась основная методика составления общего рейтинга, но с некоторыми видоизменениями — повышен вес мнения работодателей, учитывается не только количество цитирований (в этом рейтинге цитирования учитываются в расчете не на преподавателя, а на статью), но и количество публикаций в Scopus, несколько снижен вес показателей интернационализации и добавлен такой показатель качества преподавательского состава, как доля преподавателей с учеными степенями. Показатели, связанные с публикациями, в данном рейтинге составляют 15 % от общей оценки — 5% цитируемость

в расчете на 1 статью и 10 % — количество статей в расчете на 1 преподавателя.

Рейтинг Times Higher Education World University Ranking, составляемый с 2009 по 2014 г. в сотрудничестве с Thomson Reuters, претерпел по сравнению с тем, который составлялся вместе с QS, гораздо более сильные изменения, особенно в отношении показателей, связанных с публикациями.

Количественные и качественные данные о публикациях входят в интегральные показатели, характеризующие качество исследований (в целом 30 %), влияние исследований (30 %) и интернационализацию университета (7,5 %). В блоке показателей, связанных с качеством исследований, оценивается количество статей, опубликованных университетом в журналах, индексируемых в Web of Science CC, причем данный показатель рассчитывается с учетом как численности преподавателей, так и с нормализацией по областям знаний (6 %).

Влиятельность исследований рассчитывается по цитируемости статей университета. Цитируемость рассчитывается как полностью нормализованный показатель, учитывающий особенности цитирования в различных областях знаний, разницу в цитировании статей, опубликованных в разные годы, разных типов публикаций (научные статьи и обзоры, например). Кроме того, вводятся

Таблица 11

Рейтинги по научным областям

Faculty Area	Academic	Employer	Citations	H
Arts & Humanities	60 %	20 %	10 %	10 %
Engineering & Technology	40 %	30 %	15 %	15 %
Life Sciences & Medicine	40 %	10 %	25 %	25 %
Natural Sciences	40 %	20 %	20 %	20 %
Social Sciences & Management	50 %	30 %	10 %	10 %

повышающие или понижающие поправки на средний уровень цитируемости страны в целом. Такой подход позволяет достаточно корректно сравнивать цитируемость статей в разных областях наук с учетом национальных особенностей независимо от их количества и размера университета.

В показателях интернационализации университета треть, т. е. 2,5 %, приходится на долю статей, написанных в международном соавторстве.

Все показатели, связанные с публикациями, рассчитываются за 5-летний период и составляют в общей сложности 38,5 %.

Наряду с глобальным рейтингом составляются также и рейтинги по научным областям, содержащие все те же параметры, что и для общего рейтинга, но с другими весовыми коэффициентами. Так, для гуманитарных наук (Arts & Humanities) снижен вес цитируемости (15 %) и значительно выше оценивается качество образования (37,5 %). Сводные данные по весам групп показателей приведены в табл. 12.

В 2012 г. появился рейтинг университетов Азии, а осенью 2013 г. и рейтинг университетов стран БРИКС и развивающихся экономик. Оба эти рейтинга строятся по той же методике, что и глобальный рейтинг.

Кроме глобальных рейтингов университетов существуют еще и рейтинги публикационной активности. Эти рейтинги строятся на основании данных как Web of Science CC, так и Scopus.

На базе Scopus исследовательской группой SCImago из университета Гранады составляется рейтинг исследовательских организаций мира (SCImago Institutional Ranking). В рейтинг 2014 г. вошли 132 организации России (всего в рейтинге 4851 организация), из них 29 университетов (всего — 2713). Ранжирование осуществляется на основании данных о публикациях за предыдущие 5 лет и включает следующие показатели:

- Output — общее число опубликованных научных работ;
- Scientific Talent Pool — общее количество авторов из организации;
- Excellence Rate — число работ, опубликованных в изданиях, входящих в 10 % наиболее цитируемых в мире (по научным областям);
- Scientific Leadership — число опубликованных научных работ, основные авторы которых (corresponding author) указывают свою принадлежность к организации;
- Excellence with leadership — количество документов, включенных в показатель

Таблица 12

Сводные данные по весам групп показателей

Группа показателей/ научная область	Качество образования	Исследовательская репутация	Цитирование	Доход от промышленности	Интернационализация
Arts&Humanities	37,5	37,5	15	2,5	7,5
Clinical, Preclinical&Health, Life Sciences and Physical Sciences	27,5	27,5	35	2,5	7,5
Engineering & Technology	30	30	27,5	5	7,5
Social Sciences	32,5	32,5	25	2,5	7,5

Excellence rate, в которых основной автор относится к рассматриваемой организации;

- International Collaboration — международное сотрудничество (по количеству совместных с зарубежными партнерами публикаций);

- Normalized Impact — среднее цитирование научных работ организации, приведенное к среднемировому цитированию;

- High Quality Publications — число научных работ, опубликованных в самых авторитетных журналах (издания, входящие в верхний квартиль распределения по авторитетности в соответствии с рейтингом SCImago Journal Rank SJR по научным областям);

- Specialization Index — соответствие тематики научных работ специализации организации.

По данным о публикациях в Web of Science CC составляется Лейденский рейтинг, основной особенностью которого является возможность расчета всех показателей с учетом доли соавторства ученых из разных университетов. Доля соавторства в данном случае учитывается строго пропорционально авторам из разных организаций, без учета их порядка (как в Шанхайском рейтинге для Science & Nature) или «главности», обозначаемой для “corresponding author” (как в показателе Scientific Leadership в SciMago Institutional Ranking). Показатели в этом рейтинге можно подразделить на две основные группы — показатели влияния (Impact Indicators) и сотрудничества (Collaboration Indicators).

К первой группе относятся следующие:

- Mean citation score (MCS). Среднее количество цитирований публикаций университета.

- Mean normalized citation score (MNCS). Среднее цитирование публикаций университета, нормализованное по области знаний, виду публикаций, году публикации.

- Proportion top 10 % publications (PPtop 10 %). Доля публикаций университета, вхо-

дящих в 10 % самых цитируемых, в общем числе публикаций.

Вторая группа характеризует как сотрудничество с другими организациями, так и международное сотрудничество:

- Proportion collaborative publications (PPcollab). Доля публикаций университета в соавторстве с другими организациями.

- Proportion international collaborative publications (PPint collab). Доля публикаций университета в соавторстве с двумя и более странами.

- Mean geographical collaboration distance (MGCD). Среднее географическое расстояние между соавторами.

- Proportion long distance collaborative publications (PP > 1000 km). Доля публикаций университета, написанных в соавторстве с организациями, расположенными на расстоянии более чем 1000 км.

В рейтинг включаются университеты, занимающие первые 500 мест по количеству публикаций за 4 года в целом (учитываются только типы публикаций Article и Review) с возможностью проводить ранжирование по любому из показателей как с учетом доли соавторства (т. е. фракционированные показатели), так и без учета, т. е. по всему набору публикаций. С 2014 г. первичный отбор университетов для дальнейшего ранжирования осуществляется по количеству публикаций не во всех журналах, индексируемых Web of Science CC, а только в основных (core journals), правила отбора которых и список приведены на сайте <http://www.leidenranking.com/methodology/indicators#core-journals>. Таким образом, из исходного подсчета публикации, по расчетам составителей рейтинга, исключаются порядка 16 % всех публикаций, проиндексированных в базе данных Web of Science CC.

Еще один рейтинг Performance Ranking of Scientific Papers for World Universities

(HEEACT), который строится на основе публикаций, проиндексированных в Web of Science, с 2007 г. составляет Национальный университет Тайваня. Критериями отбора университетов для дальнейшего анализа и составления рейтинга являются следующие:

1. Попадание в Топ-700 организаций, представленных в аналитическом разделе Essential Science Indicators на платформе InCites.

2. Присутствие в рейтингах ARWU, QS, THE, U.S.News.

Получившийся перечень университетов анализируется по параметрам, приведенным в табл. 13, которые также берутся либо из Essential Science Indicators, либо непосредственно по базе данных Web of Science CC, и составляется окончательный рейтинг, содержащий 500 университетов.

Кроме общего рейтинга Тайваньский университет составляет и рейтинги по широким областям (6 областей) и предметным

категориям (14 категорий, относящихся к естественным и техническим наукам).

Описанные выше рейтинги не исчерпывают весь спектр существующих рейтингов университетов, но являются наиболее известными.

Таким образом, в описанных рейтингах присутствует практически весь спектр возможных показателей, от простого количества публикаций до сложных показателей нормализованного цитирования и использования фракционирования количества публикаций и цитирований для корректного учета вкладов конкретных университетов. Размер организации практически везде учитывается путем нормирования показателей на количество преподавателей (Faculty) или Academic Staff (научно-педагогических работников в российской терминологии).

Российские университеты в международных рейтингах представлены весьма скромно, при этом основные проблемы связаны

Таблица 13

Методология рейтинга Performance Ranking of Scientific Papers for World Universities (HEEACT)

Критерий	Индикаторы 2013	Вес индикатора	
Исследовательская продуктивность	Количество статей за последние 11 лет (2002–2012)	10	20
	Количество статей в текущем году (2012)	10	
Влиятельность исследований	Количество цитирований за последние 11 лет (2002–2012)	10	30
	Количество цитирований за последние 2 года (2011–2012)	10	
	Среднее количество цитирований за последние 11 лет (2002–2012)	10	
Исследовательское превосходство	h-index за последние 2 года (2011–2011)	20	50
	Количество высокоцитируемых статей (2002–2012)	15	
	Количество статей в высокорейтинговых журналах в текущем году* (2012)	15	

*Журналы, входящие в верхние 5 % по импакт-фактору в каждой предметной категории.

именно с показателями, характеризующими как количество, так и качество научных публикаций. Какие можно сделать рекомендации для сотрудников и администрации университетов по улучшению публикационных показателей с учетом всех изложенных выше способов учета и оценки публикаций и их применения для составления рейтингов?

На сегодняшний день самым реальным тактически представляется расширение международного сотрудничества, увеличивающего как долю международных публикаций, так и возможность публикаций российских авторов в рейтинговых международных журналах. Эффективным также может оказаться расширение практики публикации научных результатов в журналах открытого доступа или использование опции открытого доступа в журналах с гибридной моделью публикации. Открытый доступ потенциально способствует более быстрому цитированию научных статей, т. е. оказывает существенное влияние на показатели цитируемости, используемые при составлении рейтингов университетов.

На уровне каждого конкретного университета возможны следующие направления, способствующие как повышению публикационной активности, так и лучшему представлению научных результатов университета в международных индексах цитирования.

Административные меры:

- строгая регламентация указания названия университета как места работы автора

в публикациях, что позволит избежать ошибок и неточностей при индексации публикаций университета;

- введение требований по наличию публикаций при проведении конкурсов на замещение должностей научно-педагогических работников, позволяющее повышать уровень кадрового состава.

Стимулирующие меры:

- установление доплат или премирование в зависимости от количества публикаций в международных базах, их цитирования;
- компенсация оплаты публикаций открытого доступа в рейтинговых журналах и т. д.;
- повышение квалификации научно-педагогических работников, обучение аспирантов и студентов основам академического письма, использованию электронных ресурсов и наукометрических инструментов для планирования своей научной деятельности.

Инфраструктурные меры — расширение подписки на электронные ресурсы, позволяющие ученым знакомиться с самой актуальной научной информацией и планировать направления своих исследований в соответствии с мировыми трендами, модернизация научного оборудования и др.

Все это позволит развивать научные исследования, представляющие интерес для мирового научного сообщества и, следовательно, способствующие появлению научных публикаций высокого уровня.

Акопов А. И. Научные журналы: обзор научных работ и попытка типологической дифференциации на фоне социально-экономических и профессиональных проблем // Научно-культурологический журнал. № 12 (157). 25.08.2007. URL: <http://www.relga.ru/Environ/WebObjects/tgu-www.woa/wa/Main?textid=2024&level1=main&level2=articles>

Бредихин С. В., Кузнецов А. Ю. Методы библиометрии

и рынок электронной научной периодики. Новосибирск: ИВМиМГ СО РАН, НЭИКОН, 2012. 256 с.

Бредихин С. В., Кузнецов А. Ю., Щербак Н. Г. Анализ цитирования в библиометрии. Новосибирск: ИВМиМГ СО РАН, НЭИКОН, 2013. 344 с.

Гиляревский Р. С., Шапкин А. В., Белозеров В. Н. Рубризатор как инструмент информационной навигации. СПб.: Профессия, 2008. 352 с.

Маршакова И. В. Система связей между документами, построенная на основе ссылок: по данным Science Citation Index // НТИ. Сер. 2. 1973. № 6. С. 3–8.

Налимов В. В., Мульченко З. М. Наукометрия. Изучение развития науки как информационного процесса. М.: Наука, 1969. 192 с.

Писляков В. В., Дьяченко Е. Л. Эффект Матфея в цитировании статей российских ученых, опубликованных за рубежом // НТИ. Сер. 2. Информационные процессы и системы. 2009. № 3. С. 19–24.

Прайс Д., Бивер Д. Сотрудничество в «невидимом колледже» / пер. с англ. М. К. Петрова // Коммуникация в современной науке: сб. переводов / сост., общ. ред. и вступ. ст. Э. М. Мирского, В. Н. Садовского. М.: Прогресс, 1976.

Ранганатан Ш. Р. Классификация двоеточием. Основная классификация / пер. с англ.; под ред. Т. С. Гомолицкой. М.: ГПНТБ СССР, 1970.

Хайтун С. Д. Наукометрия. Состояние и перспективы. М.: Наука, 1983. 343 с.

Alonso S., Cabrerizo F., Herrera-Viedma E., Herrera F. hg-index: A new index to characterize the scientific output of researchers based on the h- and g-indices // Scientometrics. 2010. Vol. 82. № 2. P. 391–400.

Beall J. Predatory publishers are corrupting open access // Nature. 2012. Vol. 489. P. 179.

Berelson B. Graduate education in the United States. N. Y.: McGraw-Hill Book Co, 1960. P. 55

Bonitz M., Bruckner E., Scharnhorst A. Characteristics and Impact of the Matthew Effect for Countries // Scientometrics. 1997. Vol. 40. № 3. P. 407–422.

Bonitz M., Bruckner E., Scharnhorst A. The Matthew Index — Concentration Patterns and Matthew Core Journals // Scientometrics. 1999. Vol. 44. № 3. P. 361–378.

Bonitz M., Scharnhorst A. Competition in Science and the Matthew Core Journals // Scientometrics. 2001. Vol. 51. № 21. P. 37–54.

Börner K., Jeegar T. Maru, Robert L. Goldstone The Simultaneous Evolution of Author and Paper Networks // Proc. Natl. Acad. Sci. USA. 2004. Apr. 6; 101 Suppl 1:5266–73.

Chubin D. E., Moitra S. D. Content analysis of references: adjunct or alternative to citation counting. Social Studies of Science. 1975. Vol. 5. P. 423–441.

Garfield E. Citation Indexes for Science. A new Dimensions in Documentation through association of ideas // Science. 1955. Vol. 122.

González-Pereira B., Guerrero-Bote V. P., Moya-Anegón F. A new approach to the metric of journals' scientific prestige: The SJR indicator // Journal of informetrics. 2010. № 4 (3). P. 379–391.

Gross P. L. K., Gross E. M. College libraries and chemical education // Science. 1927. Vol. 66. № 1713. P. 385–389.

Guerrero-Bote V.P., Moya-Anegón F. A further step forward in measuring journals' scientific prestige: The SJR2 indicator // Journal of Informetrics. 2012. № 6 (4). P. 674–688.

Hirsch J. An index to quantify an individual's scientific research output // Proceedings of the National Academy of Sciences. 2005. Vol. 102. P. 16569–16572.

Kessler M. M. Bibliographic coupling between scientific papers // American Documentation. 1963. Vol. 14. Iss. 1. P. 10–25.

Latour B., Woolgar S. Laboratory life: The construction of scientific facts. N. J.: Princeton, 1986. 297 p.

Merton R. K. The Matthew Effect in Science // Science. January 1968. Vol. 5. №. 159 (3810). P. 56–63.

Merton R. K. The Sociology of Science ed by N. W. Storer. Chicago: University of Chicago Press, 1973.

Moed H. F. Citation Analysis in Research Evaluation. Springer, 2005. 346 p.

Moed H. F. Measuring contextual citation impact of scientific journals // Journal of Infometrics. 2010. Vol. 4. P. 265–277.

Nicolaisen J. Social behavior and scientific practice: Missing pieces of the citation puzzle: PhD thesis. Copenhagen: 321 Royal School of Library and Information Science, 2004. 214 p. URL: <http://books.google.com>.

Page L., Brin S., Motwani R. and Winograd T. The PageRank Citation Ranking: Bringing Order to the Web. Technical Report. 1998. Stanford InfoLab. URL: <http://ilpubs.stanford.edu:8090/422/>.

Prathap G. Hirsch-type indices for ranking institutions' scientific research output // Current Science. 2006. Vol. 91. Iss. 11. P. 1439–1439.

Price D. Little Science, Big Science. N. Y.: Columbia Univ. Press, 1963 [Издание на рус. яз.: Прайс Д. Малая наука, большая наука // Наука о науке / под ред. В. Н. Столетова. М., 1996. С. 281–384.]

Small H. Co-citation in the scientific literature: A new measure of the relationship between documents // J. Amer. Soc. Inform. Sci. 1973. Vol. 24. P. 256–269.

Thelwall M., Haustein S., Larivière V., Sugimoto C. R. Do Altmetrics Work? Twitter and Ten Other Social Web Services // PLoS ONE. 2013. Vol. 8. Iss. 5. Art. number e64841.

Wouters P. The signs of science // Scientometrics. 1998. Vol. 41. Iss. 1–2. P. 225–241.

Wouters P. Beyond the holy grail: From citation theory to indicator theory // Scientometrics. 1999. Vol. 44. Iss. 3. P. 561–580.

КАРТИРОВАНИЕ НАУКИ И ТЕХНОЛОГИИ, ПРОГНОЗ РАЗВИТИЯ



М. А. Акоев
Заместитель
директора Центра
мониторинга на-
уки и образования
Уральского феде-
рального универси-
тета имени первого
Президента России
Б. Н. Ельцина.

MAPPING SCIENCE AND TECHNOLOGY, FORECASTING RESEARCH AND DEVELOPMENT

DOI 10.15826/B978-5-7996-1352-5.0007

The chapter explores visualization of research documents relationships through subject, semantic, chronological and geographical mapping. Examples from different citation indexes and patent databases are used. Methods of analyzing a selection of documents using such visualizations are demonstrated. Possibilities and limitations of R&D forecasting are examined. The author highlights the ways to improve research productivity, and identify new promising directions for research.

В данной главе рассматриваются вопросы визуализации взаимосвязей документов в тематическом, семантическом, хронологическом и географическом представлениях на примерах публикаций из различных указателей цитирования и патентных баз. Описаны методы проведения анализа набора документов с использованием такой визуализации. В главе рассмотрены вопросы ограничений в прогнозировании научно-технического развития и аспекты повышения качества существующих научных направлений и развития новых направлений в организации.

При переходе от наукометрического анализа отдельных ученых, коллективов и организаций к их сравнению аналитик сталкивается с необходимостью выявлять закономерности в больших объемах информации. Использование приемов визуализации данных позволяет сократить затраты на нахождение зависимостей и в дальнейшем — на представление результатов анализа. Помимо построения диаграмм к ним относится также и представление сложноорганизованной информации в формате карт

(картирование), на которых можно показать связи анализируемых элементов. С вопросами картирования наукометрической информации тесно связаны задачи построения прогноза развития науки и технологии и задачи повышения качества существующих и развития новых направлений исследований в научных организациях.

4.1. Картирование науки и технологии

Картирование, являясь одним из приемов наглядного представления данных, подчиняется общим

принципам визуализации, т. е. компактному представлению большого объема информации, делающему заметными найденные в ходе анализа закономерности¹. Информация о типах и уровнях наукометрического анализа с примерами названий визуализаций приведена в табл. 14, построенной на основании классификации, приведенной в [Börner, Polley, 2014, p. 7]. В данном разделе для каждого типа из табл. 14 мы рассмотрим характерные примеры анализа с целью показать возможности и ограничения конкретных типов визуализации в применении к наукометрическим исследованиям. Каждый из представленных типов наукометрического анализа отвечает на один из вопросов о предмете исследования (вопросы приведены в скобках). На практике чаще всего встречается комбинация разных типов анализа, например, карта распределения публикационной активности организаций РАН будет одновременно отвечать на вопросы: где? с кем? и когда?

В настоящий момент наиболее исчерпывающие сведения о приемах научной визуализации содержатся в руководстве [Börner, Polley, 2014], а также в открытом онлайн-курсе тех же авторов [IVMOOC, 2014]. Примеры визуализаций, представленные в вышеуказанном руководстве и курсе, построены с использованием интегрированной платформы анализа и визуализации данных Sci²

Tool, разработанной в Cyberinfrastructure for Network Science Center Университета Индианы [Sci² Tool, 2010]. Sci² Tool представляет собой модульный набор программных средств для изучения и обработки наукометрической информации, который поддерживает проведение всех типов наукометрического анализа на всех уровнях, описанных в табл. 14. Платформу Sci² Tool можно рекомендовать для изучения приемов визуализации и дальнейшей работы, при этом необходимо учитывать, что некоторые возможности по импорту и анализу данных ориентированы на данные о науке США (например, данные о грантах NSF, привязанные к формату даты США).

Примерами статистического анализа и построения профилей авторов служат визуализации в форме разнообразных диаграмм и таблиц данных. Систематическое введение в визуализацию с использованием диаграмм можно получить, обратившись к книге [Желязны, 2009] и учебнику [Яу, 2013]. Книга [Желязны, 2009] представляет приемы использования диаграмм для коммуникаций в процессе поддержки принятия решений. В учебнике [Яу, 2013] подробно описаны технические аспекты подготовки данных для визуализации. Кратко остановимся на важных аспектах визуализации, проиллюстрировав их на примере визуализации диаграммы и таблицы исходных данных.

Визуализация числа элементов в количестве большем, чем можно удерживать в оперативной памяти, нецелесообразна. Объем оперативной памяти человека определяет «магическое» число семь плюс минус два («кошелек Миллера»). Сравните представление одной и той же информации на рис. 36а и 36б (см. с. 187). В качестве сравниваемых элементов выступают десять

¹Проект «Карта российской науки» Министерства образования и науки РФ (<http://www.mapofscience.ru/>) является информационной системой для регулярного и автоматического обновления базы данных ученых и организаций, включая показатели их деятельности, осуществления статистического анализа научно-исследовательской активности и обеспечения основы для создания аналитических материалов о состоянии российского сектора научных исследований и разработок. В «Карте российской науки» использованы отдельные методы картирования, представленные в данном разделе.

Таблица 14

Соотношение типов наукометрического анализа (в скобках указан вопрос, идентифицирующий субъект) и уровней наукометрического анализа

Тип наукометрического анализа	Уровень наукометрического анализа		
	Микро/ Индивидуальный (1–100 записей)	Промежуточный/ Локальный (101–10 000 записей)	Макро/ Глобальный (более 10 000 записей)
Статистический анализ / Построение профиля (Кто?)*	Экспертные данные о персональной научной продуктивности	Исследовательские центры, университеты, предметные области или государства	Все публикации США или все публикации в мире
Геопространственный анализ (Где?)	Индивидуальная траектория перемещения исследователя	Распределение мест работы авторов публикаций в научном журнале	Распределение публикационной активности организаций РАН
Темпоральный анализ (Когда?)	Динамика получения индивидуальных грантов	Развитие тематик институтов отделения физических наук РАН за 20 лет	113 лет физических исследований
Тематический анализ (Что?)	Ключевые слова, указанные в записях о цитируемых источниках в публикациях по теме исследования	Тематики публикаций институтов отделения химии и наук о материалах РАН	Карты ключевых слов и тематики исследований РАН
Сетевой анализ (С кем?)	Сеть организаций, ученые которых выполняют работы в соавторстве с данной научной группой	Сеть соавторства между исследователями по анализируемой тематике	Ключевые компетенции РАН

*В скобках указан вопрос, идентифицирующий субъект. В ячейках таблицы приведены примеры названий визуализаций. Таблица основана на классификации, приведенной в [Börner, Polley, 2014, p. 7], с изменением примеров.

тематик, а не отдельные характеристики в ячейках таблицы. Единственное преимущество диаграммы перед табличным представлением — это возможность визуально оценить степень отличия абсолютных значений, сравниваемых элементов сразу по нескольким показателям; например, число

публикаций в первых трех областях приблизительно одинаково, но они отличаются по нормализованной цитируемости.

Визуализация как метод представления информации не может компенсировать отсутствие или низкое качество исходных данных. Также визуализация не может вы-

явить закономерности, которые не обнаруживаются количественными методами анализа. Визуализация делает закономерности в предварительно собранных и обработанных данных более доступными для восприятия. Визуализация, по сути, представляет собой краткое обобщение результатов анализа, которое позволяет понять контекст при переходе к данным, лежащим в основе анализа. Доверие к представленным в форме визуализации результатам анализа будет высоким, если обеспечивается верифицируемость результатов анализа, т. е. получение всех элементов представляемого отчета может быть прослежено до первичных данных и методик их обработки.

Пример визуализации, представленный на рис. 36а и 36б (см. с. 187), иллюстрирует необходимость обеспечения верифицируемости данных при принятии решений. Диаграмма была подготовлена для принятия решений о выборе мероприятий по развитию десяти анализируемых направлений (названия тематик для дальнейшего обсуждения несущественны, публикации были отобраны в ручном режиме с привлечением экспертов в предметной области и представляют узкие тематические направления). Для всех анализируемых областей характерен уровень цитируемости ниже, чем среднемировой, что позволяет для всех научных групп сформулировать задачу повышения видимости их научного результата в мире. Для научных групп, которые представляют тематику Subject 10, только из анализа представленных данных можно было бы сделать вывод, что при выделении дополнительного финансирования можно ожидать роста качественных показателей публикаций при сохранении количества публикаций. К данному выводу можно прийти исходя из того, что ученые,

опубликовавшие анализируемые работы, при высоких темпах роста смогли сохранить самый высокий уровень востребованности своих результатов в анализируемых предметных областях (косвенно об этом свидетельствует значение нормализованной цитируемости публикаций). Однако при обращении к первичным данным о цитируемости публикаций данной предметной области выявляется высокий процент самоцитирования. Из этого следует, что на фоне увеличения числа публикаций по тематике группа в меньшей степени ориентирована на привлечение внешнего внимания к своим работам. Принятие решений по выделению ресурсов для данной группы требует дополнительного анализа с привлечением экспертов в этой предметной области.

Возможность визуализации данных зависит от наличия готовых инструментов построения визуальных представлений. Только опыт регулярных созданий и представлений визуализаций при наличии инструментов и подготовленных для визуализации данных гарантирует создание оптимального, с точки зрения восприятия, типа визуального представления результатов. Многие типы визуального представления результатов, обобщающие опыт многих пользователей, встроены в библиометрические базы данных или аналитические надстройки над ними (подробнее про библиометрические инструменты см. в разделе 2.1). Однако при проведении исследований с одновременным использованием разных типов анализа (табл. 14) штатных средств анализа обычно бывает недостаточно. Приходится прибегать к использованию готовых инструментов или создавать свои инструменты визуализации под конкретный вид представления или подготовку данных для визуализации. Обзор готовых

средств визуализации приведен в [Science mapping..., 2011].

Картирование — одна из сложнейших форм визуализации как по трудоемкости процесса подготовки данных, так и по восприятию и интерпретации результата потребителем. Трудоемкость подготовки данных определяется необходимостью, с одной стороны, указать координаты на карте для каждого отображаемого элемента (публикации, исследователя и т. п.), а с другой — получить «картографическую» основу, полезную для целей анализа и отражающую представление о реальности. Например, в случае геопространственных визуализаций «картографическая» основа отражает реальный мир и распределение элементов в географическом пространстве, а в случае тематических карт или схем — соотнесение тематик, которые отражают науку ровно в той степени, в которой ее отражают исходные данные, использованные для построения основы карты. Существенное отличие тематических «картографических» основ от географических в том, что мы не можем на момент построения знать все тематики исследований с учетом будущих результатов исследований. Тематические карты не обладают предсказательной силой «белых пятен» географических карт.

Основная ценность картографического способа визуализации данных заключается в возможности продемонстрировать распределение элементов в условном пространстве карты, показывая расстояния между ними. Зная распределение элементов в пространстве карты, мы можем находить в окрестности выбранного элемента все существующие элементы, которые считаются близкими, например тематически связанные публикации одного временного периода. Именно возможность определить окрестности выбранного элемента делают

картографические способы отображения полезными для анализа, так как позволяют найти эталонный набор элементов для сравнения. Отметим, что поиск окрестностей элемента не является следствием их картографического отображения. Обе возможности — как поиск, так и отображение — являются следствием моделей, лежащих в «картографических» основах. Обзор подходов к выбору моделей для построения «картографических» основ представлен в [Mapping research..., 2008].

4.1.1. Геопространственный анализ (где?)

В качестве примера геопространственного анализа построим карту мест работы авторов, указываемых в аффилиации в журнале ACM Transactions on Database Systems (TODS). Публикации в научном журнале относятся к узкой тематике, исключения составляют мультидисциплинарные журналы (подробнее обсуждение классификаций журналов см. в разделе 3.7). Следовательно, по аффилиации авторов можно отследить, где расположены научные центры, в которых выполняются работы по узкой тематике. Результат анализа представлен на рис. 37 (см. с. 188). Информация о координатах каждого места работы была получена геокодированием адресной информации из аффилиации авторов в библиографических записях. Геокодирование — это автоматизированный процесс сопоставления почтовых адресов с координатами на карте (в данном анализе использовалась служба Geocoding API Google²). За анализируемый период, с 2008 г. по первый выпуск 2014 г., в журнале была опубликована 181 работа, что составляет порядка 21 % всех публикаций в журнале с момента его создания (журнал

²<https://developers.google.com/maps/documentation/geocoding/> (дата обращения: 25.06.2014).

выходит четыре раза в год). Данные работы написали 437 авторов, которые указали 420 текстуально разных вариантов аффилиации (некоторые авторы указали более одной аффилиации). Из представленной на рис. 37 (см. с. 188) карты видно, что авторы за небольшим исключением сконцентрированы в Европе и США.

Процесс, использующий автоматическое геокодирование, не всегда возвращает правильные координаты. Например, отметка в Тихом океане у берегов Чили является результатом геокодирования следующей аффилиации: «Departamento de Ciencias de la Computación, Escuela de Ingeniería, Pontificia Universidad Católica de Chile, Chile» (при повторной обработке в июле 2014 г. геокодирование данной строки возвращало координаты в г. Сантьяго). Также при геокодировании необходимо учитывать, что не все аффилиации могут быть корректно найдены, например, строка «Microsoft Research Asia, Beijing, China» может быть распознана по координатам географического центра Пекина, а Microsoft вернет неопределенное значение в силу множества подходящих вариантов. При геокодировании адресов необходимо учитывать, что, например, для действительных членов РАН геокодирование аффилиации может возвращать координаты адреса: г. Москва, Ленинский просп., д. 14.

Применение созданных при разработке геоинформационных систем (ГИС) и активно применяемых в географии методов геопространственного анализа к наукометрическим данным находится в начальной стадии освоения. Систематическое применение методов геопространственного анализа требует использования специализированных пакетов ГИС. Для начального знакомства с темой можно порекомендовать

обратиться к форуму GIS-Lab³ и в дальнейшем перейти к чтению специализированной литературы.

4.1.2. Темпоральный анализ (когда?)

Исторически темпоральный анализ как инструмент анализа наукометрических данных появился первым, что связано с использованием SCI для установления научных приоритетов в работах 1940-х гг. Первую публично доступную программу визуализации публикаций на временной шкале с указанием связей между ними (HistCite) создал Юджин Гарфилд [HistCite, 2012]. Пример визуализации с использованием HistCite можно увидеть в статье Ю. Гарфилда «Эволюция Science Citation Index», включенной в данное руководство.

При использовании инструмента темпорального анализа в визуализации возможны два варианта представления времени: либо в качестве одной из осей, либо в последовательности анимированных изображений. Второй способ применяется для визуализации структур, которые требуют двумерного отображения, например, динамика во времени и географическое распределение публикационной активности авторов журнала. Визуализировать время для сложных структур также можно, используя цветовое кодирование времени, если это уместно. Однако этот способ менее распространен, так как затрудняет чтение карт, перегружая их визуальной информацией.

Самым простым примером темпорального анализа является столбчатая диаграмма распределения числа публикаций во времени. Для демонстрации возможностей темпорального анализа, требующего отображения на карте, приведем распределение во времени связей между кластерами публикаций авторов, прямо или косвенно

³<http://gis-lab.info>

цитирующих самую раннюю из самых цитируемых в наукометрии публикаций математика А. Дж. Лотки [Lotka, 1926] (рис. 38, см. с. 189). Визуализация выполнена в инструменте CitNetExplorer, развивающем идеи инструмента HistCite [CitNetExplorer, 2014]. CitNetExplorer позволяет визуализировать большие массивы публикаций и кластеризовать их, скрывая несущественные детали [Eck, Waltman, 2014].

Визуализация, представленная на рис. 38 (см. с. 189), была выполнена на основе публикаций за период 1945–2014 гг., полученных из Web of Science CC для 14 журналов, перечисленных в табл. 15. Всего было найдено 37 989 публикаций и 232 202 ссылки на них в анализируемом массиве. Визуализация воспроизводит пример, приведенный в лекции Н. Ван Эйка (N. Van Eck) о возможностях CitNetExplorer на конференции STI в 2014 г.

Визуализация кластеров анализируемых публикаций позволяет увидеть все ключевые фамилии авторов, оказавших наибольшее влияние на развитие наукометрии и упоминавшихся в предыдущих главах. Также на представленном рис. 38 (см. с. 189) можно увидеть связи группы публикаций по наукометрии (они выделены зеленым

цветом) с группой публикаций по информационному поиску (выделены синим цветом), например, с публикациями Г. Сэлтона (G. Salton) и с группой публикаций автора концепции тройной спирали Г. Ицковица (H. Etzkowitz) [Ицковиц, 2010] (выделены фиолетовым цветом).

4.1.3. Сетевой анализ (с кем?)

Для исследования науки как процесса научной коммуникации можно использовать инструменты сетевого анализа, развитые в дисциплине Social Network Analysis (SNA). Сетевой анализ позволяет выявить закономерности в совместных работах отдельных авторов и научных организаций. Однако возможности применения сетевого анализа для наукометрических исследований реализуются в полной мере только при наличии машиночитаемой информации о научных связях ученых, о полученных ими результатах и применении специализированных инструментов. В качестве инструмента для знакомства с возможностями сетевого анализа рекомендуется использовать Sci² Tool [Sci² Tool, 2010]. В качестве инструмента анализа и визуализации сетей большого размера можно порекомендовать бесплат-

Таблица 15

Журналы по наукометрии с 1945 г., индексируемые в Web of Science CC

American Documentation	Journal of Information Science
The Annual Review of Information Science & Technology	Journal of Informetrics
Aslib Proceedings	Journal of the American Society for Information Science
Information Processing & Management	Journal of the Association for Information Science and Technology
Information Scientist	Research Evaluation
Information Storage and Retrieval	Research Policy
Journal of Documentation	Scientometrics

ную программу Pajek [Pajek, 1996]. Однако для проведения сложных видов анализа необходимо владеть компетенциями для работы со специализированными продуктами и освоить навыки программирования. Если у аналитика отсутствуют необходимые компетенции программирования, то для преобразования и подготовки данных для анализа в проект нужно привлекать программиста. Однако аналитику желательно самостоятельно освоить основы программирования в одном из онлайн-курсов анализа и обработки данных⁴. Для большинства задач подготовки данных требуются минимальные навыки программирования. Дополнительным плюсом от освоения программирования будет улучшение взаимопонимания с программистами, которым вы ставите задачи.

В качестве примера сетевого анализа построим сеть соавторства (рис. 39, см. с. 190) для публикаций четырех авторов: Стенли Васерман (S. Wasserman), Юджин Гарфилд (E. Garfield), Алесандро Веспиньяни (A. Vespignani) и Альберт-Ласло Барабаши (A.-L. Barabasi), 361 запись, период — 1955–2007 гг., источник — Web of Science CC. Построение выполнено с использованием Sci² Tool, исходные данные визуализации входят в комплект поставки данного инструмента. Методика построения сетей описана в [Börner, Polley, 2014, p. 204–206]. Итак, соавторы образуют четыре четко выраженных кластера. Два кластера публикаций, Стенли Васермана и Юджина Гарфилда, не связаны между собой и с остальными кластерами и, таким образом, образуют изолированные «острова» соавторства. Выявление изолированных «островов» в сетях соавторства для публикаций по одной тематике может

свидетельствовать как о возможности наладить новые научные связи, так и о неверном отнесении публикаций к тематике. Выполняя анализ только на основе данных о соавторстве, без привлечения дополнительной информации, например о ко-цитировании, можно прийти к неверным выводам, так как перечисление в списке авторов не несет информации о вкладе каждого из них в написанную статью.

Визуализации сетевого анализа на промежуточном и макроуровнях (табл. 14) становятся нечитаемыми, так как содержат множество узлов с одной-двумя связями, загромождающими построенную карту. Концентрация связей позволяет визуально подчеркнуть центры анализируемой активности, однако затрудняет общее восприятие и делает метки нечитаемыми. Для представления карт, построенных по результатам сетевого анализа, рекомендуется сокращать концентрацию связей, используя вместо них цветовую раскраску, например тепловую карту (heatmap), или используя в качестве дополнительного параметра кольцо вокруг узла с данными.

Сетевой анализ не замыкается только на анализе библиометрических данных. Ученые, активно применяя современные коммуникационные возможности при работе с информацией, дают аналитикам возможность использовать при анализе данные, сбор которых ранее был затруднен или невозможен. В качестве примера можно привести сетевой анализ и визуализацию истории переходов читателей между статьями на сайтах издательств [Clickstream Data..., 2009]. За два года, с 2007-го по 2008-й, была собрана информация о более чем одном миллиарде обращений к страницам полнотекстовых ресурсов, данные о переходах и идентификации пользователей

⁴<https://www.coursera.org/courses?categories=stats>

представили издательства. Анализ, выполненный в статье, базируется на предположении, что большинство переходов на страницы издательств с целью получения полного текста были предприняты для чтения запрашиваемой работы как минимум с целью ознакомления. Полученная в результате анализа карта (рис. 40, см. с. 191) представляет данные о связях между журналами на основе анализа динамики перехода читателей между страницами изданий. Связи между журналами отражают статистически значимое число переходов в процессе чтения.

На карте сетевого анализа можно увидеть ожидаемую концентрацию переходов между журналами одной тематики и переходы между журналами традиционно связанных тематик. Однако из карты следует «неожиданный» результат — в центре карты область социально-гуманитарных журналов оказалась связана с областями естественных, технических и медицинских наук, что не соответствует картам, которые строятся только на основании данных о цитировании публикаций. В качестве дополнительного примера сетевого анализа в применении к наукометрии гуманитарных дисциплин можно привести исследование Р. Коллинза [Collins, 2000]. В рамках данного исследования были выявлены и проанализированы связи между философами на протяжении более чем двух тысяч лет и было показано, что в философии наблюдаются процессы, идентичные тем, которые мы видим в естественных науках. В качестве примера выбора источников данных для проведения сетевого анализа внутри организации можно обратиться к диаграмме, представленной в разделе 1.5.

4.1.4. Тематический анализ (что?)

Ответы на вопросы «Что является предметом исследований?» и «Как связаны

предметы исследований между собой и с практикой?» — это священный Грааль наукометрии. Основным препятствием в ответах на оба вопроса является научный процесс, который ориентирован на порождение компактного представления новых знаний для нужд практики, а не для облегчения наукометрических исследований (подробнее см. в разделе 1.2). Предмет исследований надежно идентифицируется только комбинацией трех элементов: текстов с представлением результатов исследований, контекста привлеченных для аргументации ссылок на факты и имплицитных знаний в голове сообщества исследователей и практиков, работающих над предметом исследований. Современные методы, используемые для идентификации предмета исследований, являются грубым приближением, ориентированным на сокращение числа документов, которые необходимо просмотреть и прочитать в процессе поиска информации. Достигнутое качество аппроксимации определяется прежде всего потребностями поиска информации, а не проведением наукометрических исследований. Для записей баз данных о публикациях считается достаточным, если классификация статьи базируется на журнальной классификации. Единственное неудобство возникает при классификации статей из мультидисциплинарных журналов (подробнее обсуждение классификаторов см. в разделе 3.7).

Возможности поиска информации в базах данных публикаций, основанных на словах в названии статей и источнике, в котором опубликован текст, могут быть дополнены поиском по ключевым словам. Ключевые слова способствуют уточнению тематики публикации, но не подменяют содержимое публикации. Два ограничения препятствуют широкому использованию ключевых

слов для идентификации предмета исследования. Это, во-первых, необходимость указывать достаточное число ключевых слов для полноты идентификации тематики статьи, а во-вторых, проблема унификации и синонимии с целью повышения точности поиска. Для многих российских статей по экономике одно из наиболее популярных ключевых слов — это «экономика». Преодолеть подобные ограничения может механизм приписывания ключевых слов на основе информации из библиографии, указанной в публикации, например инструмент KeyWords Plus, используемый в Web of Science CC. К сожалению, подобные механизмы не могут полностью заменить работу автора текста или издателя при генерации поисковых образов документов для всех записей в базе.

Несмотря на ограничения, присущие ключевым словам, их можно использовать прежде всего для проведения поиска, а также ориентироваться на них при выявлении тематик исследований. Для примера визуализируем по ключевым словам изменение тематики публикаций авторов из Казахстана с 2004 по 2014 г. (см. рис. 41 — с. 192), используя в качестве инструмента облако ключевых слов Wordle⁵ [Participatory..., 2009]. Число публикаций авторов из Казахстана, отраженных в Web of Science CC, за указанный период удвоилось (периоды сравнения — пять лет). Облако ключевых слов строится по значению поля KeyWords Plus, которое указано менее чем в половине записей. Размер ключевого слова отражает долю публикаций в общем числе публикаций. Из динамики ключевых слов и факта просмотра соответствующих им публикаций можно сделать вывод, что тема моделирования и астрофизики в Ка-

захстане на подъеме, а интерес к экологии водных ресурсов снизился.

На представленной визуализации можно увидеть и основное ограничение инструмента облака ключевых слов, а именно отсутствие унификации ключевых слов. Слово «модель» в единственном и множественном числе одновременно присутствует на диаграмме, и без обращения к исходным публикациям выяснить контекст применения этих понятий невозможно. Использовать визуализацию в виде облака ключевых слов желательно только на массиве, относящемся к одной предметной области.

Альтернативой облаку ключевых слов, лишенной их недостатков, является инструмент картирования на основе модели векторного пространства (Vector space model, VSM). Обзор современных методов приведен в [An approach to..., 2009]. Примером использования модели векторного пространства является построение патентных ландшафтов. Его применение возможно, если все записи на публикации в базе снабжены ключевыми словами, обеспечивающими полноту и точность описания публикации для поиска. В качестве примера такой базы данных можно привести международную реферативную патентную базу данных Derwent World Patents Index, размещенную на информационной платформе Thomson Innovation Thomson Reuters. Каждая запись о патенте снабжена описанием патентной формулы в форме ключевых слов, составленным экспертами вручную (в дополнение к патентной формуле, описанной в тексте патента), которое позволяет идентифицировать существенное сходство патентов, выданных в разной юрисдикции и разным правообладателям. Единицей анализа в Derwent World Patents Index выступает патентное семейство — пул

⁵<http://www.wordle.net/>

патентов, обладающих идентичной патентной формулой.

Пример патентного ландшафта приведен на рис. 42 (см. с. 193). В основе построения патентного ландшафта лежит метод измерения расстояний между двумя патентными семействами. Расстояние между ними определяется на основании числа ключевых слов описания патентной формулы, на которое отличаются два описания. Патентный ландшафт строится по пулу предварительно отобранных патентов, т. е. не существует единой картографической основы, на которую можно нанести патенты и увидеть их глобальное взаиморасположение. Если у двух патентных семейств, для которых вычисляется расстояние, нет общих ключевых слов, то расстояние вычисляется как минимальный путь из всех возможных путей, образуемых пулом анализируемых патентов, включающих оба семейства, между которыми вычисляется расстояние.

С помощью цветов на патентном ландшафте кодируются названия тематических областей. Области не изолированы между собой, например, область гражданского строительства в анализируемом пуле граничит с материаловедением, информационно-телекоммуникационными технологиями, электроникой и полупроводниками. Изменением оттенка и изолиниями на ландшафте показаны области с высокой патентной активностью, условные вершины на карте подписаны для ориентации. Так же как и остальные виды картографической визуализации, патентные ландшафты не позволяют принимать решения без дополнительного анализа патентных семейств. Например, условная низина на ландшафте, в которой у организации наблюдается концентрация патентов, может служить указанием на потенциально пер-

спективную область получения патентов для данной организации, но также может свидетельствовать о том, что область осталась в развитии. Возможность построения патентных ландшафтов подтверждает возможность выявления сходных патентов и их последующего анализа.

Основная ценность библиографических баз данных, обеспечивающих точность и полноту поиска по предмету исследований, заключается прежде всего в возможности снизить затраты специалистов на поиск публикаций. Если информация к каждой записи добавляется единообразно и в машиночитаемом виде, то возникает возможность найти близкие по тематике записи и использовать их для расчета средних показателей, которые могут быть референтными для сравнения показателей научной результативности как для отдельных ученых, так и для научных коллективов (подробнее см. в разделах 3.6 и 3.9). Подобные базы позволяют автоматизированно выявлять закономерности в данных о публикациях и на их основании косвенно судить о динамике приоритетов развития науки в ретроспективном режиме.

Эталонным примером подобных баз выступали реферативные журналы ВИНИТИ, в которых для каждой отраженной в них записи приведены сведения о классификации записей, ключевых словах и написан реферат специалистом в предметной области, который дает представление о содержимом работы. Создание и поддержание в актуальном состоянии подобных баз данных требует привлечения большого числа квалифицированных специалистов. Недостатком подхода к созданию баз «с ручным» реферированием является также и задержка с обновлением информации в базе, так как от момента появления публикации до

появления реферата на нее проходит полгода и более. Эффективно работать над развитием научного направления с подобными задержками затруднительно. Однако для задач мониторинга трендов развития науки и технологии подобный подход эффективен.

В некоторых областях науки возможно альтернативное решение по идентификации предметов исследований, снижающее зависимость от человеческого фактора при создании поискового описания документов. Например, в органической химии активно используются подходы по идентификации предмета исследований на основании структурной формулы соединений, упоминаемых в статье. В качестве подобной базы данных можно привести базу Chemical Abstracts (SciFinder), выпускаемую Американским химическим обществом. Используя единый подход к описанию структурных формул органических соединений, можно добиться однозначности в идентификации предмета исследований, но без обращения к тексту статьи невозможно получить однозначное понимание контекста, в котором упоминается данное вещество. Как и в примере с реферативными журналами ВИНТИ, в Chemical Abstracts создание поискового описания требует квалифицированного труда, однако поскольку для идентификации соединений в тексте не требуется полностью читать статью, обновление базы происходит без существенных задержек. Можно привести в качестве примера несколько баз, использующих аналогичные подходы к созданию описаний, основанных на возможности идентификации заболеваний по международной классификации болезней (International Statistical Classification of Diseases and Related Health Problems) или индекса художественных произведений.

Использованию идентификации как инструмента для поиска всех научных публикаций зачастую препятствует отсутствие надежных методов однозначной идентификации предметов исследований и высокие затраты на создание и актуализацию подобных баз. Затраты на создание баз данных для областей с разработанными методами идентификации в будущем могут быть снижены в результате развития методов автоматического извлечения информации, однако универсальных подходов к созданию методов однозначной идентификации на данный момент не создано.

Компромисса между скоростью создания поискового аппарата, его качеством, обеспечивающим полноту и точность поиска, и стоимостью актуализации библиографических баз данных сумели достичь при использовании такого инструмента, как цитирование — естественное для научной публикации средство указания на факты, привлекаемые для доказательства утверждений (обсуждение связи фактов и ссылок приведено в [Латур, 2013, с. 64–106], также см. раздел 3.3). Science Citation Index как инструмент в первую очередь сделал доступным перспективный поиск информации, а также позволил использовать информацию, аппроксимирующую предмет исследований публикации для автоматического анализа информации.

Использование информации о том, с какими свойствами цитировались публикации в анализируемом массиве публикаций, позволяет наглядно показать, как связаны группы публикаций. В качестве примера на рис. 43 (см. с. 194) приведена карта публикаций авторов СПбГУ, группировка работ выполнена по журналам, в которых они были опубликованы. Для визуализации использована программа VOSviewer [VOSviewer,

2013], позволяющая представлять результаты анализа методами «библиографического сочетания», ко-цитирования и соавторства [Eck, Waltman, 2010]. Основы перечисленных методов анализа рассмотрены в разделе 3.4. Визуализация методом «библиографического сочетания» позволяет судить о том, как связаны публикации по совместно используемым источникам информации: например, для карты на рис. 43 (см. с. 194) связь осуществляется по ссылкам на публикации в одних и тех же журналах. Классификация тематики публикаций закодирована цветом и соответствует классификации, используемой в Web of Science CC (см. раздел 3.7). На карте видно, что публикации по математике, физике, геологии и химии образуют последовательность сильно связанных звеньев одной цепи. К этим статьям примыкают публикации по астрономии и отдельные публикации по биологии. При этом основная часть биологических публикаций образует отдельный «архипелаг», что может говорить о том, что биологическая тематика и остальные исследования не имеют существенных общих предметов изучения в университете. Без проведения дополнительного анализа публикаций невозможно сделать вывод о причинах подобной обособленности.

Построение визуализации связей между публикациями на макроуровне способствует пониманию связей между научными дисциплинами. Часто подобным визуализациям дают название — карта науки. Такое громкое название, скорее, отражает амбиции авторов. В качестве примера на рис. 44 (см. с. 195) показана карта науки, построенная методом ко-цитирования. Принцип, определяющий размер и цветовое кодирование кругов, отражающих число публикаций в каждом журнале в выбранный год, тот же, что и на пре-

дыдущем рисунке. Взаимное расположение кругов на диаграмме определяется на основании числа цитирований идентичных наборов публикаций в работах, опубликованных в каждой паре журналов. В основе анализа, базирующегося на ко-цитировании, лежит предположение, что статьи связаны не только общим предметом исследований, но и контекстом, в котором происходит исследование общего предмета, если есть более чем одна общая цитируемая работа. Необходимо обратить внимание на отличие связей, существующих между научными дисциплинами в представленной карте науки, от визуализации связей между журналами, построенных на основании анализа истории переходов между журналами на сайтах издательств (см. рис. 40 — с. 191). Данный факт можно интерпретировать следующим образом: ссылки на социальные и гуманитарные исследования не указываются в работах по дисциплинам, связанным на карте чтения как неприемлемые для «строгих» дисциплин. Однако возможны и другие объяснения, например, что представители остальных дисциплин читают социальные и гуманитарные исследования для общего развития. Однозначно ответить на вопрос о причинах противоречий между картами, отражающими два взгляда на науку, без дополнительных исследований не представляется возможным.

После визуализации связей и взаимного расположения журналов на основании анализа ко-цитирования их публикаций можно перейти к построению классификаций журналов, притязающих на естественность по способу своего построения. Классификация журналов, построенная на основе ко-цитирования публикаций, отражает связи между классами журналов, определяемых на основании общности предмета исследований. Примером подобной классифи-

кации является карта науки, созданная для University of California, San Diego (UCSD) и представленная на рис. 45 (см. с. 196). Методика построения базируется на анализе ко-цитирования публикаций за три периода с 2001 по 2010 г. по данным Web of Science CC и уточнена по данным Scopus. Детали методики построения и история построения карты описаны в [Design and Update..., 2012]. Итоговая классификация содержит 13 категорий и 554 подкатегории. Увеличение дробности классификации и использование классов ко-цитирования в качестве основания для выделения приводит к появлению интересных классов (табл. 16). Названия категорий свидетельствуют о том, что публикации в журналах, которые входят в представленные подкатегории, существенно изолированы от остальных публикаций в категориях. Отметим, что в последующих обновлениях классификации к подкатегориям «российская математика» и «российская химия» добавилась и «российская физика».

Картирование с использованием готовой тематической топоосновы, например пред-

Таблица 16

Примеры категорий классификации UCSD

Номер UCSD	Подкатегория	Категория
393	Mathematical Science (Russia)	Math&Physics
491	Chemistry (Russia)	Chemistry

Источник [DesignandUpdate..., 2012]; данные опубликованы на сайте <http://sci.cns.iu.edu/ucsdmap/>

ставленной на рис. 45 (см. с. 196), позволяет сравнить степень специализации анализируемого массива публикаций со всеми публикациями мира, а фиксированное расположение элементов классификации на диаграмме позволяет быстро читать и срав-

нивать карты, построенные для разных «пулов» публикаций. На рис. 46 (см. с. 197) представлен результат анализа публикаций авторов из СПбГУ, которые уже анализировались на рис. 43 (см. с. 194). Сравнивая два метода визуализации, можно отметить, что на карте, построенной на основании информации только из анализируемых публикаций, не заметны исследования в области мозга, медицины, а также не охвачены социально-гуманитарные области. Таким образом, дополнительным достоинством метода картирования на основании готовой тематической топоосновы является возможность увидеть относительно небольшие группы исследований, которые в случае применения другого метода визуализации теряются на фоне публикаций, относящихся к тематикам с самым большим числом публикаций.

Описанный выше метод, лежащий в основе построения естественной классификации журналов на основании ко-цитирований, можно применить и для классификации отдельных публикаций. Классификация отдельных публикаций позволяет надежно определить категорию, к которой относится публикация, не на основании классификации журнала, а на основании тематической окрестности каждой публикации. Это позволяет преодолеть сложность в классификации статей в мультидисциплинарных журналах. Используя данный подход, можно не ограничиваться уже предложенными категориями, а строить существенно более подробную систему классификации. Однако данный метод как в отношении журналов, так и в отношении статей обладает существенным недостатком. Если мы построили классификацию на массиве данных, например, за 2001–2010 гг., то надежность классификации публикаций, написанных

существенно позднее данного периода, будет сомнительной, ведь за это время может существенно поменяться схема связей между категориями. Однако обновление классификации на основании более свежих данных приводит к тому, что элементы классификации изменяются. Изменения категорий верхнего уровня, может быть, и незначительные, но для узких категорий, используемых для классификации отдельных публикаций, становится невозможным проведение сравнений между классификациями, построенными на массивах публикаций разных лет.

Представленные в разделе техники визуализации информации о науке и технологии путем картирования наукометрической информации позволяют облегчить понимание результатов анализа, однако они не могут подтвердить существование закономерностей, которые не следуют из исходных данных. Доступность современных инструментов визуализаций позволяет с минимальными усилиями создавать большое число диаграмм, которые привлекут внимание к представлению результатов анализа, однако могут содержать дублирование уже представленной информации. В большом числе визуализаций можно скрыть недостатки проведенного анализа и сформировать искаженное понимание представленных результатов.

4.2. Прогнозирование и развитие

Лица, принимающие решения при обсуждении методов картирования науки и технологии, неявно считают, что подобные методы обладают свойствами географических карт, важнейшим из которых была идентификация «белых пятен», позволяющая «планировать» научные и технологические достижения. Тематические «картографиче-

ские» основы принципиально отличаются от географических тем, что мы не можем на момент построения карты знать все потенциально возможные тематики исследований, из которых можно было бы сложить карту для прогнозирования развития науки и техники.

Одно исключение, при котором возможно плановое управление направлением развития с использованием для прогноза методов «наукометрии», представлено в форме схемы системной динамики (рис. 32а). Речь идет о догоняющем развитии науки отдельной страны. Если в мире есть технологический лидер, который обладает достаточным объемом ресурсов для проведения исследований по широкому спектру направлений, то догоняющая страна может сэкономить свои усилия на развитие собственных перспективных научных исследований за счет мониторинга изменения тематик исследований у лидера и соответственно перераспределять имеющиеся научные ресурсы для воспроизведения уже полученных результатов, «подражая» поведению научного лидера. Если у научной тематики возникает возможность прикладного применения, то подобный подход может гарантировать, что в стране найдется специалист, который разбирается в проблеме или может быстро переключиться на нее за счет того, что он занимается тематически близкой проблематикой. У предложенной схемы есть два недостатка. Во-первых, всегда будет наблюдаться отставание от лидера, так как доступные для мониторинга изменения тематик на основе видимых результатов зачастую появляются в момент, когда авторы публикации работают над развитием тематики или уже отказались от нее. Задержка возникает вследствие необходимости донесения проблематики до конкретного ис-

следователя, который может находиться вне контекста исследований. Во-вторых, в целях экономии ресурсов эксперты могут пожертвовать тематиками, развиваемыми у лидера, которые не соответствуют высокоуровневым приоритетам практической деятельности в стране. Эти тематики могут быть востребованы в дальнейшем, но в момент начала работ их необходимость неочевидна.

На схеме (рис. 32а) представлено формирование прогноза развития науки и техники на основании мониторинга результатов уже выполненных исследований. Эксперты в данной схеме играют роль оракулов при мониторинге научных результатов, получаемых в стране-лидере. В предложенной модели возможно внесение небольших корректировок в прогноз развития за счет учета потребностей развития локальных исследований (указано пунктирной стрелкой), однако эта коррекция незначительна. Неявно в схеме предполагается, что эксперты при составлении прогноза развития учитывают потребности развития приоритетных направлений техники, если это необходимо.

Обратим внимание на потенциальную проблему в рассматриваемой схеме, а именно на отрыв экспертов от практической деятельности и полное их сосредоточение на анализе входного документального потока информации для составления плана раз-

в. руетс. я нав.к поиска новых направлений исследований и развиваются иждивенческие настроения в отношении получения финансирования. Использование принципа, заложенного в схеме на рис. 32а, предусматривает установление такого уровня развития науки и техники, при котором развитие переходит на методы, используемые страной-лидером. Обсуждение уровней

развития науки в стране описано в [Tracking scientific..., 2014].

Схема формирования направлений развития науки и техники, реализуемая в странах-лидерах, представлена на рис. 32б. Она визуализирует процедуру, описанную в разделе 1.1. В предложенной схеме эксперты находятся в тесном контакте с исследователями, а часто и сами являются исследователями и находятся в контексте тематик исследований. Данные наукометрии привлекаются как вспомогательный инструмент для калибровки и настройки понимания экспертом цели экспертизы при формировании экспертных заключений. В результате

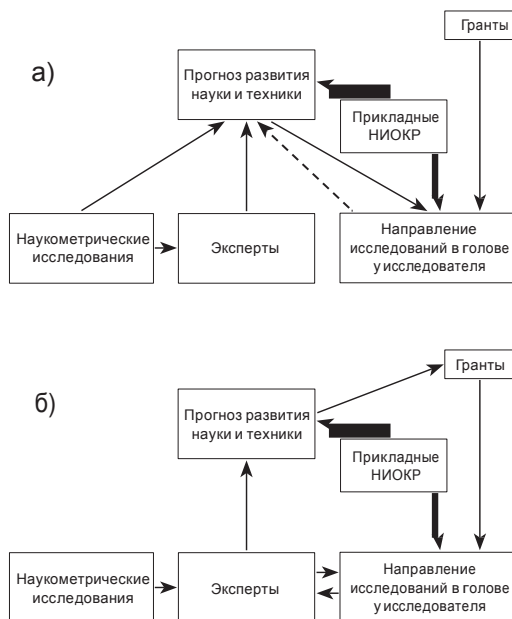


Рис. 32. Варианты управления тематикой научных исследований: а) плановое управление с использованием наукометрии для прогноза; б) схема подстройки направлений исследований, формирующая цикл управления через изменение приоритетов грантового распределения ресурсов

процесса экспертизы формируются направления развития науки и техники, которые используются при изменении приоритетов в выделении объемов ресурсов для научных исследований по тематике (на схеме представлены в форме грантов).

Если возможности применения схемы рис. 32а в развитии науки ограничены, нельзя ли применить ее для развития технологии? Если правообладатель не заинтересован в получении локального патента, то это хороший шанс получить патент в своей юрисдикции и наладить выпуск продукции по имеющимся описаниям. Однако не во всех технологических областях все продукты защищаются патентами. В исследовании инновационных продуктов, получивших в период 1977–2004 гг. ежегодную премию R&D 100 Awards журнала Research and Development, было обнаружено, что на 90,9 % продуктов не были получены патенты [Reassessing patent..., 2013]. Вопросы полноты отражения знания технологий в патентах требуют дополнительных исследований, так же как и вопрос о выявлении связи между научным и технологическим знанием (см. обсуждение в разделе 1.2).

Если схема рис. 32а не работает в рамках страны, может ли она сработать в рамках отдельной научной организации? Да, может, но только на начальном этапе развития. Например, созданный в 2009 г. Университет науки и технологии имени короля Абдулы (King Abdullah University of Science and Technology, KAUST), определив тематический план своего развития и привлекая высококвалифицированных ученых со всего мира, смог увеличить объем и качество своих научных результатов до уровня, позволившего ему войти в топ-500 Шанхайского рейтинга. Сможет ли университет продолжить свое развитие, если прекратит-

ся подпитка высокорейтинговыми учеными, покажет время.

При принятии решения о развитии новых научных направлений необходимо учитывать доступность ресурсов человеческого капитала и специфику региона, понимание которой не может быть достигнуто средствами наукометрии. Такие данные можно получить исключительно через экспертов и уже работающих в организации ученых, знакомых с местной спецификой и местной научной составляющей.

При определении направлений развития нужно выбирать тематики, которые позволят сформировать жизнеспособные научные группы. Критерии для выбора тематики, которая позволяет сформировать жизнеспособную научную группу, основываются не столько на выявленных методами наукометрии перспективных направлениях, сколько на наличии возможности создать группу ученых, которая сможет привлечь к обсуждению и использованию своих результатов коллег и практиков вне организации (см. обсуждение в разделе 1.2).

Работа научной группы по новому направлению начинается с чтения публикаций по теме исследований. Чем лучше группа исследователей знакома с уже вышедшими публикациями, тем лучше понимание членами группы контекста исследований, проводимых в мире, и тем выше шансы предложить для обсуждения новые результаты, которые отвечают на актуальные вопросы. В качестве средства ранней диагностики развития научного направления можно предложить отслеживание динамики чтения. Связь между чтением и последующими публикациями статистически значима и подтверждалась в большом числе исследований университетов и научных организаций, ориентированных на создание ново-

го знания. Динамика увеличения числа обращений к полнотекстовым ресурсам может задолго до появления первых публикаций сигнализировать о том, что процесс развития нового направления запущен.

Для членов новой научной группы также требуются и навыки представления научных результатов (см. обсуждение в разделе 1.5). Привлечение внимания к результатам своих исследований и, как следствие, повышение уровня качества научных результатов подразумевает систематическое чтение и обсуждение актуальных научных статей, поездки на семинары/конференции для общения с коллегами и получение обратной связи о потребностях от местной промышленности и общества.

Навыки систематического чтения научной литературы, обсуждения и представления новых научных результатов не исчерпывают всех навыков, необходимых исследователям в научных группах. Список универсальных навыков исследователей, на которые нужно обращать внимание при обучении новых научных сотрудников или привлечении существующих, был определен в рамках исследования мобильности научных сотрудников на основе результатов анализа факторов, способствующих присвоению степени PhD в Европейском союзе, см. табл. 17 [Transferable Skills..., 2012, p. 20].

Предложенный выше процесс формирования жизнеспособных научных групп для развития новых научных направлений в организации требует учета потребности промышленности (практики). Хорошим индикатором востребованности уже выполненных исследований в практике является регулярное выполнение научной группой прикладных работ в интересах внешнего заказчика. Однако для новых научных групп этот индикатор не является релевантным, так как

группой еще не получены результаты, которые могут привлечь практиков. Исключение составляют группы, которые создаются изначально для решения практических задач, а в дальнейшем переходят к развитию научной тематики. Для выявления и учета долгосрочных потребностей практики в научных исследованиях можно обратиться к концепции тройной спирали Г. Ицковица [Ицковиц, 2010].

В основе концепции лежит модель взаимодействия государства, университета и бизнеса, названная автором тройной спиралью. Автор концепции на многочисленных примерах создания инновационных центров в США, Швеции и Бразилии демонстрирует, что в разных странах и географических регионах одного государства наблюдаются различные формы реализации модели тройной спирали. Но общим моментом во всех рассмотренных примерах является формирование условий непрерывного создания в регионе новых фирм, реализующих идеи, разработанные в стенах университетов. Ключевым условием появления новых фирм автор считает создание в регионе университета нового типа — предпринимательского. Прослеживая исторические причины успешности двух эталонных предпринимательских вузов, Массачусетского технологического института и Стэнфордского университета, можно сделать вывод, что видимые и активно воспроизводимые в других университетах структуры, такие как центры трансфера технологий, инкубаторы и технопарки, являются вторичными по отношению к концепции предпринимательских университетов. Университет нового типа не просто создает региональный пул экспертов — преподавателей университета, но инициирует передачу своих разработок в региональную промышленность, способ-

Перечень универсальных навыков для исследователей

Источник: [Transferable Skills..., 2012, p. 20]

Категория навыков	Перечень универсальных навыков
Когнитивные способности	• Креативность (творческий подход) и абстрактное мышление. • Решение задач.
Навыки коммуникаций	• Письменные и устные навыки общения и подачи информации. • Навыки общения и ведения диалога с неспециалистами (связи с общественностью). • Навыки преподавания / группового обучения. • Научный подход к формированию политики организации и законотворчеству.
Предпринимательские навыки	• Бизнес-навыки (основы предпринимательской деятельности, инновации, коммерциализация новых товаров, интеллектуальная собственность и патентное дело).
Межличностные навыки	• Групповая деятельность / работа в команде. • Навыки репетиторства / индивидуального обучения. • Ведение переговоров. • Налаживание связей (установление контактов) и неформальное общение.
Организационные навыки	• Навыки ведения проектов и планирование рабочего времени. • Планирование карьеры.
Исследовательские компетенции	• Навыки планирования и управления исследовательской работой. • Знакомство с методами и технологиями из областей научного знания, лежащих вне тематики квалификационной работы. • Научная этика и принципы научной честности. • Написание заявок на получение грантов.

ствуя созданию новых фирм из числа своих ученых и выпускников, если в регионе нет фирм, готовых реализовать идеи. Автором концепции университета как инициатора передачи технологий в промышленность был Уильям Бартон Роджер, основатель и первый президент Массачусетского технологического института, изложивший свои идеи в докладе 1846 г. Спустя сто лет данная концепция была реализована в Стэнфордском университете благодаря Фредерику Терману, который, работая во время войны в Массачусетском технологическом инсти-

туте, мог наблюдать эффективность концепции в военных разработках.

Государство в предложенной модели выступает не просто как гарант правил игры, но и как источник инвестиций с высоким риском и горизонтом получения результатов десять и более лет. По мнению Г. Ицковица, гарантией того, что инвестиции со столь долгим сроком появления первых результатов будут эффективны, является то, что каждый элемент в тройной спирали обеспечивает процесс уточнения целей исследований с учетом потребностей практики, ко-

которые могут возникнуть в будущем. Модель тройной спирали демонстрирует, что идеи исследований в университетах возникают в результате общения представителей региональной промышленности и вузовской науки в рамках ассоциаций, поддерживаемых региональными властями, но финансирование идеи получают только в том случае, если проходят экспертизу на предмет возможности практического применения на горизонте десятилетий.

Заключение

Методы визуального представления наукометрической информации позволяют сформировать начальные гипотезы о закономерностях при анализе наукометрической информации и представлении конечных результатов анализа. Обнаруженные на диаграммах закономерности обязательно требуют подтверждения в рамках дальней-

ших исследований с опорой на методы количественного анализа, представленные в главах 2 и 3 данной книги. Показана связь между приемами компактного представления результатов наукометрического анализа с логикой развития и совершенствования научного процесса. Проведенный в главе анализ свидетельствует, что при принятии решений о векторе развития науки на макроуровне нельзя опираться исключительно на методы наукометрии, привлечение экспертного мнения принципиально важно для получения адекватного результата. Развитие новых научных направлений на уровне организаций зависит в большей степени от наличия и уровня развития человеческого капитала и в меньшей степени от определения перспективных направлений развития науки и техники.

Желязны Дж. Говори на языке диаграмм. Пособие по визуальным коммуникациям. М.: Манн, Иванов и Фербер, 2009. 320 с.

Ицкович Г. Тройная спираль: университеты — предприятия — государство: инновации в действии. Томск: Изд-во Томского гос. ун-та систем управления, 2010. 237 с.

Латур Б. Наука в действии: Прагматический поворот. СПб.: Изд-во Европейского университета в Санкт-Петербурге, 2013. 416 с. [Оригинал: Latour B. Science in Action: How to Follow Scientists and Engineers Through Society. Cambridge, Mass.: Harvard University Press, 1988. 288 p.]

Яу Н. Искусство визуализации в бизнесе. Как представить сложную информацию простыми образами. М.: Манн, Иванов и Фербер, 2013. 352 с.

An approach to discovering new technology opportunities: Keyword-based patent map approach / S. Lee, B. Yoon, Y. Park // *Technovation*. 2009. Vol. 29. № 6–7. P. 481–497.

Börner, K., Polley, D.E. Visual Insights: A Practical Guide to Making Sense of Data. Cambridge, Massachusetts: The MIT Press, 2014. 312 p.

CitNetExplorer - Analyzing citation patterns in scientific literature. [2014]. URL: <http://www.citnetexplorer.nl/Home> (дата обращения: 22.07.2014).

Clickstream Data Yields High-Resolution Maps of Science / J. Bollen et al. // *PLoS ONE*. 2009. Vol. 4. № 3. P. e4803.

Collins R. The Sociology of Philosophies: A Global Theory of Intellectual Change. Belknap Press of Harvard University Press, 2000. 1098 p. [Русский перевод: Коллинз Р. Социология философий: глобальная теория интеллектуального изменения. Новосибирск: Сибирский хронограф, 2002. 1280 с.]

Design and Update of a Classification System: The UCSD Map of Science / K. Börner et al. // *PLoS ONE*. 2012. Vol. 7. № 7. P. e39464. [Материалы классификации опубликованы по адресу <http://sci.cns.iu.edu/ucsdmap/>]

Eck N. J. van, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping // *Scientometrics*. 2010. Vol. 84. № 2. P. 523–538.

Eck N. J. van, Waltman L. CitNetExplorer: A new software tool for analyzing and visualizing citation networks // *Journal of Informetrics*. 2014. Vol. 8. № 4. P. 802–823.

HistCite / ThomsonReuters. [2012]. URL: <http://interest.science.thomsonreuters.com/forms/HistCite/> (дата обращения: 22.07.2014).

IVMOOC: Information Visualization [Открытый онлайн курс]. [2014]. URL: <http://ivmooc.cns.iu.edu/> (дата обращения: 21.06.2014).

Lotka A. J. The Frequency Distribution of Scientific Productivity // *Journal of the Washington Academy of Sciences*. 1926. № 16. P. 317–323.

Mapping research specialties / S. A. Morris, B. Van der Veer Martens // *Annual Review of Information Science and Technology*. 2008. Vol. 42. № 1. P. 213–295.

Pajek / Vladimir Batagelj, Andrej Mrvar. 1996.

Participatory visualization with wordle / F. B. Viegas, M. Wattenberg, J. Feinberg // *Visualization and Computer Graphics, IEEE Transactions on*. 2009. Vol. 15. № 6. P. 1137–1144.

Reassessing patent propensity: Evidence from a dataset of R&D awards, 1977–2004 / R. Fontana, A. Nuvolari, H. Shimizu, A. Vezzulli // *Research Policy*. Ser. Eco-

nomics, innovation and history: Perspectives in honour of Nick von Tunzelmann. 2013. Vol. 42. № 10. P. 1780–1792.

Sci² Tool : A Tool for Science of Science Research and Practice. [2010]. URL: <https://sci2.cns.iu.edu/user/index.php> (дата обращения: 23.07.2014).

Science mapping software tools: Review, analysis, and cooperative study among tools / M. J. Cobo, A. G. Lopez-Herrera, E. Herrera-Viedma, F. Herrera // *Journal of the American Society for Information Science and Technology*. 2011. Vol. 62. № 7. P. 1382–1402.

Tracking scientific development and collaborations – The case of 25 Asian countries / Henk Moed, Gali Halevi // *Research Trends*. 2014. № 38. P. 25–30.

Transferable Skills Training for Researchers / OECD. Paris: Organisation for Economic Co-operation and Development, 2012. 151 p.

VOSviewer // Centre for Science and Technology Studies, Leiden University, 2013. URL: <http://www.vosviewer.com>. (дата обращения: 25.06.2014).

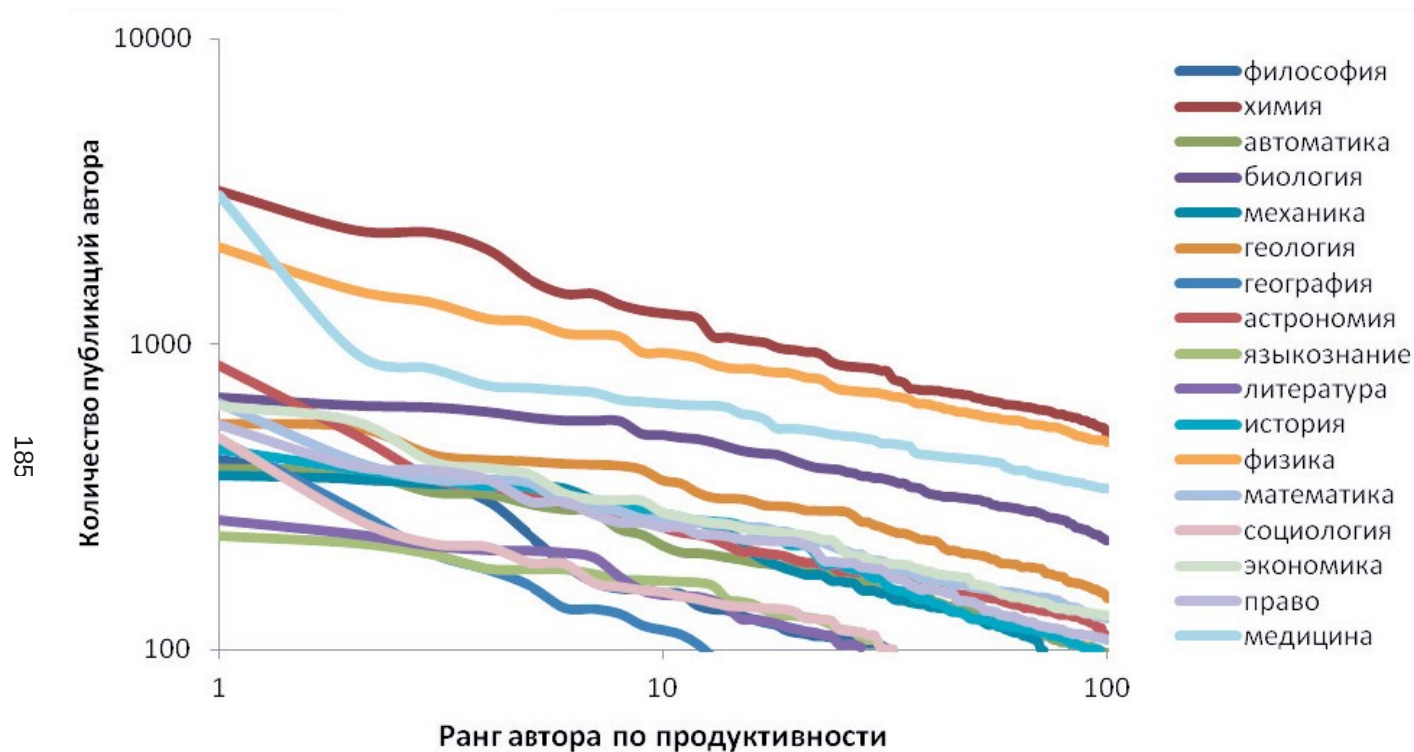


Рис. 33. Распределение авторов по количеству публикаций по научным областям по данным РИНЦ.
Пояснение см. на с. 124

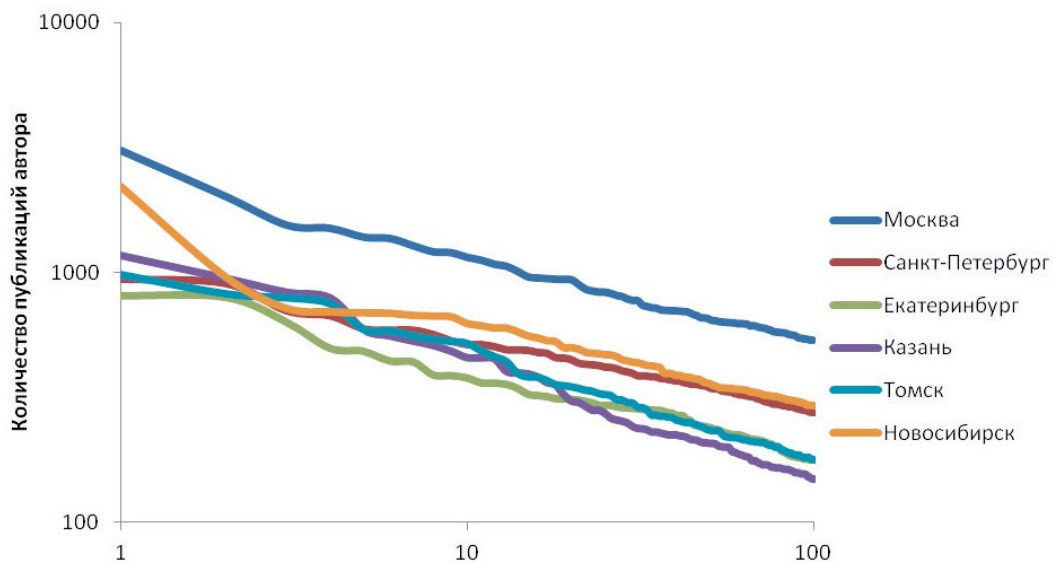


Рис. 34. Распределение авторов по количеству публикаций по городам по данным РИНЦ.
Пояснение см. на с. 124

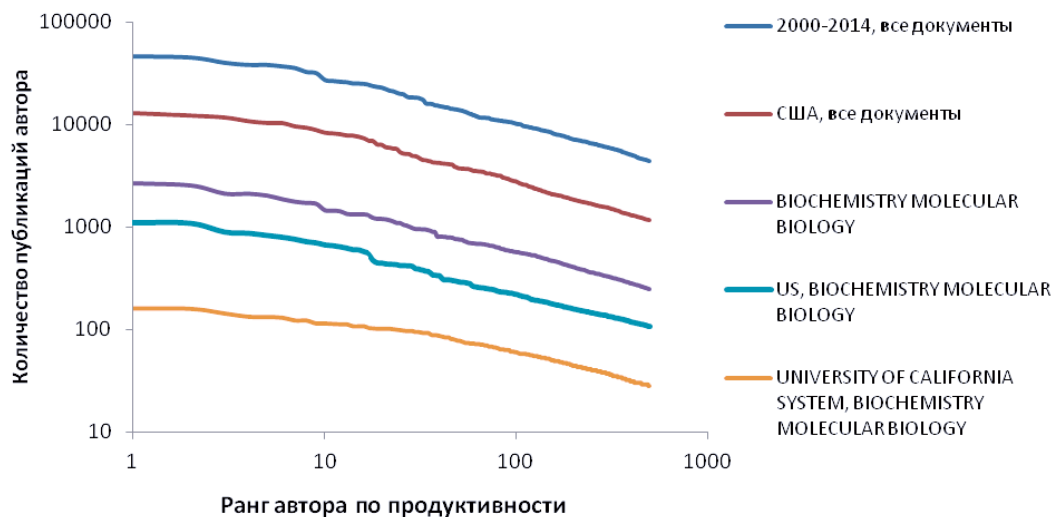
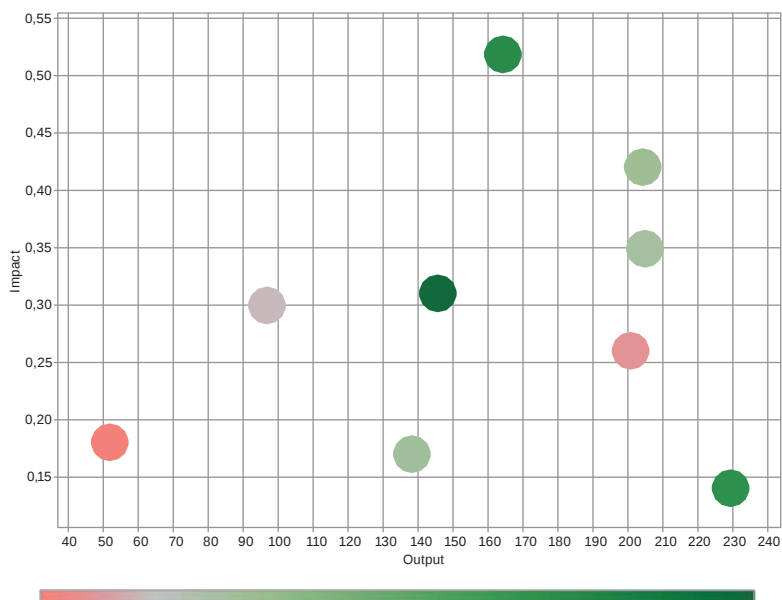


Рис. 35. Распределение авторов по количеству публикаций по данным Web of Science CC.
Пояснение см. на с. 124



a)

Предметная область	Количество публикаций	Нормализованная цитируемость	Динамика роста за десять лет
Subject_1	205	0,35	13,50%
Subject_2	201	0,26	-12,10%
Subject_3	204	0,42	17,00%
Subject_4	229	0,14	72,60%
Subject_5	97	0,3	-2,00%
Subject_6	138	0,17	15,60%
Subject_7	146	0,31	110,60%
Subject_8	10	0,83	800,00%
Subject_9	52	0,18	-20,70%
Subject_10	164	0,52	78,00%

б)

Рис. 36. Представление сравнительных количественных и качественных характеристик публикаций по предметным областям (см. пояснение на с. 165–167):

а) графическое представление предметных областей на пузырьковой диаграмме, где по горизонтали представлено количество публикаций (Output), а по вертикали – нормализованная цитируемость публикаций (Impact, см. раздел 2.3); цветом закодирована динамика роста за десять лет; предметная область Subject_8 исключена из диаграммы по причине малого числа публикаций по сравнению с остальными областями;

б) исходные данные, показанные на предыдущей диаграмме; цветом закодирован диапазон значений по столбцам

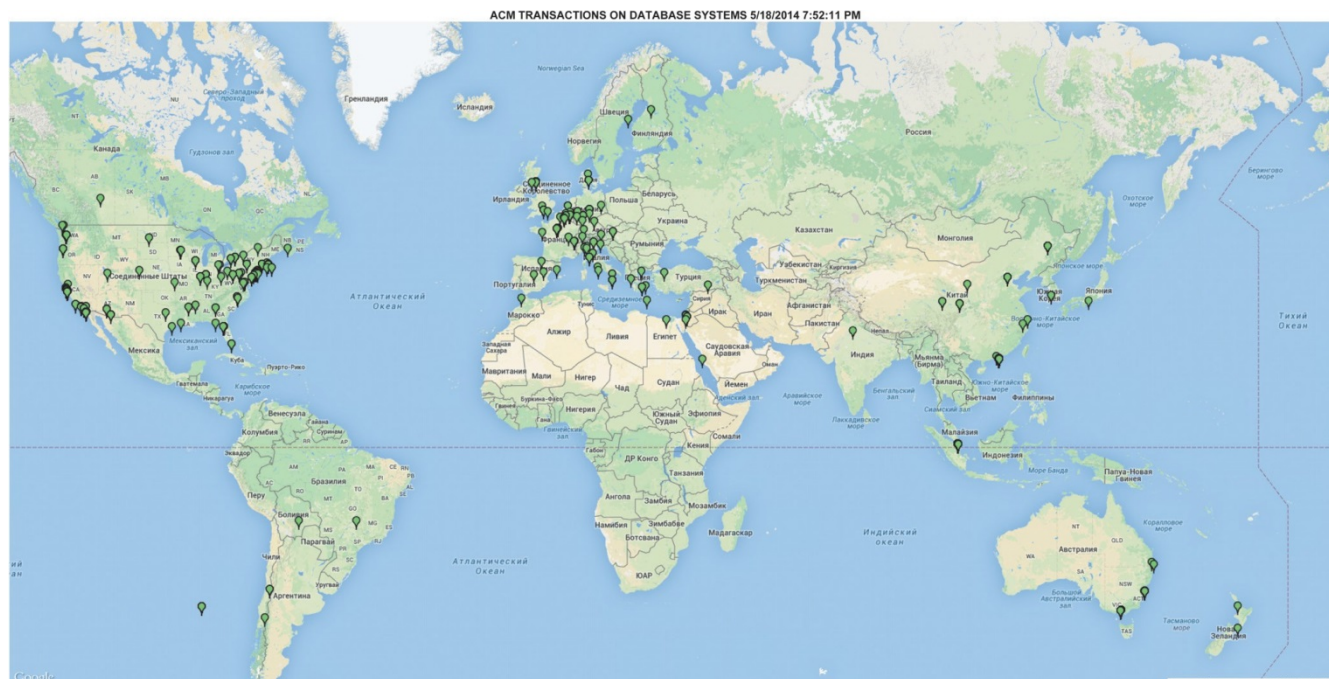
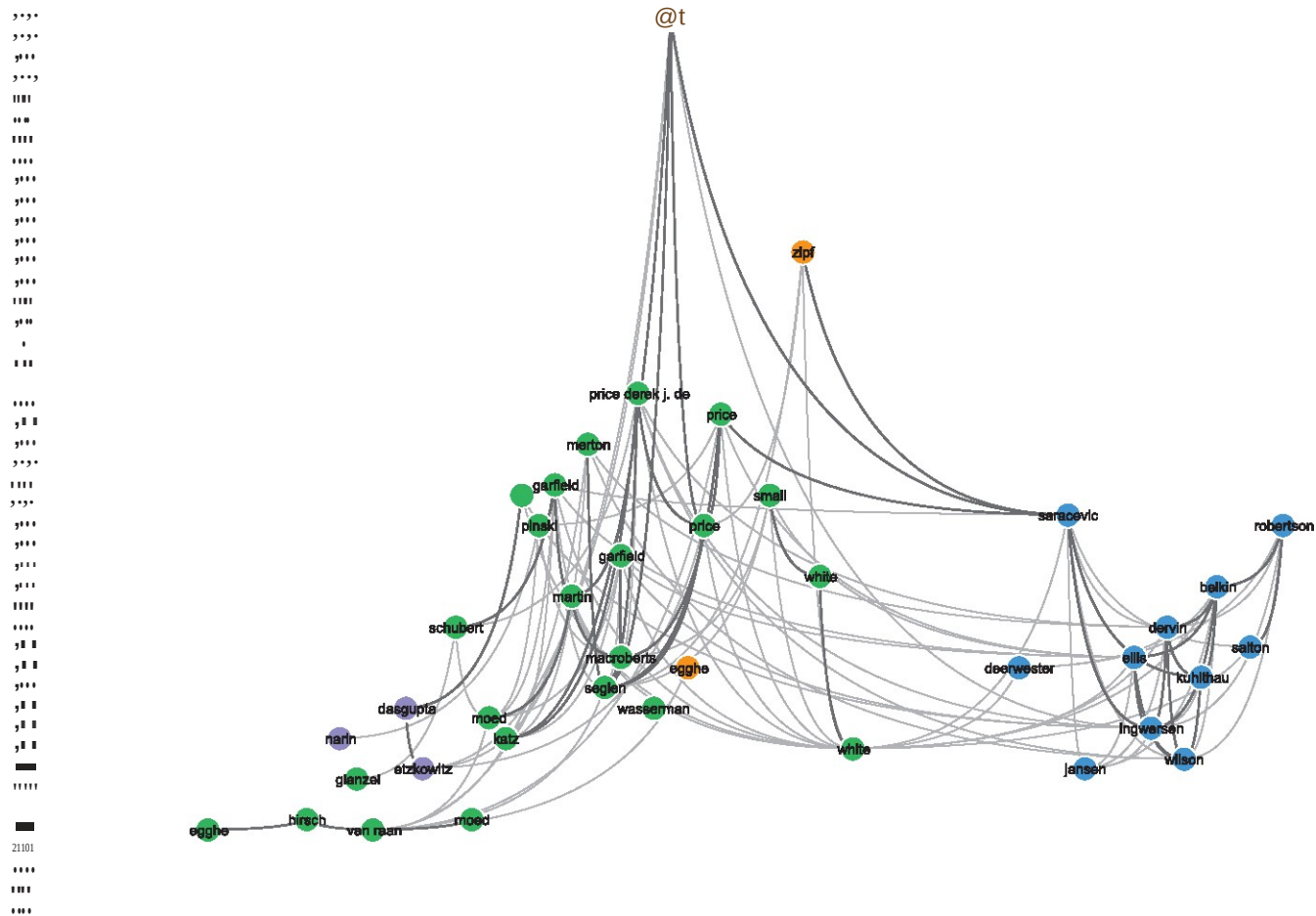


Рис. 37. Географическое распределение авторов ACM Transactions on Database Systems (TODS) 2008–2014 гг.
Пояснение см. на с. 168—169



P1-1c. 38. (8531-1 Me>K,QY KnaCTepaMI-1 ny6nf1KalJf1, npilMO f1nf1 KOC8eHHO Lj1Tf1pyf0Uj1X pa6ury [lotka, 1926]8 MaCCf18e ny6nf1KalJf13a ne-
 pl-10,11 1945–2014 . npe,QCTa8neHO 4284 ny6nf1KalJf1f1 f128482 CCb1nKI-1. B Ka YeCT8e MeTOK KnaCTepo8 f1CnOnb3yeTCil cj:JaMf1nf1il nep8010
 a8TOpa. VICX0,QHbIMaCCI-18 ny6nf1KaLlf1f1Onf1CaH8 TeKCTe.
 noilCheHf1e CM. HaC. 170

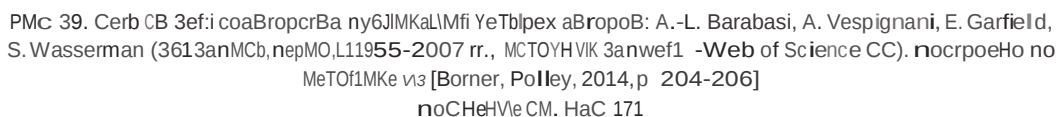




Рис. 40. Карта науки построена на основании истории переходов между страницами публикаций, размещенными на сайтах издательств [Clickstream Data..., 2009].

Карта опубликована под лицензией Creative Commons.

Пояснение см. на с. 172

1- Z 1.1.1 O

[illegible]

10 раз (20% от 8Сех VK333HHbiX 8 ny6JIV1K3LVI>1X). npoaH3JIV13V1p083HO 4737 ny6JIVIK3LVIi13a 2004–2014 rr., 1, [e xOT>I 6bl 0, [MH 3BTOp
YK333JI8 acjxpV1flV13L\ VVI Ka3aXCT3H. VICTO'-IHVIK 3anMcei1: Web of Science CC. no>ICHeHVle CM. Ha c 173

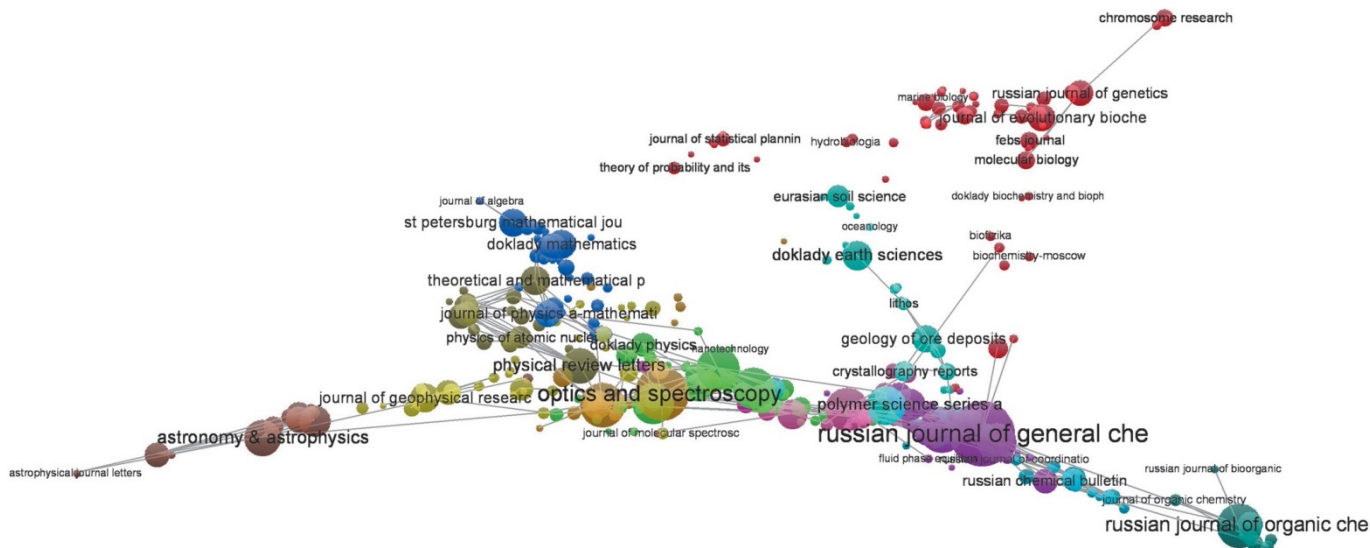


Рис. 43. Визуализация методом «библиографического сочетания» по журналам массива публикаций авторов СПбГУ (10785 записей, период 2004–2013 гг., источник записей – Web of Science CC, построено с использованием VOSviewer). Цветом закодированы тематические категории; площадь круга отражает число публикаций в источнике; расположение и линии связи указывают на силу связи публикаций, цитирующих работы, опубликованные в одном и том же журнале.

Пояснение см. на с. 175–176

Рис. 44. Карта науки, построенная методом ко-цитирования для 5000 базовых научных журналов (публикации за период 2007 г., анализ цитирований за 1997–2006 гг., источник записей – Web of Science CC, построено с использованием VOSviewer). Цветом закодированы тематические категории; площадь круга отражает число публикаций в источнике; расположение указывает на силу связи цитирования одинакового набора публикаций; описание метода построения представлено в [Eck, Waltman, 2010]. Обработанные данные для построения карты включены в состав поставки VOSviewer.

Пояснение см. на с. 176

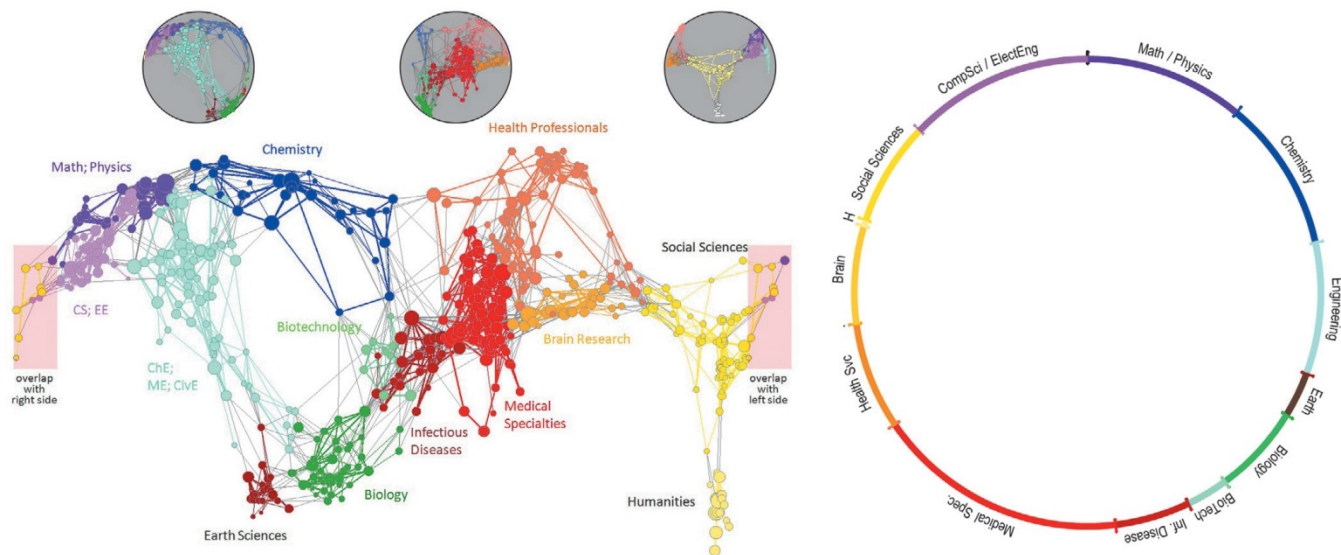


Рис. 45. Визуализация карты науки UCSD (2001–2010 гг.). Слева – отображение карты науки в двумерной проекции Меркатора (левая сторона карты имеет продолжение справа), сверху – три врезки с отображением на 3D-сфере, справа – отображение на одномерное колесо. Источник: [Design and Update..., 2012]. Карта опубликована под лицензией Creative Commons, права на классификацию и карту принадлежат University of California, San Diego.

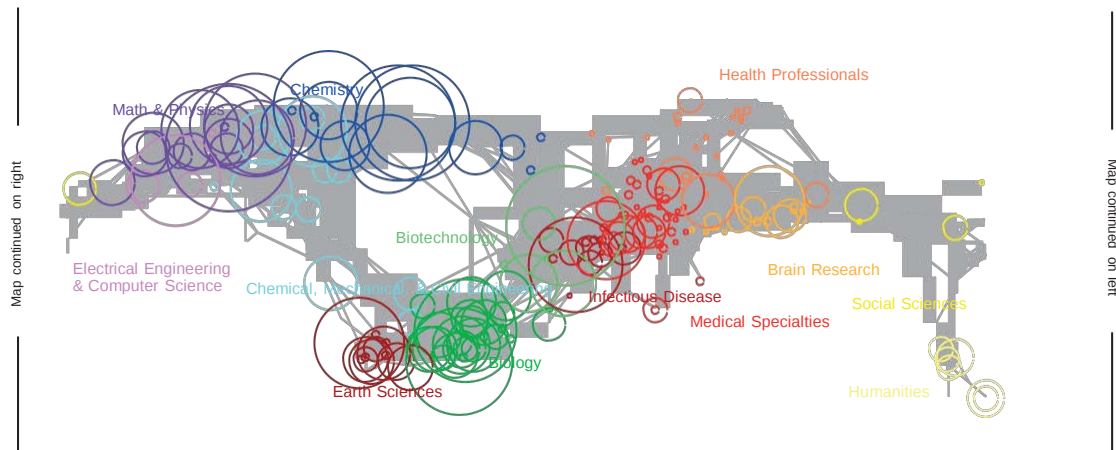
Пояснение см. на с. 177

Topical Visualization

Generated from 10785 Unique ISI Records

154 out of 1 595 publications were mapped to 202 subdisciplines and 13 disciplines.

19, 2014 | 10:46 PM YEKT

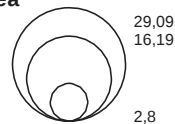


2008 The Regents of the University of California and SciTech Strategies.
Map updated by SciTech Strategies, OST, and CNS in 2011.

Legend

Circle area: Fractional Journal Count
Unclassified = 1441
Minimum = 0
Maximum = 36
Color: Discipline
See end of PDF for color legend.

Area

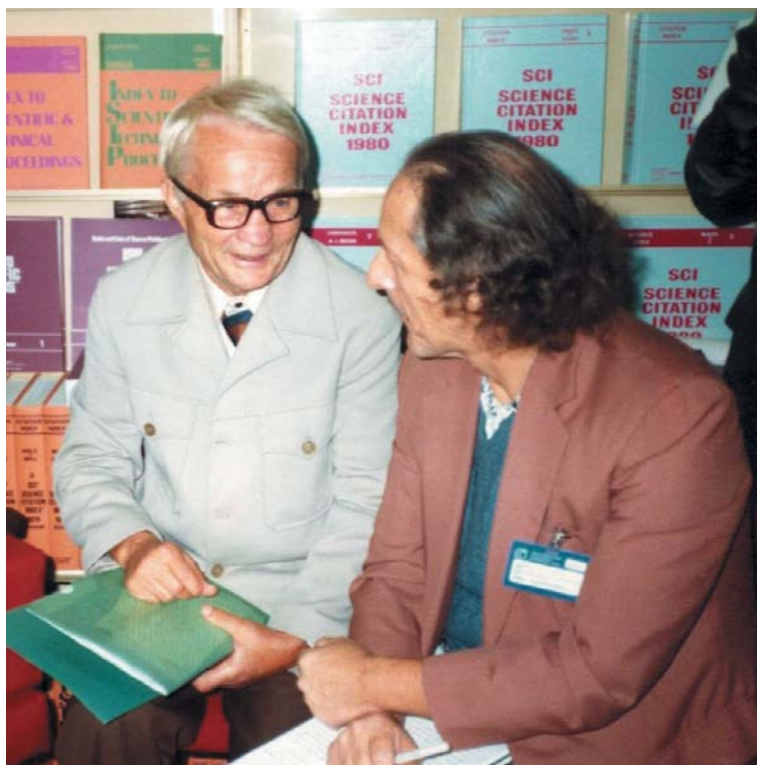


How To Read This Map

The *UCSD map of science* depicts a network of 554 subdiscipline nodes that are aggregated to 13 main disciplines of science. Each discipline has a distinct color and is labeled. Overlaid are circles, each representing all records per unique subdiscipline. Circle area is proportional to the number of fractionally assigned records. Minimum and maximum data values are given in the legend.

CNS (cns.iu.edu)

Рис. 46. Визуализация массива публикаций авторов СПбГУ на карте классификации UCSD (10785 записей, период – 2004–2013 гг., источник записей – Web of Science CC, построено с использованием Sci² Tool). Цветом закодированы категории; площадь круга отражает число публикаций в журналах подкатегории; расположение соответствует карте подкатегорий журналов, представленной на рис. 45. Описание метода построения представлено в [Börner, Polley, 2014, p. 140–141]. Пояснение см. на с. 177

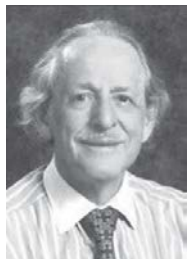


Ю. Гарфилд и В. В. Налимов,
сентябрь 1981 г., Международная выставка-ярмарка в Москве



Ю. Гарфилд и В. А. Маркусова,
2002 г., Москва

ИХ ОШИБКИ, НАШИ ОШИБКИ И ВАШИ ОШИБКИ



Ю. Гарфилд

Впервые опубликована в Garfield. E. Errors Theirs, Ours and Yours // Current Contents. 1974. № 25. June 19. P. 5–6.

Перевод выполнен по изданию Essays of an Information Scientist. Vol. 2. 1974–76. P. 80–81.

Переводчик Валентина Александровна Маркусова.

Перевод на русский язык выполнен с любезного разрешения автора.

Недавно автор написал нам о статье, которую он опубликовал. В содержании журнала в заглавии его статьи было пропущено слово “touch” из сочетания “light touch”. В результате этой ошибки он получил огромное количество запросов на репринты от людей, интересующихся исследованиями, связанными со зрением. Надеюсь, что мой призыв к читателям обратить внимание на этот случай поможет автору получить запросы от исследователей, которые заинтересованы в его тематике. Однако мы с радостью смогли отметить, что эта досадная ошибка в содержании журнала не попала в «Еженедельный предметный указатель» (Weekly Subject Index) и в Science Citation Index (SCI).

Понятно, что ISI (Institute for Scientific Information — Институт научной информации) будет совершать собственные ошибки. Порог допустимости этих ошибок является экономическим вопросом. Я всегда настаиваю на том, что мы должны охватывать не менее 100 % каждого выпуска научного

журнала, который мы обрабатываем. Пропуск даже одного выпуска непростителен. Возможно, проще вообще не обрабатывать журнал. С другой стороны, если в среднем в журнале публикуется 300 статей ежегодно, то допустимо ли упущение в 0,33 %? Может быть, некоторые читатели готовы позволить нам пропустить одну статью из 300, но наша цель — иметь порог ошибок ниже этого. Я знаю, что нет легкого способа определения уровня наших ошибок. Даже наших обращений к редакторам 5000 научных журналов, которые мы обрабатываем, недостаточно. Значительная часть этой проблемы — точное определение вида статей, которые мы индексируем. У журналов, которые содержат только исследовательские статьи в стандартном формате, нет проблем. Но как много вы можете назвать журналов, которые никогда не публикуют заметки редактора, рецензии на книги, заметки о конференциях, объявления, некрологи, письма к редактору и другие более эфемерные

материалы? «Письмо к редактору» может означать практически все, что угодно. «Письмо к редактору» в журнал Nature или в журнал Science отличается от письма к редактору журнала по растениеводству или научно-популярного журнала.

Ошибки в написании или пропуски в адресе автора являются достаточно серьезными ошибками. Мы обрабатываем более миллиона адресов ежегодно. До сих пор тысячи журналов отказываются давать полный адрес. (Не обращались ли вы случайно к редактору какого-нибудь журнала с этой проблемой?) В то время как мы пытаемся предоставить полный адрес, включая почтовый индекс, когда это возможно, есть много случаев, когда это невозможно сделать. В некоторых университетах подразделения разбросаны по городу или штату, и у каждого из них может быть свой почтовый индекс.

Существует общее мнение, что не следует открыто говорить о недостатках, неадекватности, неудачах и т. п. Большинство публикаций, а также организаций следуют совету Наполеона «стирать грязное белье у себя дома». Однако я полагаю, что ISI имеет особые отношения с опытными читателями Current Contents (далее — CC. — Прим. ред.). Мы осознаем свои достижения и признаем свои недостатки. Мы не можем удовлетворить всех. Поэтому для нас важно, чтобы вы знали, в чем мы можем ошибиться и насколько обоснована ваша уверенность в CC и ISI, если речь идет о ваших индивидуальных требованиях. Но в то же время мы рассчитываем на то, что CC является для вас частью большой информационной системы, к которой большинство из вас имеет доступ.

Экономический прессинг, а также фрагментация научных исследований по мульт-

тидисциплинарным направлениям или кластерам все чаще приводят нас к необходимости исключать некоторые журналы из одного или нескольких изданий CC. Каждый раз, когда мы принимаем такое решение, некоторые читатели чувствуют себя обиженными или уязвленными. Эти читатели полагают, что это «серьезное» упущение.

Универсальной удовлетворительной альтернативой этому могла бы быть обработка всех журналов в одном громадном выпуске CC. Его пришлось бы выпускать ежедневно. Я хотел проверить эту идею несколько лет назад и предложил Национальному научному фонду (National Science Foundation) опубликовать научную ежедневную газету — Daily Newspaper of Science. Но об этом в другой раз.

Есть и другие ошибки, которые мы не можем контролировать без чрезмерных затрат. Предположим, слово в заглавии статьи напечатано с ошибкой в содержании журнала, но не внутри журнала. Мы производим довольно значительное «редактирование» содержания, чтобы исключить пробелы или другие ненужные характеристики. Однако мы не сравниваем по буквам содержание с заглавиями каждой статьи внутри журнала. Каждое слово заглавия, фамилия автора, адрес организации и т. д. отдельно вводятся в банк данных ISI. Для этого используется непосредственно текст заглавия и содержание статьи. Таким образом, даже если есть ошибка в оглавлении журнала, правильное написание слова все равно будет введено в банк данных и наш еженедельный предметный индекс (Weekly Subject Index).

Адреса многих организаций являются технически правильными, но они не унифицированы. Когда мы недавно закончили ежегодное издание ISI Who is Publishing in

Science — WIPIS (Кто публикуется в науке), то после сортировки компьютером нам пришлось скорректировать 75 000 названий! Иначе в банк были бы включены дубликаты: Prof. R. E. Jones at Columbia University, Physics Dep., 118 St. & Broadway, New York City, 10 027 и Prof. R. E. Jones, Columbia University, Dep. of Physics, 118 St. & Broadway, New York, N. Y., 10027. Проблема еще более усложняется, если два человека с одной и той же фамилией и инициалами работают в одном и том же университете. Как много людей, которых зовут Т. YamaKawa, работают в Университете Токио? Если вы найдете серьезную ошибку в написании вашего адреса в СС или WIPIS, пожалуйста, сообщите нам об этом. Мы опубликуем правильный адрес, который потом будет включен в наш кумулятивный индекс.

Часто грамматические ошибки и ошибки в написании встречаются в переводе на английский содержания, опубликованных в неанглоязычных журналах. Я уверен, что если бы нам в ISI пришлось переводить с английского языка на русский, японский, немецкий, французский и др. языки, то мы совершали бы подобные ошибки. Тем не менее имидж ISI не улучшается в глазах читателей, когда они находят подобные ошибки и винят нас в этом. Например, советский журнал «Доклады»¹ обрабатывается для подготовки СС в его оригинальной версии, и мы воспроизводим содержание на английском

языке, которое публикуется в этом журнале. Должно быть, это нелегкая задача для персонала Академии наук СССР переводить на английский язык статьи по стольким научным дисциплинам. Мы можем только восхищаться такой многогранностью.

Однако другие советские журналы не заслуживают такой похвалы, но их квалификация по переводу выше, чем многих других неанглоязычных журналов, публикуемых на Западе. Мы можем только надеяться, что они консультируются с авторами статей в использовании терминологии. Нам трудно критиковать качество перевода заголовков статей, когда существует так много журналов, которые решительно отказываются переводить на английский язык содержание журнала. Поскольку мы стремимся улучшить качество нашей работы, мы обратились к атташе по науке ряда стран и попросили помощи в решении этой проблемы. Мы надеемся, что они также смогут помочь в опубликовании полных адресов авторов из этих стран, включая почтовые индексы.

И наконец, я обращаюсь за помощью к нашим читателям, поскольку многие из них являются авторами. Сделайте так, чтобы редакторы ваших любимых журналов обеспечивали полные и точные данные, чтобы ваши коллеги из других областей знания смогли вас найти, когда ваша статья появится в СС.

¹В то время журнал «Доклады АН СССР» не разбивался на серии, как это осуществляется сейчас компанией ВИНИТИ-НАУКА.

РУССКИЕ ИДУТ!



Ю. Гарфилд

Часть 1. Сто самых влиятельных советских ученых 1973–1988. Впервые опубликована в: Garfield E. The Russians are Coming. Part 1. The Red-Hot 100 Soviet Scientists, 1973–1988 // Current Contents. 1990. № 24, June 11. P. 202–215.

Часть 2. Пятьдесят самых цитируемых советских статей за 1973–1988 гг. в Science Citation Index и взгляд на исследовательские фронты 1988 г. Впервые опубликована в: Garfield E. The Russians are Coming. Part 2. The Top 50 Soviet Papers Most Cited in the 1973–1988 Science Citation Index and a Look at 1988 Research Fronts // Current Contents. 1990. № 25, June 18. P. 216–226.

Переводчик Валентина Александровна Маркусова

Сокращенный перевод. Перевод на русский язык выполнен с любезного разрешения автора.

Часть 1

Советская наука и гласность Горбачева

Во многих смыслах 1990 г. — это решающий год для Советского Союза. Изменения, которые происходят в СССР, невозможно было предугадать пять лет назад, и все, что происходит в начале 1990 г., говорит о том, что этот год тоже будет полон событиями. Кто мог подумать, что многопартийная система, создание западного типа президентства и правительства будут предложены Генеральным секретарем советской коммунистической партии?

Гласность и перестройка вызывают надежду не только на улучшение политических и экономических условий, но и большую открытость и свободу советской науки. Действительно, в последние 24 месяца новые научные связи были установлены между СССР и Западом, в особенности между СССР и ее соперником в холодной

войне — США. Беспрецедентным событием было участие советской делегации, возглавляемой вице-президентом АН СССР академиком К. Фроловым, в ежегодном собрании Американской ассоциации по достижениям в науке (American Association for the Advancement of Science — AAAS), прошедшей в феврале 1988 г. в Бостоне. Это событие открыло новую главу в научном диалоге между СССР и США. С тех пор различные приглашения по научному обмену были подписаны между различными научными делегациями СССР и США.

В 1989 г. рост научных связей между СССР и США привел к пактам по научному сотрудничеству, в том числе между рядом университетов, например Brookhaven National Laboratory (Upton, NY) и Институтом физики высоких энергий (Протвино), в которых в середине 1990-х гг. 15 американских и 15 советских физиков

Moscow

будут выполнять совместный эксперимент на советском ускорителе. В августе Геологическая служба США (US Geological Survey) и АН СССР подписали контракт о работе по совместному предупреждению землетрясений, включая будущую установку станций совместного мониторинга в обеих странах. Национальная академия наук США (NAS) и АН СССР собираются создать общий комитет по экологии.

В течение прошедших 12 месяцев влияние гласности сказалось на опубликовании материалов, до этого неизвестных на Западе, материалов о неудачах и даже катастрофах научных проектов, финансируемых правительством СССР.

Согласно данным, обнародованным министром финансов В. Павловым, правительство собирается выделить на гражданскую науку 17,34 млрд долл. в 1990 г. по сравнению с 11,5 млрд долл. в 1989 г. (т. е. наблюдается беспрецедентный рост с 1,5 до 2,3 %). Нужно подождать, чтобы увидеть, как гласность повлияет на количество статей советских ученых, но есть положительные признаки того, что более свободное и широкое распространение результатов исследований в СССР возможно.

Имеется несколько примеров создания новых англоязычных журналов, которые являются совместными изданиями АН СССР и Королевского общества Великобритании по химии (журналы Biomedical Science, Mendeleev Communications).

Институт ISI (Institute for Scientific Information) всегда поддерживал научные обмены по информатике. В марте 1987 г. институт посетил профессор С. П. Капица. Другими приглашенными учеными были морской биолог профессор А. Пудовкин из Института биологии моря во Владивостоке (ноябрь 1987 – март 1988 г.), математик

В. Налимов из МГУ (октябрь – декабрь 1988 г.), заведующий информационным отделом ВИНТИ В. Маркусова (сентябрь 1989 – январь 1990 г.). Маркусова внесла значительный вклад в подготовку этой публикации. Совсем недавно И. Маршакова-Шайкевич из Института философии, которая независимо от Генри Смолла открыла «анализ на основе ко-цитирования», также была гостем ISI.

Поскольку интерес ко всему происходящему в СССР огромен, мы решили выполнить анализ цитируемости советской науки. Нельзя сказать, что мы игнорируем страны Восточной Европы, но их работы имеют ограниченный импакт в наших исследованиях по публикациям с высокой цитируемостью.

Методология и вопросы, связанные с институтской аффилиацией

Мы решили проанализировать советскую литературу за 15-летний период. Данные для нашего исследования мы взяли из ISI Science Indicators. Эта база данных содержит сведения о 4,5 млн процитированных статей за 1973–1988 гг. и является частью массива базы данных SCI за 1945–1988 гг. Все статьи из этого массива были процитированы по крайней мере один раз. Около 3,83 млн статей из этого массива никогда не были процитированы за этот период времени. Для нашего исследования были отобраны 180 тыс. статей¹, опубликованных

¹Все журналы, используемые для подготовки Индекса, называются журналами-источниками (Source Journals), а содержащиеся в них статьи – статьями-источниками. Статьи, опубликованные в этих журналах, называют статьями-источниками (source article). Авторы этих статей-источников называют цитирующими авторами (citing authors). Ссылки, содержащиеся в этих статьях, называют цитируемыми статьями (cited articles), а их авторов – цитируемыми авторами (cited authors).

советскими учеными за 1973–1988 гг. Для анализа и дальнейшего обсуждения из этого массива были отобраны 892 статьи, процитированные не менее 50 раз. Наш анализ не был ограничен выбором только первого автора, учитывались все соавторы.

Хочу отметить, что советские ученые, опубликовавшие высокоцитируемые статьи до 1973 г., отсутствуют в этом файле. Например, нобелевский лауреат Лев Ландау был вторым по цитируемости в нашем исследовании из 250 самых цитируемых авторов за 1961–1975 гг. по массиву SCI. Однако Ландау, умерший в 1968 г., публиковал статьи с 1930 г. до середины 60-х гг. и поэтому отсутствует в нашем списке. Недавно появилась очень интересная рецензия на книгу о его отношениях с П. Капицей.

Дополнительным препятствием в нашем исследовании явилась идентификация автора статьи. Когда советские ученые работают за границей по научному обмену и в статье ставят адрес организации, в которой они работают за рубежом, эта статья не может попасть в файл советских ученых. По этой причине в нашем анализе отсутствует статья известного советского математика академика Л. Фадеева *Operator anomaly for the Gauss law*, опубликованная в 1984 г. и процитированная к концу 1988 г. более 160 раз. Но Фадеев проставил адрес организации, в которой он работал в Италии — *Scuola Normale Superiore di Pisa*.

Проблемы с адресами — наша постоянная забота. Worse всего, когда адрес вообще не проставлен. Многие советские журналы вообще не содержат адреса авторов. Обычно это объясняется соображениями секретности. Советский биохимик и историк науки Ж. Медведев, работающий в Национальном институте по медицине в Лондоне, именно так объясняет эту проблему.

Многие задаются вопросом, зачем беспокоиться об адресах авторов, если они все равно не могут иметь личные контакты. Я только могу надеяться, что гласность позволит советским ученым давать не только их адреса, но и телефоны, и факсы.

Еще одна проблема с идентификацией авторов связана с тем, что два или три автора могут иметь одну и ту же фамилию и инициалы. Обычно в большинстве случаев эта проблема легко разрешается при обращении к журналу, если они работают в областях, очень далеких друг от друга, например иммунологии и переработки нефти. Однако если они работают в одном и том же направлении, мы должны полагаться только на адрес организации автора, чтобы отличить их друг от друга. Очевидно, что эта стратегия менее эффективна, если мы имеем дело с советскими авторами, у которых вообще нет адресов. Мы не можем обращаться к каждому автору и просить дополнить библиографию его статьи, это отнимает много времени и слишком дорого стоит, поэтому нет никакой гарантии, что эта проблема может быть решена. Мы сделали все, что было в наших силах, чтобы гарантировать точность в нашем списке наиболее цитируемых советских ученых, и уверены, что эта проблема была сведена к минимуму.

Сто советских ученых — все ли члены академии?

В таблице 1 приведен список 100 самых цитируемых советских ученых за 1973–1988 гг. В колонках слева направо расположены сведения о количестве статей за 15-летний период (B), а также среднее количество ссылок на статью (импакт, C) и количество ссылок на каждого автора, полученных в 1988 г.² (D).

²Звездочкой отмечены ученые, являющиеся членами АН СССР, а крестиком — члены-корреспонденты.

Учитывая колоссальное влияние АН СССР в выборе политики в науке и технике, можно было ожидать ученых с наиболее высоким импактом именно среди членов академии. Но это не так. Четыре наиболее цитируемых советских ученых не являются членами академии или ее членами-корреспондентами: физики В. С. Летохов, В. И. Захаров, В. А. Шифман и А. И. Вайнштейн. Только 35 ученых из списка топ-100 являются членами академии. Трое академиков – биохимик Ю. А. Овчинников,

А. Н. Несмеянов (бывший президент АН СССР) и ядерный физик Я. Б. Зельдович – умерли. В нашем списке есть также умерший академик Академии медицинских наук В. М. Жданов. Следует отметить, чтобы быть честными, что не все члены не советских академий являются высокоцитируемыми учеными. Когда в 1981 г. мы изучали список тысячи самых цитируемых авторов, только 240 человек из 736 были идентифицированы как члены Американской академии наук (NAS).

Таблица 1

Список ста самых цитируемых советских ученых за 1973–1988 гг.

Автор	A	B	C	D
1. Letokhov V. S. Institute of Spectroscopy Troitsk	4,575	270	16,9	248
2. Zakharov V. I. Theoretical and Experimental Physics Institute (A) Moscow	4,401	129	34,1	14
3. Shifman M. A. Theoretical and Experimental Physics Institute (A) Moscow	4,268	101	42,3	303
4. Vainshtein A. I. Theoretical and Experimental Physics Institute (A) Moscow	4,181	78	53,6	64
5 * Ovchinnikov Y. A. Shemyakin M. M. Bioorganic Chemistry Institute (A) Moscow	4,082	113	36,1	286
6. +Polyakov A. M. L. D. Landau Theoretical Physics Institute Moscow	3,980	20	199,0	576
7. * Kochetkov N. K. Zelinskii N. D. Organic Chemistry Institute (A) Moscow	3,326	329	10,0	241
8. + Voronkov M. G. Institute of Organic Chemistry (A) Irkutsk	3,290	624	5,3	263

Автор	A	B	C	D
9. * Prokhorov A. M. Institute of General Physics (A) Moscow	3,177	490	6,5	52
10. * Georgiev G. P. V. A. Engelhardt Molecular Biology Institute (A) Moscow	3,140	123	25,5	69
11. Linde A. D. P. N. Lebedev Physics Institute (A) Moscow	2,777	53	52,4	329
12. * Korshak V. V. A. N. Nesmeyanov Organoelemental Compounds Institute (A) Moscow	2,626	617	4,3	25
13. Berezin I. V. M. V. Lomonosov State University Department of Chemistry Moscow	2,432	220	11,1	110
14. + Pudovik A. N. A. E. Arbuзов Organic and Physical Chemistry Institute (A) Kazan	2,389	548	4,4	116
15. * Zefirov N. S. M. V. Lomonosov State University Department of Chemistry Moscow	2,188	253	8,6	210
16. * Nesmeyanov A. N. A. N. Nesmeyanov Organoelemental Compounds Institute Moscow	1,997	226	88,4	375
17. * Kostyuk P. G. A. A. Bogomolets Physiology Institute (U) Kiev	1,924	55	35,0	247
18. Lippmaa E. Chemical and Biological Physics Institute (E) Tallin	1,887	72	26,2	102
19. +Skulachev V. P. M. V. Lomonosov State University A. N. Belozerski Laboratory of Molecular Biology and Bioorganic Chemistry Moscow	2,681	135	19,8	123
20. Struchkov Y. T. A. N. Nesmeyanov Organoelemental Compounds Institute (A) Moscow	3,303	464	7,1	24
21. +Sunyaev R. A. Space Research Institute (A) Moscow	1,877	35	53,6	87

Автор	A	B	C	D
22. Martinek K. M. V. Lomonosov State University Department of Chemistry Moscow	1,842	118	15,6	208
23 Kiselev A. V. M. V. Lomonosov State University Department of Chemistry Moscow	1,763	182	64,6	178
24. Altshuler B. L. B. P. Konstantinov Nuclear Physics Institute (A) Leningrdsd	1,813	36	50,4	393
25. *Reutov O. A. M. V. Lomonosov State University Department of Chemistry Moscow	1,733	206	8,4	48
26. *Zeldovich Y. B. Institute of Physics Problems (A) Moscow	1,732	91	19,0	952
27. Bergelson L. D. Cardiology Research Center (M) Moscow	1,708	176	66,5	64
28. *Kabanov V. A. A. V. Topchiev Petrochemical Synthesis Institute (A) Moscow	1,622	224	7,2	51
29. Shashkov A. S. N. D. Zelinskii Organic Chemistry Institute Moscow	1,597	171	9,3	38
30. +Privalov P. L. Institute of Protein Research (A) Pushchino-on-Oka	1,59a	61	26,1	218
31. + Fradkin E. S. P. N. Lebedev Physics Institute (A) Moscow	1,652	76	21,7	388
32. * Basov N. G. P. N. Lebedev Physics Institute (A) Moscow	1,544	199	7,8	316
33. * Ivanov V. T. M. M. Shemyakin Bioorganic Chemistry Institute (A) Moscow	1,478	90	16,4	23
34. Belov N. V. M. V. Lomonosov State University Department of Physics Moscow	1,475	409	3,6	45
35. + Larkin A. I. L. D. Landau Theoretical Physics Institute	1,475	55	26,8	23

Автор	A	B	C	D
36. Novikov V. A. Theoretical and Experimental Physics Institute (A) Moscow	1,463	56	26,1	62
37. *Razuvaev G. A. N. I. Lobachevskii State University Chemical Research Institute Gorki	1,454	220	52,1	85
38. Aronov A. G. B. P. Konstantinov Nuclear Physics Institute Gatchina	1,449	54	26,8	45
39. Spitsyn V. I. Physical Chemistry Institute (A) Moscow	1,397	293	4,8	96
40. +Tsvetkov V. N. Institute of Macromolecular Compounds Leningrad	1,367	143	9,6	185
41. Shklovskii B. I. A. F. Ioffe Physical Technical Institute (A) Leningrad	1,347	51	26,4	131
42. * Bayev A. A. Institute of Molecular Biology Moscow	1,338	77	17,4	14
43. *Goldanskii V. I. Chemical Physics Institute (A) Moscow	1,309	136	9,6	167
44. Zamolodchikov A. B. L. D. Landau Theoretical Physics Institute (A) Moscow	1,295	38	34,1	188
45. + Petrov A. A. Physical Technical Institute (A) Moscow	1,280	283	4,5	47
46. Magi M. Institute of Chemistry and Biological Physics (E) Tallin	1,251	40	31,3	21
47. * Plate N. A. M. V. Lomonosov State University Department of Chemistry Moscow	1,220	131	9,3	108
48. + Bunkin F. V. General Physics Institute (A) Moscow	1,197	137	8,7	121

Автор	A	B	C	D
49. Shuryak E. V. Novosibirsk Nuclear Physics Institute (A) Novosibirsk	1,196	40	30,0	232
50. Krishtal O. A. A. A. Bogomolets Physiology Institute (U) Kiev	1,192	30	40,0	35
51. + Bystrov V. F. M. M. Shemyakin Bioorganic Chemistry Institute (A) Moscow	1,185	55	21,5	76
52. Kuzmin V. A. Chemical Physics Institute (A) Moscow	1,181	148	8,0	58
53. Drachev L. A. M. V. Lomonosov State University Belozerskii A. N. Laboratory of Molecular Biology and Bioorganic Chemistry Moscow	1,179	54	21,8	34
54. Akhrem A. A. Institute of Bioorganic Chemistry (B) Minsk	1,170	213	5,5	68
55. Minkin V. I. Rostov State University Physics and Organic Chemistry Research Institute Rostov	1,170	198	5,9	66
56. * Sagdeev R. Z. Space Research Institute (A) Moscow	1,168	119	9,13	117
57. Bakayev V. V. Novosibirsk Medical Institute Novosibirsk	1,149	22	52,2	6
58. Shuvalov L. A. Institute of Crystallography (A) Moscow	1,141	181	6,3	17
59. Petrovskii P. V. A. N. Nesmeyanov Organoelemental Compounds Institute Moscow	1,135	137	8,3	0
60. Efros A. L. A. F. Ioffe Physical Technical Institute (A) Leningrad	1,119	57	19,6	94
61. Dmitriev B. A. N. F. Gamaleya Epidemiology and Microbiology Institute (M) Moscow	1,117	89	12,6	29

Автор	A	B	C	D
62. *Gorkov L. P. L. D. Landau Theoretical Physics Institute (A) Moscow	1,116	54	20,7	111
63. Belavin A. A. L. D. Landau Theoretical Physics Institute (A) Moscow	1,114	18	61,9	235
64. Lipatov Y. S. Institute of the Chemistry of Macromolecular Compounds (U) Kiev	1,112	196	5,7	134
65. + Zakharov V. E. L. D. Landau Theoretical Physics Institute (A) Moscow	1,109	61	18,1	275
66. Ryskov A. P. Molecular Biology Institute (A) Moscow	1,090	51	21,4	4
67. Shuvalov V. A. Institute of Soil Science and Photosynthesis(A) Pushchino-on-Oka	1,089	54	20,2	4
68. Tseytlin A. A. P. N. Lebedev Physics Institute (A) Moscow	1,087	41	26,5	51
69. Nametkin N. S. A. V. Topchiyev Petrochemical Synthesis Institute (A) Moscow	1,082	125	8,7	0
70. Levanyuk A. P. A. V. Shubnikova Crystallography Institute (A) Moscow	1,072	73	14,7	59
71. Nefedov V. I. N. S. Kurnakov General and Inorganic Chemistry Institute (A) Moscow	1,063	105	10,1	70
72. Skryabin K. G. Molecular Biology Institute (A) Moscow	1,062	40	26,6	9
73. Davankov V. A. A. N. Nesmeyanov Organoelemental Compounds Institute (A) Moscow	1,040	84	12,4	34
74. Shibaev V. P. M. V. Lomonosov State University Department of Chemistry Moscow	1,039	108	9,6	53
75. Zhdanov V. M. D. I. Ivanovskii Virology Institute (M)	1,035	138	7,5	34

Автор	A	B	C	D
76. Klimov V. V. A. N. Bakh Biochemistry Institute (A) Moscow	1,021	67	15,2	55
77. Lutsenko I. F. M. V. Lomonosov State University Department of Chemistry Moscow	1,001	191	5,2	20
78. Abdulaev N. G. M. V. Shemyakin Bioorganic Chemistry Institute (A) Moscow	997	37	27,0	12
79. Kumakhov M. A. M. V. Lomonosov State University Institute of Nuclear Physics Moscow	990	50	19,8	35
80. Voloshin M. B. Theoretical and Experimental Physics Institute Moscow	987	44	22,4	62
81. Yermakov Y. I. Institute of Catalysis Novosibirsk	982	95	10,3	36
82. Orlov S. N. Central Research Laboratory Ministry of Public Health USSR Moscow	969	67	14,5	43
83. Boreskov G. K. Institute of Catalysis (A) Novosibirsk	964	109	8,8	46
84. Postnov Y. V. Central Research Laboratory Ministry of Public Health USSR Moscow	957	47	20,4	79
85. * Spirin A. S. Institute of Protein Research (A) Pushchino-on-Oka	952	60	15,9	65
86. *Knorre D. G. Institute of Bioorganic Chemistry (A) Novosibirsk	937	76	12,3	34
87. Ptitsyn O. B. Institute of Protein Research (A) Pushchino-on-Oka	932	57	16,4	48
88. *Chazov E. I. Moscow Cardiology Center Moscow	932	42	22,2	118
89. Kachanov V. A. Serpukhov High Energy Physics Institute Serpukhov	925	82	11,3	0

Автор	A	B	C	D
90. Feigina M. Y. M. M. Shemyakin Institute of Bioorganic Chemistry (A) Moscow	924	23	40,2	0
91. Osipov O. A. Rostov Don University Physical and Organic Chemistry Institute Rostov Don	923	178	5,2	31
92. Sokofov V. I. A. N. Nesmeyanov Organoelemental Compounds Institute (A) Moscow	919	117	7,9	443
93. Shakura N. I. Shternberg Astronomy Institute Moscow	916	14	65,4	77
94. Makarov G. N. Spectroscopy Institute(A) Troitsk	908	41	22,1	1
95. Brandt N. B. M. V. Lomonosov State University Department of Physics Moscow	903	124	7,3	115
96. Saks V. A. Cardiovascular Surge-Institute (M) Moscow	902	51	17,7	33
97. Rubin A. B. M. V. Lomonosov State University Department of Biophysics Moscow	902	125	7,2	13
98. Samoson A. Institute of Chemical Physics and Biophysics (E) Tallin	897	21	42,7	31
99. Karden A. D. M. V. Lomonosov State University A. N. Belozerskii Laboratory of Molecular Biology and Bioorganic Chemistry Moscow	897	33	27,0	2
100. Frankkamenetskii M. D. Institute of Molecular Genetics Moscow	880	32	27,5	44

(A) = Academy of Sciences

(M) = Academy of Medical Sciences

(E) = Academy of Sciences EsSSR

(U) = Academy of Sciences UkSSR

(B) = Academy of Sciences BeSSR

Кто есть кто в советской науке: физики доминируют

Среди ста наиболее высокоцитируемых советских ученых (табл. 1) треть — физики (35 человек), 32 человека — ученые из области наук о живой природе, 30 химиков и трое занимаются исследованиями космоса. Среди 35 физиков десять являются членами академии: члены-корреспонденты Ф. В. Бункин, Е. С. Фрадкин, А. И. Ларкин, А. М. Поляков и В. Е. Захаров; академики нобелевский лауреат Н. Г. Басов, В. И. Гольдандский, Л. П. Горьков, нобелевский лауреат А. М. Прохоров и покойный Я. Б. Зельдович. Все они работают в институтах, расположенных в Москве. Не члены академии в подавляющем большинстве работают также в Москве (16), но есть и представители других городов: в Ленинграде (3), Таллине (3), Гатчине (1), Новосибирске (1), Свердловске (1), Серпухове (1) и Троицке (1). Ж. Медведев полагает, что доминирование физиков было бы большим, если бы физики-ядерщики могли публиковать свои работы за рубежом. Более того, он указывает, что многие физики-ядерщики (а также физики, работающие в других областях оборонного значения) не могут публиковать свои работы в советских журналах. Без сомнения, то же самое относится к западным ученым, работающим в оборонных областях.

Среди 32 ученых, занимающихся исследованиями в науках о жизни, четыре наиболее высокоцитируемых ученых являются членами академии. Это покойный академик Ю. А. Овчинников, молекулярный биолог Г. П. Георгиев (автор комментария в Citation Classics), физиолог П. Г. Костюк и биохимик В. П. Скулачев. Среди этих 32 ученых в Москве работают 23 человека, двое в Киеве, в Пушино-на-Оке — четверо (в том числе академик Спирин, автор комментария

в Citation Classics), двое в Новосибирске и один в Минске.

Среди 30 химиков двенадцать являются членами академии. Подавляющее большинство из них работает в Москве, по двое в Новосибирске и в Ростове-на-Дону, по одному в каждом из следующих городов: Ленинград, Иркутск, Казань, Горький, Киев, Таллин и Троицк.

Все трое ученых, занимающихся космическими исследованиями, являются членами Академии наук и работают в Москве.

В общем среди топ-100 советских ученых 71 человек работает в Москве, по четыре в Ленинграде, Новосибирске и Пушино-на-Оке, по три человека в Киеве и Таллине. На рис. 1 представлена карта СССР и отмечены города, ученые которых опубликовали не менее 250 статей в 1989 г. В 1989 г. для подготовки SCI использовались свыше 3200 научных журналов, среди которых 127 — советские журналы, публикуемые на русском или английском языке.

Факторы, влияющие на цитируемость советских ученых

Ознакомившись с моей статьей, Жорес Медведев заметил, что большинство ссылок на работы советских ученых были сделаны в советских журналах и очень мало ссылок из американских и западных журналов. Медведев также отметил, что наиболее цитируемые и наиболее плодотворные авторы, Летохов и Воронков, имеют значительный процент самоцитируемости. Он отмечает, что в СССР почти обязательно цитировать так называемых «классиков» и ученых, которые имеют большое влияние. Например, у Чазова, бывшего министра здравоохранения, даже его статья в газете «Аргументы и факты» включена в список его работ. Все это

только потому, что он был министром, а не благодаря его научным достижениям. Другим фактором, влияющим на цитируемость, является должность главного редактора научного журнала. Например, это относится и к цитируемости академика В. П. Скулачева, главного редактора журнала «Биохимия», и к цитируемости покойного академика В. М. Жданова, бывшего главным редактором журнала «Вопросы вирусологии».

Советские ученые с наиболее высоким импактом не обязательно опубликовали большее количество работ

Самый высокий импакт среди советских авторов за 15-летний период принадлежит физику-ядерщику А. Полякову, члену-корреспонденту АН СССР. 20 его статей получили 4000 ссылок, цитируемость каждой статьи в среднем – 200. У кристаллографа Белова самый низкий из топ-100 импакт



Рис. 1. Карта СССР, на которой отмечены города, ученые которых опубликовали не менее 250 статей в 1989 г.

(3,6 ссылки). Средний импакт в этом списке – 23,3 ссылки. Для членов академии этот импакт – 24,6 ссылки. Однако если из этого списка исключить супер-звезду А. Н. Полякова, то средний импакт академии – 19,4 ссылки. После Полякова самый высокий импакт имеют шесть ученых: покойный академик Несмеянов (88,3 ссылки), кардиолог Бергельсон³ (66,5 ссылки), астроном Шакура (65,4 ссылки), физик Киселев (64,6 ссылки), физик Вайнштейн и специалист по исследованию космоса Сюняев (у каждого из них по 53,6 ссылки на статью). Авторы, опубликовавшие наибольшее количество работ, это физики и химики: Воронков, опубликовавший 624 статьи, Коршак (617 статей), Пудовик (548 статей), Прохоров (490 статей), Стручков (464 статьи), Белов (409 статей) и Кочетков (329 статей). 53 ученых, включенных в табл. 1, опубликовали 100 или менее статей, а 30 ученых – от 100 до 200. Из 17 ученых, опубликовавших более 200 статей, одиннадцать являются членами академии.

Преобладание физиков – признак советского милитаризма?

Физика – это та область, в которой в СССР были выполнены первоклассные исследования в течение последних 50 лет. Советские физики – уважаемые, влиятельные ученые, они выполняли научные исследования на передовом крае мировой науки. Данные, представленные в этой статье, подтверждают это: из свыше 100 наиболее цитируемых ученых 35 – физики. Обычно в списке самых цитируемых ученых, не разделенных по областям науки, ученые в области наук о жизни доминируют. Противоположный случай

³Ю. Гарфилд ошибся, его специальностью была биохимия.

с учеными СССР. Мы попросили Медведева прокомментировать это. Он предположил, что в прошлом в Советском Союзе военный истеблишмент сконцентрировал усилия национальной науки на проектах в области физики, таких как создание ядерного оружия. Физики, по мнению Медведева, имели самое лучшее финансирование по сравнению с любой другой областью науки. В то же время существовало идеологическое вмешательство, следствием чего было катастрофическое влияние на биологию (деятельность Т. Лысенко), химию (отрицалась теория Л. Поллинга) и кибернетику.

Согласно мнению нобелевского лауреата Р. Хоффмана (R. Hoffman), милитаризм не единственная причина доминирования физики в советской науке. По его мнению, другим фактором является традиция в советском (русском) обществе, подразумевающая огромное уважение к ученым. «Роль модели очень важна в советской (русской) истории. Три великих физика сформировали советскую науку: Л. Ландау, И. М. Френкель и П. Л. Капица, и не только физику, поскольку они привели советскую молодежь в науку. Знаменитость и престиж науки предшествовали периоду Второй мировой войны.

Превосходство советской физики подтверждается, если мы посмотрим на список нобелевских лауреатов. Из десяти Нобелевских премий, которыми были удостоены советские ученые, семь относятся к физике. В 1958 г. эту премию получили И. М. Франк, И. Е. Тамм и П. Черенков. Ландау получил премию в 1962 г. за работы в области конденсированных сред, особенно по жидкому гелию. Басов, Прохоров и американец Ч. Тоунс (Charles

Townes) получили премию в 1964 г. (Прохоров занимает 9-е место в списке наиболее цитируемых советских ученых, см. табл. 1). П. Капица, который разделил в 1978 г. премию с американцами А. Пензиасом (Arno Penzias) и Р. Улсоном (Robert Wilson), был удостоен награды за изобретения и открытия в области низких температур.

Среди советских физиков-лауреатов наблюдается интересная тенденция в премиях за 1958 и 1964 гг., врученных за исследования, связанные с лазерами. Тамм и Франк нашли объяснение фосфоресцированию жидкости при ее облучении гамма-лучами, явление, открытое Черенковым. Это открытие оказало огромное влияние на лазерные технологии и было удостоено премии в 1958 г. Исследования Басова, Прохорова и Ч. Тоунса были направлены на развитие лазеров и были удостоены премии в 1964 г. Самый высокоцитируемый в массиве ISI Science Indicators советский ученый Летохов является пионером в области лазерной спектроскопии.

Во второй части моего очерка мы проанализируем наиболее цитируемые статьи, институты, исследовательские фронты и журналы.

Часть 2.

Пятьдесят статей из СССР, принадлежащих к Citation Classics

В таблице 2 представлен список пятидесяти наиболее цитируемых статей, опубликованных советскими учеными, из массива ISI Science Indicators. Большинство из этих идентифицированных статей могут быть отнесены к Citation Classics, поскольку их цитируемость превышает порог в 200 ссылок. Из этих пятидесяти статей к тематике, связанной с физикой, относятся 28 статей. Эти статьи относятся к исследованиям по физике элементарных частиц, сверхпроводимости, электропроводности твердых веществ и физике низких температур. К исследованиям по наукам о жизни относятся 15 статей, и они посвящены изучению структуры, формы и активности молекул белка, хроматина и аминокислот (aminoacyl-tRNA). Несколько статей по наукам о жизни посвящены исследованиям прохождения (передачи) через мембраны красных кровяных телец, электрических токов в нейронах моллюсков (electrical currents in mollusk neurons). Пять статей по химии связаны с исследованиями по анализу абсорбции ядра и использованию ядерно-магнитного резонанса для изучения структуры силикатов ядра и углеводных соединений. Оставшиеся две статьи посвящены наукам о земле и исследованию космоса.

Таблица 2

Список пятидесяти наиболее цитируемых статей, опубликованных советскими учеными, из массива ISI Science Indicators

Количество ссылок	Библиографические данные
839	Shifman M. A., Vainshtein A. I. & Zakharov V. I. QCD and resonance physics: theoretical foundations. Nucl. Phys. B 147:385-447, 1979.
730	Shakura N. I. & Sunyaev R. A. Black holes in binary systems: observational appearance. Astron. Astrophys. 24:337-55, 1973.
682	+Polyakov A. M. Particle spectrum in quantum field theory. L'ETP Lett.-Engf. Tr. 20:194-5, 1974.

Количество ссылок	Библиографические данные
677	+Polyakov A. M. Quantum geometry of bosonic strings. Phys. Lett. B 103:207-10, 1981.
615	Shifman M. A., Vainshtein A. I. & Zakharov V. L. QCD and resonance physics: applications, Nucl. Phys. B 147:448-518, 1979.
533	+Chlrikov B V. Universal instability of many-dimensional oscillator syate.ms. Phys. Rep. –Rev. Sect. Phys. Lat. 52:263-379, 1979.
503	Linde A. D. A new inflationary universe scenario: a possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. Phys. far. B 108:389-93, 1982.
437	+Polyakov A. M. Quark confinement and topology of gauge theories. Nucl. Phys. B 120:429-58, 1977.
417	Belavin A. A., +Polyakov A. M. & Zamolodchikov A. B. Infinite conformal symmetry in two-dimensional quantum field theory. Nucl. Phys. B 241:333-80, 1984.
399	Altshuler B. L., Aronov A. G. & Lee P. A. Interaction effects in disordered Fermi systems in two dimensions. Phys. Rev. Lat. 44:1288-91, 1980.
394	Matveev V. A., Muradyan R. M. & +Tavkhelidze A. N. Automodellism in the large-angle elastic scattering and structure of hadrons. Mr. Nuovo Cimento 7:719-23, 1973.
377	+Polyakov A. M. Interaction of Goldstone particles in two dimensions: applications to ferromagnets and massive Ymg-Mills fields. Phys. Lat. B 5979-81, 1975.
369	Privalov P. L. Thermodynamic approach to problem of stabilization of globular protein structure: calorimetric study. J. Mol. Biol. 6565-84, 1974.
366	Gerr R. G., Yanovsky A. I. & Struchkov Y. T. Perfection of the system of crystallographic programs in the laboratory of X-ray structure: analysis of the A. N. Nesmeyanov Institute of Organoelemental Compounds of the USSR Academy of sciences. Kristallografiya SSSR 28:1029-30, 1983.
359	Altshuler B. L., Khmel'nitzkii D., +Larkin A. I. & Lee P. A. Magneto-resistance and Hall effect in a disordered two-dimensional electron gas. Phys. Rev. B—Condensed Matter 22:5142-53, 1980.
337	Varshavsky A. J., Bakayev V. V. & Georgiev G. P., Heterogeneity of chromatin subunits in vitro and location of histone H 1. Nucl. Acid. Res. 3:477-92, 1976.
336	+Polyakov A M. Compact gauge fields and infrared catastrophe. Phys. Lest. B 59:82-4, 1975.
331	Linde A. D. Phase transitions in gauge theories and cosmology. Rep. Progr. Phys. 42:389-437, 1979.
324	*Ovchinnikov Y. A., Abdulaev N. G., Feigina M. Y., Kiselev A. V. & Lobanov N. A. Structural basis of the functioning of bacteriorhodopsin: overview. FEBS Lert. 100:219-24, 1979.
304	Ivanov V. I., Minchenkova L. E., Schyolkina A. K. & Poletayev A. L. Different conformations of double-stranded nucleic acid in solution as revealed by circular dichroism, Biopolymers 12:89-110, 1973.

Количество ссылок	Библиографические данные
300	+Polyakov A. M. Quantum geometry of fermionic strings. Phys. Lett. B 103:211-3, 1981.
296	Novikov V. A., +Okun L. B., Shifman M. A., Vainshtein A. I., Voloshin M. B. & Zakharov V. L. Charmonium and gluons. Phys. Rep. — Rev. Sect. Phys. Lett. 41:1-133, 1978.
293	Lvov B. V. Electrothermal atomization: way toward absolute methods of atomic-absorption analysis. Spectrochim. Acta Pt. B—At. Spec. 33:153-93, 1978.
292	*Kostyuk P. G., + Kriashal O. A. & Shakhovarov Y. A. Separation of sodium and calcium currents in somatic membrane of mollusk neurons. J. Physiol-London 270:545-68, 1977.
280	Lippmaa E., Magi M., Samoson A., Engelhardt G. & Grimmer A.-R. Structural studies of silicates by solid-state high-resolution ²⁹ Si NMR. J. Am. Chem. Soc. 102:4889-93, 1980.
279	Yagubskii E. B., Shchegolev I. F., Laukhin V. N., Kononovich P. A., Karatsova M. V., Zvyaginskaya A. V. & Boravov L. L. Normal-pressure superconductivity in an organic metal (BEDT-TTP)2I3 bis(ethylene dithiol) tetrathiofulvalene triiodide]. JETP Lett.—Engl. Transl. 39:12-6, 1984.
271	*Migdal A. B. Pion fields in nuclear matter. Rev. Mod. Phys. 50:107-72, 1978.
269	Shifman M. A., Vainshtein A. I. & Zakharov V. L. QCD and resonance physics: RHO-omega mixing. Nucl. Phys. B 147:519-34, 1979.
264	Sun S. S., Nesbitt R. W. & Sharaskin A. Y. Geochemical characteristics of mid-ocean ridge basalts. Earth Planet. Sci. Lett. 44:119-38, 1979.
256	Serbinenko F. A. Balloon catheterization and occlusion of major cerebral vessels. J. Neurosurg. 41:125-45, 1974.
252	Shuryak E. V. Quantum chromodynamics and the theory of superdense matter. Phys. Rep. — Rev. Sect. Phys. Lett. 61:71-158, 1980.
249	Burstein E. A., Vedenkina N. S. & Ivkova M. N. Fluorescence and the location of tryptophan residues in protein molecules. Photochem. Photobiol. 18:263-79, 1973.
247	Vaskovsky V. E., Kostetsky E. Y. & Vasendin I. M. Universal reagent for phospholipid analysis. J. Chromatog. 114:129-41, 1975.
246	Lubimov V. A., Novikov E. G., Nozik V. Z., Tretyakov E. F. & Kozik V. S. An estimate of the epsilon-E mass from the beta-spectrum of tritium in the valine molecule. Phys. Lett. B 94:266-8, 1980.
246	Rubakov V. A. Adler-Bell-Jackiw anomaly and fermion number breaking in the presence of a magnetic monopole. Nucl. Phys. 203:311-48, 1982.
236	Lim V. I. Algorithms for prediction of alpha-helical and beta-structural regions in globular proteins. J. Mol. Biol. 88:873-94, 1974.
235	Shifman M. A., Vainshtein A. I. & Zakharov V. I. Asymptotic freedom, light quarks and origin of $\delta t = 1/2$ rule in non-leptonic decays of strange particles. Nucl. Phys. B 120:316-24, 1977.
228	Starobinsky A. A. Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations. Phys. Lett. B 117:175-8, 1982.

Количество ссылок	Библиографические данные
226	Kukhtarev N. V., Markov V. B., Odulov S. G., Soskin M. S. & Vinetskii V. L. Holographic storage in electrooptic crystals. 1. Steady-state. <i>Ferroelectrics</i> 22:949-60, 1979.
225	Fesenko E. E., Kolesnikov S. S. & Lyubarsky A. L. Induction by cyclic GMP of cationic conductance in plasma membrane of retinal rod outer segment. <i>Nature</i> 313:310-3, 1985.
221	Friedenstein A. J., Chailakhyan R. K., Latsinik N. V., Panasyuk A. F. & Keiliss-Borok I. V. Stromal cells responsible for transferring the microenvironment of hematopoietic tissue: cloning in vitro and retransplantation in vivo, <i>Transplantation</i> 17:331-40, 1974.
214	Atiyah M. F., Hitchin N. J., Drinfeld V. G. & Manin Y. I. Construction of instantons. <i>Phys. Lett. A</i> 65:185-7, 1978.
211	Abelev G. L. Alpha-fetoprotein as a marker of embryospecific differentiations in normal and tumor tissues. <i>Transplant. Rev.</i> 20:3-37, 1974.
205	Pustnov Y. V., Orlov S. N., Shevchenko A. & Adler A. M. Altered sodium permeability, calcium-binding and Na-K-ATPase activity in red blood cell membrane in essential hypertension. <i>Pflugers Arch. – Eur. J. Physiol.</i> 371:263-9, 1977.
204	*Kostyuk P. G. & Krishtal O. A. Effects of calcium and calcium-chelating agents on inward and outward current in membrane of mollusk neurons. <i>J. Physiol. – London</i> 270:569-80, 1977.
201	Bugomolnyi E. B. Stability of classical solutions, <i>Sov. J. Nucl. Phys. – Engl. Tr.</i> 24:449-54, 1976.
201	Efros A. L. & Shklovskii B. I. Coulomb gap and low-temperature conductivity of disordered systems. <i>J. Phys. C—Solid State Phys.</i> 8:1A9-51, 1975.
201	Kisselev L. L. Aminoacyl-tRNA synthetases: some recent results and achievements. <i>Advan. Enzymol. Relat. Areas Mol.</i> 40:141-238, 1974.
198	+Polyakov A. M. String representations and hidden symmetries for gauge fields. <i>Phys. Lett. B</i> 82:247-50, 1979.
193	Shashkov A. S. & Chizhov O. S. C13 NMR spectroscopy in chemistry of carbohydrates and related compounds. <i>Bioorg. Khim.</i> 2:437-97, 1976.

В первой части мы говорили о доминировании физики в анализируемом массиве высокоцитируемых статей (892 статьи) за 1973–1988 гг.: 43 % относятся к физике, 32 % – к биологическим наукам и 19 % – к химии. Сходные данные были получены при наукометрическом изучении значительно большего массива (140 тыс. статей) мирового потока за 1981–1985 гг., выполненном венгерскими специалистами Т. Брауном (Т. Braun), В. Глэнцелом (W. Glänzel)

и А. Шубертом (А. Schubert), работающими в Будапеште. Граф в виде пирога по распределению советских статей по предметным категориям в обоих исследованиях представлен на рис. 2.

Кроме распределения статей по предметным областям из ряда стран А. Шуберт и его коллеги проанализировали импакты этих стран, т. е. среднее количество ссылок на одну статью в ряде областей знания. СССР занимал второе место в мире после США

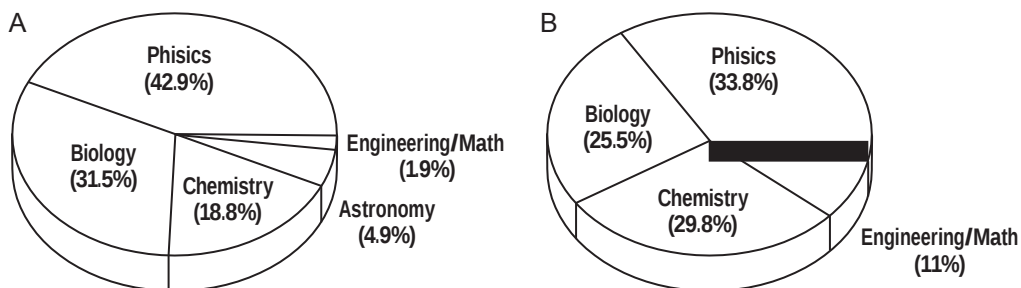


Рис. 2. Распределение статей советских ученых по предметным категориям в исследованиях 892 статей за 1973–1988 гг. (А) и массива из 140 тыс. статей мирового потока за 1981–1985 гг. (Б)

по количеству статей по физике и химии. Но СССР был шестым по импакту по физическим наукам после США, Федеративной Республики Германии, Великобритании, Японии и Франции. В области химии СССР занимал седьмое место в мире по импакту после США, Японии, ФРГ, Великобритании, Франции и Канады. СССР был седьмым по количеству статей по наукам о жизни, а по их импакту – на шестнадцатом месте.

Десять топ-статей: физика, астрономия/астрофизика

Все десять самых цитируемых статей посвящены физике или астрофизике. Шесть статей посвящены исследованиям по физике, в частности тематике квантовой хромодинамики (QCD). Эта область исследований ищет объяснение, почему кварки, которые считаются фундаментальной составляющей материи, объединяются, формируя наблюдаемые закономерности элементарных частиц, таких как протон и нейтрон. Сильное взаимодействие между кварками является тематикой фундаментальной науки о том, как атомные ядра связаны друг с другом. Теория (QCD) привлекает исследователей из-за своей математики, практиче-

ски идентичной той, которая используется в квантовой электродинамике (QED) и единой теории слабых и электромагнитных взаимодействий. Физики, занимающиеся частицами, надеются, что математическое сходство между этими теориями может быть индикатором того, что «универсальная теория поля» близка к реализации.

Три наиболее цитируемые статьи по проблемам QCD (табл. 1) были опубликованы учеными из Института теоретической физики АН СССР им. Л. Д. Ландау: это М. А. Шифман, А. И. Вайнштейн, В. И. Захаров. Все три этих физика находятся в группе из четырех самых цитируемых советских авторов в массиве за 1973–1988 гг.

Автором четырех статей, связанных с тематикой QCD, является сотрудник этого же института член-корреспондент А. М. Поляков⁴. В списке пятидесяти самых цитируемых советских статей (табл. 2) А. М. По-

⁴Профессор А. М. Поляков с 1989 г. работает в Princeton University. За свои открытия в сфере теории поля, теории струн, магнитных монополей и другие достижения удостоен премии Fundamental Physics Prize (3 млн долл.), учрежденной бывшим сотрудником ФИАН Ю. Мильнером, является лауреатом многих научных премий, в 2005 г. избран действительным членом в U.S. National Academy of Sciences.

лякову принадлежат 8 статей. Вайнштейн, Захаров и Шифман опубликовали по пять статей каждый. Каждый из следующих перечисленных ученых опубликовал по две статьи из этого списка: математик В. Л. Альтулер из Института ядерной физики АН СССР им. В. П. Константинова (Ленинград), физик А. Д. Линде из Физического института им. А. П. Лебедева АН СССР, П. Г. Костюк и О. А. Кристал (оба из Института физиологии им. А. А. Богомольца АН Украинской ССР).

Вторая по количеству ссылок статья «Черные дыры в бинарной системе: появление при наблюдении» была опубликована совместно Н. И. Шакурой из Института астрономии им. Штернберга и членом-корреспондентом Р. А. Сюняевым из Института космических исследований АН СССР. Предполагается, что черные дыры — это последствие погасших звезд, когда силы гравитации настолько сильны, что даже свет не может исчезнуть.

Другая статья по астрофизике, имеющая седьмой ранг, была опубликована Линде и связана с выдвинутой им идеей непрерывно расширяющейся Вселенной. Линде, Шакура и Сюняев входят в группу наиболее цитируемых советских ученых, о чем было написано в первой части статьи.

Шестая, наиболее цитируемая статья опубликована Б. В. Чириковым из Института ядерной физики СО АН СССР. Эта статья посвящена механизму неустойчивости в колебательных системах (Арнольда).

Десятая по цитируемости статья посвящена эффекту взаимодействия в беспорядочных системах Ферми, статья была опубликована совместно двумя советскими учеными Альтулером и Ароновым (Институт ядерной физики им. В. П. Константинова) и американцем Р. А. Лее (Bell Laboratories, New Jersey).

Научные центры с высоким импактом

Был составлен список из 25 организаций, ученые которых опубликовали 50 самых цитируемых статей. В этот список вошли одиннадцать институтов АН СССР, опубликовавших 36 таких статей (72 %). Среди неакадемических институтов выделяется ИТЭФ, опубликовавший 6 статей, три статьи были опубликованы учеными из двух институтов АН Украины и одна статья учеными АН Эстонии. Три института Академии медицинских наук (АМН) СССР опубликовали четыре статьи. Почти все организации, опубликовавшие самые цитируемые статьи, расположены в Москве или Ленинграде. В таблице 3 представлен список институтов с самым высоким импактом, т. е. средней цитируемостью одной статьи. В этом списке доминируют институты, выполняющие исследования по физике и ядерной физике. Первое место занимает Институт теоретической физики им. Л. Ландау со средней цитируемостью одной статьи — 15,86 ссылки. Этот институт также находится на первом месте по количеству самых цитируемых статей. Среди других институтов ФИАН им. Лебедева, Институт биорганической химии им. М. М. Шемякина АН, Институт органической химии им. Н. Д. Зелинского АН, Физико-технический институт им. А. И. Иоффе АН и ИТЭФ.

Советские научные публикации — выход из русского обледенения

Был составлен список журналов, опубликовавших пятьдесят самых цитируемых статей. В этом списке только два советских журнала (ЖЭТФ и «Биохимия»), двенадцать из США, по восемь из Великобритании и Нидерландов и по одному из Италии и Дании. Все эти журналы публикуются на английском языке. Из 239 научных журналов, содержащих проанализированный

Таблица 3

Список институтов с самым высоким значением импакта

Rank	Institute	Number of Papers	Number of Citation	Citation Impact
1	L. D. Landau Theoretical Physics Institute	1,254	19,896	15,86
2	Theoretical and Experimental Physics Institute	1,001	13,324	13,31
3	M. M. Shemyakin Bioorganic Chemistry Institute	1,203	10,490	8,71
4	P. N. Lebedev Physics Institute	4,615	32,742	7,09
5	I. V. Kurchatov Institute of Atomic Energy	1,812	11,246	6,20
6	N. D. Zelinskii Organic Chemistry Institute	1,408	8,647	6,14
7	Joint Institute for Nuclear Research (Dubna)	2,729	16,702	6,12
8	A. F. Ioffe Physical Technical Institute (Leningrad)	5,539	28,153	5,08
9	M. V. Lomonosov State University	16,952	82,080	4,84
10	L. Y. Karpov Physicochemical Research Institute	2,165	9,964	4,60

массив советских статей за 1973–1988 гг., только 18 % статей были опубликованы на русском языке. Оставшиеся 82 % статей были или опубликованы в западных журналах или в переводе на английский. В. Маркусова во время ее пребывания в ISI отмечала, что опубликование статьи на английском привлекает более широкую аудиторию читателей. Кроме того, временной лаг от представления статьи в редакцию до ее опубликования в советских журналах может составлять 2–2,5 года, в то время как на Западе десять лет назад он составлял от 8 до 18 месяцев. В настоящее время, с учетом технологического совершенствования издательского процесса, он может быть короче. Среди 32 журналов только один, *Bioorganicheskaya Khimiya* («Биоорганиче-

ская химия»), относится к наукам о жизни. Может быть, поэтому, как я упоминал в первой части моей статьи, создано совместное советско-английское предприятие для выпуска журнала *Biomedical Science*.

Советские исследовательские фронты — акцент на изучение кристаллов, прикладную химию и изучение сердца

В этом разделе я буду обсуждать исследовательские фронты в массивах SCI и Social SCI только за 1988 г., в которых советские публикации составляют не менее 55 % всех цитирующих статей за этот год. Эти фронты — небольшая часть базы данных исследовательских фронтов (всего их 8177). Исследовательские фронты форми-

руют по специальной методике. Этот метод, называемый кластерингом, создан на основе ко-цитирования документов. Статьи, которые часто совместно процитированы текущими статьями, создают «ядро» специальности. Цитирующие статьи создают исследовательский фронт, название которого выбирается из фраз, встречающихся в этих цитирующих статьях.

Советские статьи составляют 45 исследовательских фронтов. Мы полагаем, что специальности, в которых 55 % из всех цитирующих статей являются советскими, могут рассматриваться как области, в которых советские ученые особенно активны. Доминирующими являются исследования по физике (20 фронтов), материаловедению (13 фронтов), с прикладной химией связаны 11 фронтов. Исследования по наукам о жизни и медицине представлены также одиннадцатью фронтами, причем четыре из них по сердечно-сосудистым заболеваниям. Другие исследовательские фронты представляли: геологию (один фронт), психологию и генетику дрозофилы (один фронт) и генетику личинок сальмониды (один фронт).

Необходимо отметить, что только один исследовательский фронт по исследованию Вселенной (Low flying quadropole vibration of superfluid nuclei and cross section for universe beta decay) не имел советских авторов среди ядра цитируемых статей.

Классика цитирования

На сегодняшний день 18 советских ученых дали нам комментарии по поводу своих самых высокоцитируемых статей, опубликованных в разделе Citation Classics. Эти комментарии относятся к публикациям более раннего периода времени, чем тот, который мы обсуждали в этой статье. Исключение

составляют академик В. И. Гольданский (Институт химической физики АН СССР) и академик А. С. Спирин (Институт исследований белка АН СССР). Среди авторов комментариев восемь являются членами АН.

Мне хотелось бы узнать мнение советских ученых о тех ученых, кого нам бы следовало номинировать как авторов высокоцитируемых статей. Только, пожалуйста, не забудьте указать их адрес, тогда мы сможем обратиться непосредственно к авторам и пригласить написать комментарий.

Улучшат ли перестройка и гласность советскую науку?

Наше исследование показало, что физика является «королевой» в советской науке. Науки о жизни по сравнению с ней занимают второе место как по количеству публикаций, так и по импакту. Вероятной причиной этого является отрицательное влияние академика Т. Д. Лысенко (1898–1976), которого поддерживали Сталин и Хрущев. Псевдонаука Лысенко контролировала с конца 30-х до конца 60-х гг. аграрную науку и науки о жизни.

По мнению историка советской науки и биолога Ж. Медведева, даже теперь, когда влияние «лысенковщины» рассеялось, продолжительный эффект политизации еще существует в советской науке.

Секретность — это еще одна причина, которая влияет на развитие советской науки и не дает советским ученым воспользоваться выгодами от научных и технологических достижений. В недавнем интервью академик В. Овдеевский, председатель советского комитета по продвижению переговоров, отметил, что распространенная атмосфера секретности не дает возможности людям, работающим в одном направлении промышленности, узнать о результатах исследований в другом, замедляя, таким

образом, распространение и внедрение инновационных технологий. Технологии, разработанные для военных целей, которые могли бы принести пользу гражданскому сектору, не раскрываются из-за соображений секретности. Такие технологии включают обработку деталей, сварку, укрепление поверхностей, выращивание кристаллов, а также использование гибких автоматизированных производственных линий. Академик Р. Сагдеев, бывший директор ИКИ АН СССР, отмечает, что до перестройки ученые, работающие на военную промышленность, были избалованы, в то время как ученые, работающие в гражданских областях, имели устаревшее оборудование, низкую зарплату и невысокий шанс продвижения по служебной лестнице или международного признания. Но реструктуризация привела к худшему, по мнению Сагдеева. Вместо того чтобы поднять всех ученых до уровня военных стандартов, все ученые были понижены до уровня гражданских стандартов.

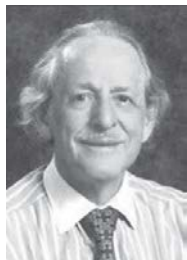
В попытке получить социальные и экономические выгоды от научных исследований советские научные институты поощряются для укрепления связей с промышленностью. Таким примером может служить совместное предприятие американского мозгового центра (think tank) — Arthur D. Little (Cambridge, MA) и АН СССР. Совместное

предприятие займется поиском советских разработок в фундаментальных науках, привлечет западных партнеров для их развития и передаст лицензию компаниям, которые сделают конечный продукт. Наиболее обещающие области разработок — материаловедение и исследования лазеров.

Другим направлением в осуществлении реформ является введение конкурсного финансирования. Вместо финансирования через академические институты часть финансирования распределяется комитетами по науке и технике. В прошлом году было подано 6200 заявок на выполнение проектов, и 3800 из них были удовлетворены. По мнению Р. Сагдеева, часть денег все еще вкладывается в убыточные предприятия.

Эта статья представляет собой обзор советской науки, и в ней обсуждаются только ученые и статьи, имеющие высокий импакт. В ней отмечены только некоторые факторы, препятствующие советской науке в достижении ее полного потенциала. У советской науки много проблем, но реформы Горбачева вселяют в западную и советскую науку большие надежды. Можно надеяться, что в ближайшем будущем произойдет как рост научных обменов между советскими и западными учеными, так и распространение электронного оборудования, такого как факсы и персональные компьютеры.

ЭВОЛЮЦИЯ SCIENCE CITATION INDEX



Ю. Гарфилд

Впервые опубликована в: Eugene Garfield The evolution of the Science Citation Index // CONTRIBUTIONS to SCIENCE, 5 (1): p. 63–70 (2009)

Institut d'Estudis Catalans, Barcelona

DOI: 10.2436/20.7010.01.60

Переводчик Валентина Александровна Маркусова.

Перевод на русский язык выполнен с любезного разрешения автора.

Введение

Впервые концепция использования Science Citation Index (SCI) как интеллектуального инструмента, облегчающего процесс распространения и поиска научной литературы, была обнародована в 1955 г. Старшее поколение ученых наверняка вспомнит уже существовавший к тому времени информационный продукт под названием Current Contents, бывший первой революционной «идеей», которая сделала возможной практическую реализацию SCI. Примечательно, что Current Contents все еще еженедельно публикуется в печатной форме, хотя его электронная версия существует более десяти лет. Даже ярые поклонники Current Contents вряд ли помнят, какую роль сыграл компьютер в его (Current Contents) создании, позволяя осуществлять еженедельный выпуск, где одновре-

менно с названиями статей также приводился указатель слов и указатель адресов авторов. В те дни обычные указатели появлялись через шесть месяцев, а иногда даже и через три года после опубликования научной литературы. Мы же в это время получали ежегодно более десяти миллионов запросов на репринты из разных стран мира после прочтения Current Contents.

Тем не менее успех SCI связан не с его первоначальной функцией — инструмента информационного поиска, а с его последующим использованием в качестве инструмента для измерения научной продуктивности, благодаря побочному продукту — базе данных Journal Citation Reports (JCR) и ранжированию научных журналов по импакт-фактору IF.

Междисциплинарная база данных SCI имеет две цели: первая — определить, что каждый ученый опубликовал, и вторая —

определить, где и как часто статьи¹ этого ученого были процитированы. Следовательно, SCI всегда делится на две части по фамилиям авторов статей: Указатель авторов статей-источников — Source Author Index и Указатель цитируемых авторов — Citation Index. Существует дополнительная возможность определить, в какой организации и стране были опубликованы статьи автора и как часто они были процитированы. Это особенно важно, потому что трудно найти полный список публикаций конкретного автора.

The Web of Science (WoS), являясь электронной версией SCI, связывает эти две функции: публикации авторов могут быть расположены в хронологическом порядке, по названию журнала или по частоте цитирования. Также можно проводить поиск по фамилиям ученых, опубликовавших статьи в течение определенного периода. В таблице 1 приведен список, полученный в результате поиска ученых, публиковавших работы в течение 70–85 лет. Например, за строкой Izaak M. Kolthoff следует молекулярный биолог Michael Heidelberger, чья последняя работа появилась в 2004 г., незадолго до его смерти в возрасте 104 лет.

Когда SCI был опубликован в 1964 г., Ирвинг Шер (Irving Sher) и я уже начали использовать библиографические ссылки для создания топологических карт, называемых историограммами, чтобы исследовать возможность использовать индексы цитирования в создании мини-историй научных направлений. Совсем недавно благодаря

¹Все журналы, индексируемые для подготовки, называются журналами-источниками, а содержащиеся в них статьи называются статьями-источниками.

Citing author — цитирующий автор, т. е. автор статьи-источника.

Cited author — цитируемый автор, т. е. автор, работа которого была процитирована в статье-источнике.

мощным современным компьютерам с гигабайтами памяти была создана программа под названием HistCite. Эта запатентованная программа находилась в разработке около пяти лет и будет доступна для продажи начиная с февраля 2009 г. Благодаря этой программе, выполнив поиск в WoS и загрузив в нее полученный файл, можно автоматически построить историограмму. Собирая все соответствующие релевантные цитируемые документы по предмету поиска в WoS, HistCite представляет собой коллективную память цитирующих авторов и производит визуальное описание истории по данной тематике. Ключевой вопрос, который часто возникает, связан со способностью метода цитирования отождествлять (извлекать) все релевантные публикации по данной теме.

В предвоенные времена и задолго до появления молекулярной биологии практика цитирования была не так стандартизирована, как сегодня, и неявное использование ссылок было довольно распространенным явлением. В результате явное цитирование ранней релевантной работы не всегда может быть найдено.

Гейдельберг (Heidelberger) был одним из первых молекулярных биологов. Он еще до Второй мировой войны, работая вместе с Освальдом Т. Авери (Oswald T. Avery) и другими учеными в Институте Рокфеллера (в том числе с Colin M. Macleod и с Maclyn McCarthy), опубликовал пионерскую работу по истории ДНК. Фактически эта работа является ключевым звеном в генеалогической истории статьи Уотсона — Крика (J. Watson & Crick) по структуре двойной спирали ДНК². Мы использовали HistCite, чтобы проследить неявную связь между этой статьей

²В 1955 г. Уотсон и Крик были удостоены Нобелевской премии за раскрытие структуры ДНК.

Уотсона — Крика и работой Авери и других соавторов по пневмококковой ДНК, опубликованной в 1944 г.

Те, кто знаком с историей, знают, что Джим Уотсон несколько лет назад наконец признал, что сожалеет, что они не процитировали работу Авери 1944 г. в своей знаменитой статье, опубликованной в 1953 г. Он и Крик очень торопились отправить статью в печать и не выполнили обычной проверки списка цитируемой литературы. Для того чтобы продемонстрировать значимость работы Авери, в действительности известную современным исследователям, мы создали несколько файлов с помощью HistCite, выполнив по-

иск в SCI, размещенный в WoS. Для того чтобы исследовать историческую связь между работой Гейдельберга и его соавтора О. Авери, необходимо было отредактировать тысячи таких неявных ссылок. На рис. 1 представлена историограмма, созданная на основе связей между работой Гейдельберга, Авери, Уотсона и Крика.

Продемонстрировав, как поисковая система WoS может быть использована для отслеживания истории развития научного направления, мы теперь можем обратиться к вопросу о часто упоминаемом импакт-факторе журнала. Ежегодный Указатель цитируемости научных журналов (Journal Citation Reports (JCR)) официально был

Таблица 1

Ученые, опубликовавшие статьи в течение 69 и более лет

Scientist	Birth/Death	Pub Years	Years Pub
Izaak Maurits (Piet) Kolthoff (analytical chemist)	1894–1993	1917–2002	86
Michael Heidelberger (organic chemist–immunologist)	1888–1991	1909–1993	85
Melvin Guy Mellon (chemist)	1893–1993	1920–2003	84
Ernst Mayr (geneticist)	1904–2005	1923–2005	83
Michel Eugene Chevreul (chemist)	1786–1889	1808–1889	82
Carl S. Marvel (polymer chemist)	1894–1988	1917–1996	80
Joel H. Hildebrand (chemist)	1881–1983	1907–1983	77
Linus Pauling (chemist)	1901–1994	1923–1998	76
John Carew Eccles (neurophysiologist)	1903–1997	1929–1992	74
Donald Coxeter (mathematician)	1907–2003	1930–2001	72
Charles Scott Sherrington (physiologist)	1857–1952	1882–1952	71
Hans Albrecht Bethe (physicist)	1906–2005	1934–2004	71
Alexander Kossiakoff (engineer; guided missile expert)	1914–2005	1935–2005	71
Norman Hackerman (chemist)	1912	1936–2006	71
Michael DeBakey (cardiac surgeon)	1908	1937–2006	70
Gerhard Herzberg (chemist)	1904–1999	1924–1992	69
Herman Mark (polymer chemist)	1895–1992	1922–1990	69

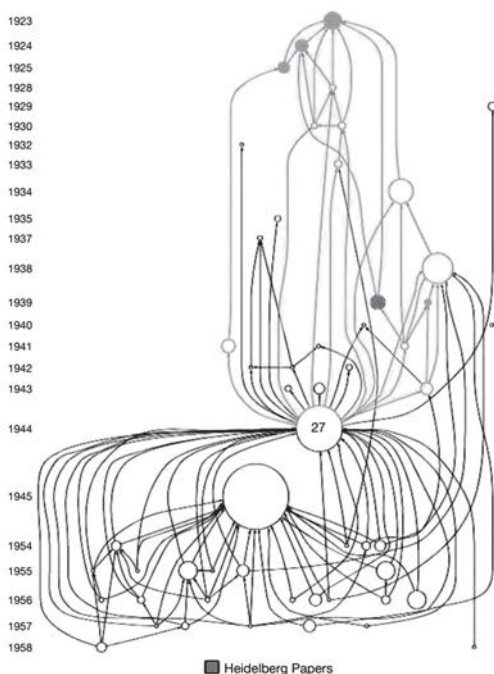


Рис. 1. Историограмма, созданная на основе связей между работой Гейдельберга (Heidelberger), Авери (Avery), Уотсона и Крика

выпущен в свет в 1975 г., хотя мы производили эти данные уже более десяти лет.

JCR возник как статистическая коллекция данных Индекса цитирования журналов (Journal Citation Index), который, в свою очередь, явился результатом пересортировки Индекса цитирования авторов (Author Citation Index), т. е. вместо сортировки в алфавитном порядке фамилий авторов, файл был отсортирован по названиям журналов, в которых были опубликованы работы авторов. Когда в начале 1960-х гг. этот тест был впервые выполнен, мы обнаружили, что журналы, уже индексируемые в Current Contents, включали либо те, которые публиковали большое число статей, либо те, которые были процитированы больше всех.

Однако необходим был простой метод, позволявший сравнивать толстые (большие) журналы, такие как Nature, Science и JAMA, с небольшими журналами, как, например, Annual Reviews. В первые дни Current Contents мы уделяли особое внимание молекулярной биологии и биохимии. Мы заметили, что 25 % всех ссылок в литературе, опубликованной в текущем году, были на статьи, опубликованные только 2–3 года назад. Именно поэтому было принято решение использовать цитирование за два года, предшествующих текущему, в качестве основы для расчета импакт-фактора за текущий год. Таким образом, импакт-фактор — это среднее число ссылок на опубликованную статью. Тем не менее мы также осознавали, что относительно небольшие, но важные, обзорные и специализированные журналы могут не попасть в выборку, если мы сосредоточимся исключительно на общем числе публикаций или на общем количестве цитирований [1]. «Импакт-фактор» журнала был создан в качестве метода сравнения журналов независимо от их размера или частоты цитируемости. На рисунке 2 сопоставлены три таблицы, содержащие сведения по трем журналам в области наук о живой природе, отсортированных по:

- (А) наиболее цитируемым журналам в 2008 г.;
- (В) количеству статей, опубликованных в 2008 г.;
- (С) значениям импакт-факторов за 2008 г.

Термин «импакт-фактор» постепенно эволюционировал, особенно в Европе, как характеристика влияния как журнала, так и автора. Эта неоднозначность часто вызывает проблемы, так как использовать импакт-факторы для сравнения журналов — это одно дело, и совсем другое дело — ис-

пользовать IF для сравнения авторов. В то время как индивидуальный автор обычно публикует в среднем небольшое количество статей (хотя есть некоторые феноменально производительные авторы), импакт-факторы журналов обычно включают относительно большие популяции статей и ссылок. Импакт-фактор журнала основан на двух элементах: числитель, который представляет собой количество ссылок, сделанных в текущем году на любые статьи из этого журнала, опубликованные в течение двух лет, предшествующих году обследования; и знаменатель — это общее число статей (статьи-источники), опубликованных за те же два года.

Импакт-фактор может быть легко рассчитан только на основе статей предыдущего года, что давало бы лучшее представление о быстро изменяющихся областях

(A)

Abbreviated Journal Title	Total Cites	Articles	Impact Factor
Nature	443,967	899	31.434
P Natl Acad Sci USA	416,018	3508	9.380
Science	409,290	862	28.103
J Biol Chem	407,492	3761	5.520
J Am Chem Soc	318,252	3242	8.091
Phys Rev Lett	310,717	3905	7.180
Phys Rev B	250,465	5782	3.322
New Engl J Med	205,750	356	50.017
Appl Phys Lett	179,925	5449	3.726
Astrophys J	177,571	2128	6.331
J Chem Phys	164,492	2763	3.149
Lancet	148,106	289	28.409
Circulation	143,852	607	14.595
Cell	142,064	348	31.253
Angew Chem Int Edit	139,534	1797	10.879
J Geophys Res	129,836	2860	3.147
Cancer Res	125,341	1228	7.514
J Immunol	123,910	1860	6.000
Blood	122,032	1237	10.432
J Neurosci	120,933	1438	7.452

(B)

Abbreviated Journal Title	Total Cites	Articles	Impact Factor
J Biol Chem	407,492	3761	5.520
P Natl Acad Sci USA	416,018	3508	9.380
J Am Chem Soc	318,252	3242	8.091
J Chem Phys	164,492	2763	3.149
J Immunol	123,910	1860	6.000
Angew Chem Int Edit	139,534	1797	10.879
J Agr Food Chem	51,062	1670	2.562
J Neurosci	120,933	1438	7.452
Biochemistry US	94,645	1437	3.379
Org Lett	46,502	1403	5.128
Nanotechnology	16,291	1397	3.446
Macromolecules	80,559	1379	4.407
J Virol	86,021	1293	5.308
Blood	122,032	1237	10.432
Cancer Res	125,341	1228	7.514
Brain Res	56,664	1187	2.494
Zootaxa	2639	1118	0.740
World J Gastroentero	10,822	1112	2.081
Neurosci Lett	28,223	1080	2.200
Nuclei Acids Res	86,787	1070	6.878

(C)

Abbreviated Journal Title	Total Cites	Articles	Impact Factor
CA-Cancer J Clin	7522	19	74.575
New Engl J Med	205,750	356	50.017
Annu Rev Immunol	15,519	24	41.059
Nat Rev Mol Cell Bio	19,628	84	35.423
Physiol Rev	17,865	40	35.000
JAMA-J Am Med Assoc	114,250	225	31.718
Nature	443,967	899	31.434
Cell	142,064	348	31.253
Nat Rev Cancer	18,908	85	30.762
Nat Genet	61,812	215	30.259
Annu Rev Biochem	16,889	31	30.016
Nat Rev Immunol	15,775	86	30.006
Nat Rev Drug Discov	10,062	62	28.690
Lancet	148,106	289	28.409
Science	409,290	862	28.103
Nat Med	48,632	141	27.553
Annu Rev Neurosci	10,132	23	26.405
Nat Rev Neurosci	15,642	71	25.940
Nat Immunol	25,245	133	25.113
Cancer Cell	12,985	78	24.962

Рис. 2. Журналы в области наук о живой природе, отсортированные следующим образом: (A) — наиболее цитируемые в 2008 г.; (B) — по числу статей, опубликованных в 2008 г.; (C) — по величине импакт-фактора в 2008 г.

исследования; или же на основе более длительного периода цитирования статей-источников, но такой показатель не был бы достаточно оперативным.

Важно отметить, что корреспонденция, письма в редакцию, новости, некрологи, редакционные интервью не включаются при подсчете числа статей-источников в JCR. Тем не менее, поскольку числитель включает ссылки на подобные «более эфемерные» виды публикаций, это приводит к некоторым искажениям. Обычно только небольшое количество журналов подвержено таким искажениям, и оно колеблется в пределах 5–10 % [2]. Кроме того, в настоящее время JCR содержит ссылки из более чем 5000 журналов; поэтому обсуждение настолько незначительных отклонений в JCR не представляется целесообразным.

Наукометрия и журналология (Journalology)

Анализ цитирования расцвел за последние три десятилетия и превратился в научную область — наукометрию, которая теперь имеет свое собственное общество — International Society for Scientometrics and Informetrics (ISSI, <http://www.issi-society.info>) и научный журнал *Scientometrics*, публикующийся с 1978 г. Более 15 лет назад Стив Локк (Steve Lock) назвал применение наукометрии к оценке журналов журналологией («Journalolog») [3].

Все исследования цитирования должны быть нормализованы, чтобы принять во внимание такие переменные, как предметная область знания, период полужизни научной литературы в этой области и плотность цитирования [4]. Период полужизни (количество ретроспективных лет, за которые было процитировано 50 % научной литературы) значительно больше для журнала по физио-

логии, чем для физического журнала. Для некоторых областей знания импакт-факторы JCR за двухлетний период могут дать / или не дать той полной картины, которую дают пятилетние или десятилетние импакт-факторы. Тем не менее, когда журналы изучаются в рамках дисциплинарных категорий, ранжирование на основе значений импакт-факторов за один, 7 или 15 лет существенно не отличается [5]. Другими словами, когда журналы были изучены по предметным категориям, то обнаружилось, что ранги журналов по физиологии значительно улучшились с течением лет, но рейтинги в пределах категории физиологии существенно не изменились. Плотность цитирования — это среднее количество ссылок, процитированных в статье-источнике. Плотность цитирования (R/S) значительно ниже для математических журналов, чем для журналов по молекулярной биологии. Существует широко распространенное, но ошибочное мнение, что размер научного сообщества, значительно влияет на импакт-фактор журнала. Это предположение игнорирует тот факт, что большое научное сообщество (т. е. большое число авторов) производит большое количество ссылок, а эти ссылки используются большим числом цитируемых статей. Большинство статей в большинстве областей не очень хорошо цитируются, в то время как некоторые статьи в небольших научных областях могут иметь значительное влияние, особенно когда они имеют междисциплинарное значение.

Хорошо известно, что существует неравномерное распределение ссылок в большинстве областей в соответствии с правилом 80/20, т. е. 20 % статей дают около 80 % ссылок. Следует повторить, что среднее число ссылок на статью и показатель оперативности (immediacy) цитирования, а не количество авторов или статей в области

являются важными [6]. Размер области знания, как правило, увеличивает количество «высокоцитируемых» («super-cited») статей. И в то время как несколько классических

методологических статей превышают высокий порог цитирования, этого не происходит с тысячами других методических и обзорных статей. Обзорные статьи, как

Таблица 2

Краткий список высокоцитируемых статей в области наук о жизни

Authors	Title	Source	Year	Volume	Page	Hits
Lowry O. N., Rosebrough N. J., Farr A. L., Randall R. J.	Protein Measurement with the Folin Phenol Reagent	Journal of Biological Chemistry	1951	193	265	293,328
Laemmli U. K.	Cleavage of Structural Proteins During Assembly of Head of Bacteriophage T4	Nature	1970	227	680	192,022
Bradford M. M.	Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding	Analytical Biochemistry	1976	72	248	120,179
Sanger F., Nicklen S., Coulson A. R.	DNA Sequencing with Chain Terminating Inhibitors	Proceedings of the Na- tional Acad- emy of Sci- ence USA	1977	74	5463	63,909
Chomczynski O., Sacchi N.	Single-Step Method of RNA Isolation by Acid Guanidinium Thiocyanate Phenol Chloroform Extraction	Analytical Biochemistry	1987	162	156	55,987
Towbin H., Staehelin T., Gordon J.	Electrophoretic Transfer of Proteins from Polyacramide Gels to Nitrocellulose Sheets – Procedure and Some Applications	Proceedings of the Na- tional Acad- emy of Sci- ence USA	1979	76	4350	48,671
Folch J., Lees M., Stanley G. H. S.	A Simple Method for the Isolation and Purification of Total Lipides from Animal Tissues	Journal of Biological Chemistry	1957	226	497	35,646
Southern E. M.	Detection of Specific Sequences among DNA Fragments Separated by Gel Electrophoresis	Journal of Molecular Biology	1975	98	503	31,273

правило, цитируются примерно в два раза чаще, чем другие статьи, но опубликование обзорных статей не обязательно повысит влияние импакта вашего журнала. Если посмотреть на таблицу 2, в которой представлен краткий список высокоцитируемых статей в области наук о жизни, то можно увидеть, что лидирует статья Лоури. Статья О. Лоури (O. N. Lowry)³ недавно обсуждалась в журнале *Biological Chemistry* [7], но авторы не упомянули комментарий Лоури по поводу того, что его самая цитируемая статья в истории науки, по его мнению, не является наиболее важной работой [8].

Неравномерность распределения ссылок повторяется как заклинание критиками импакт-фактора. Некоторые редакторы хотели бы видеть импакт, рассчитанный исключительно на основе самых цитируемых работ, таким образом, малоцитируемые работы следовало бы проигнорировать. Однако неравномерность распределения ссылок характерна для большинства журналов, а поэтому это не должно существенно повлиять на рейтинги журналов. Кто-то хотел бы видеть ранжирование по географическим областям, поскольку SCI обвиняют в доминировании англоязычных журналов. Еврофилы хотели бы иметь возможность сравнить свои журналы по языкам или по географическим группам, особенно в общественных и гуманитарных науках. Время, необходимое для рецензирования рукописей, может также оказывать влияние: если процесс работы над рукописью задерживается, то ссылки в статьях не появятся в течение двухлетнего окна цитирования, необходимого для JCR [9, 10]. С другой стороны, появление статей

на ту же тему в том же номере журнала может оказать положительное влияние.

Для большей точности было бы желательно провести анализ каждой статьи в журнале, чтобы оценить разницу в величине импакта разных типов редакционных документов и учесть это [11].

Другие возражения в отношении импакт-факторов связаны и с системой классификации журналов, используемой в JCR. В идеальной системе следовало бы иметь возможность сравнивать журналы с идентичными профилями. Но на самом деле редко можно найти два журнала с идентичными семантическими или библиографическими профилями. Эвристические, в какой-то степени несколько субъективные методы отнесения журналов к предметным категориям, используемые в ISI, отнюдь не являются совершенными, в действительности даже специалисты используют анализ цитируемости для обоснования своего решения. Недавно были выполнены попытки сгруппировать журналы более объективно, опираясь на двусторонний анализ отношений цитируемости между журналами, чтобы уменьшить влияние названий журналов на их классификацию [12]. Например, анализ цитируемости показал, что журнал *Journal of Experimental Medicine* был ведущим журналом по иммунологии и даже в наши дни по-прежнему входит в пятерку лидирующих журналов по иммунологии по показателю их импакт-факторов. В JCR недавно добавили новую функцию, которая позволяет пользователям более точно установить предметную категорию журнала на основе взаимосвязей цитируемости. Рисунок 3 показывает общую формулу для расчета взаимосвязей цитируемости между двумя журналами и коэффициента взаимосвязи, выражающего его средний максимум и ми-

³Методическая статья Лоури (O. N. Lowry) *Protein measurement with the Folin Phenol reagent* // *J. Biol. Chem.* 1951. 193–265 была признана самой цитируемой работой (300 тыс. ссылок).

нимум. Этот коэффициент отражает, как часто журнал цитирует и цитируется каждым из журналов, с которым он сравнивается, и учитывает размеры журналов (количество опубликованных статей), а также сколько раз каждый журнал цитирует другой. Тем не менее с использованием метода связанности JCR некоторые журналы могут быть отнесены к другим предметным категориям JCR. Рассмотрим это на примере журнала *Circulation*. Этот журнал имеет самый высокий импакт-фактор среди журналов в области кардиологии. Однако мы выявили, что журнал *New England Journal of Medicine* (NEJM), являющийся журналом по общим проблемам медицины, занимает седьмой ранг среди наиболее связанных с кардиологией журналов. До сих пор можно было только догадываться о близости NEJM к этой или другой тематике.

$$R_{1>2} = \frac{C_{1>2} \times 10^6}{\text{Ref}_1 \times \text{Pap}_2}$$

$$R_{1<2} = \frac{C_{1<2} \times 10^6}{\text{Ref}_2 \times \text{Pap}_1}$$

$$R_{\text{coeff}} = \sqrt{R_{1>2} \times R_{1<2}}$$

C — количество ссылок;

Ref₁ — количество публикаций, процитированных в журнале 1;

Ref₂ — количество публикаций, процитированных в журнале 2;

Pap₁ — количество публикаций, содержащихся в журнале 1;

Pap₂ — количество публикаций, содержащихся в журнале 2.

Рис. 3. Общая формула для расчета взаимосвязей цитируемости между двумя журналами и коэффициент взаимосвязи, отражающий средний максимум и минимум

Многие недостатки импакт-факторов журналов были устранены в другой базе данных ISI, называемой *Journal Performance Indicators* (JPI, <http://scientific.thomson.com/products/jpi>). Эта ежегодная компиляция охватывает период с 1981-го по текущий год. В отличие от JCR, эта база данных связывает каждую статью-источник с уникальными ссылками на нее, позволяя рассчитать импакт более точно. В знаменатель включаются только ссылки на оригинальные статьи, и это дает возможность получить кумулятивный импакт, охватывающий более длительный период. Для анализируемого периода с 1999 по 2004 г., в таблице 3 указано, что значение кумулятивного импакта для статей, опубликованных в JAMA в 1999 г., было равно 84,5. Данное значение получено путем деления 31 257 ссылок в период с 1999 по 2004 г. на 370 статей, опубликованных в 1999 г. В этом году в JAMA всего вышло 1905 публикаций, в том числе 630 писем и 253 редакционных текста, цитирования которых не включены в расчет импакта по JPI. Несущественность возможных искажений итогового значения импакта, вносимых неучетом ссылок на все редакционные материалы в SCI, показана в докладе González и Campanario, исследователей из Университета Алкала (Alcalá) [13].

Кроме библиотечкарей, использующих импакт-фактор журнала для принятия решения о том, какие научные журналы необходимо приобрести, он также используется авторами для выбора журнала, в который они представляют свои статьи для опубликования. Как правило, журналы с высоким импакт-фактором являются также и самыми престижными, хотя восприятие престижа является довольно туманным понятием. Специалисты по библиотечным наукам утверждают, что числитель

Таблица 3

Данные JPI по импакту цитирования журнала JAMA (все статьи-источники),
распределение по годам с 1981 по 2004 г.
(БД ISI, Journal Performance Indicators, массив 2004 г.)

$$\frac{31,257}{370} = \frac{\text{Citations received 1999–2004}}{\text{Articles published in JAMA in 1999}} = 84.5$$

Rank	Year	Impact	Citations	Papers
1	1981	29.57	16,291	551
2	1982	35.53	20,358	573
3	1983	40.11	22,219	554
4	1984	35.26	21,791	618
5	1985	35.05	18,436	526
6	1986	48.76	24,576	504
7	1987	44.70	26,688	597
8	1988	48.40	30,009	620
9	1989	55.79	34,979	627
10	1990	54.83	35,968	656
11	1991	47.19	30,389	644
12	1992	58.48	34,389	588
13	1993	65.55	38,349	585
14	1994	70.54	39,148	555
15	1995	81.99	45,094	550
16	1996	60.16	32,908	547
17	1997	58.19	32,821	564
18	1998	75.20	37,372	497
19	1999	84.48	31,257	370
20	2000	56.71	21,040	371
21	2001	49.98	18,842	377
22	2002	42.84	16,921	395
23	2003	19.09	7311	383
24	2004	3.34	1174	351

при расчете импакт-фактора сам по себе является более релевантным. Vensman утверждал, что само по себе количество ссылок, посчитанных за двухлетний период, — лучший показатель значимости и экономической эффективности журнала, чем импакт-фактор [14]. Импакт-фактор журнала также может быть полезным при сравнении ожидаемой и реальной частоты цитирования. Таким образом, когда ISI готовит персональный доклад о цитируемости, то предоставляются данные об ожидаемом импакте цитируемости не только для конкретного журнала, но и для конкретного года, потому что импакт-факторы могут меняться из года в год.

Использование импакт-фактора журнала вместо актуального числа цитирований самой статьи является довольно спорным вопросом. Поскольку статьи, опубликованные недавно, возможно, не имели достаточно времени, чтобы быть процитированными в работе, то появляется соблазн использовать импакт-фактор журнала в качестве суррогатного инструмента оценки. Предположительно само принятие документа для публикации в журнал с высоким импактом расценивается как показатель престижа. Как правило, когда рассматривается недавняя библиография работ автора, импакт-факторы журналов, в которых автор публиковался, используются вместо реального подсчета цитируемости. Эта практика началась около десяти лет назад, когда администраторы решили, что они могут оценить будущий импакт недавно опубликованной статьи по импакту журнала, в котором была опубликована статья. Это особенно актуально для молодых ученых, так как многие из их работ, перечисленных в C.V., опубликованы в течение периода времени, используемого для расчета импакта, и боль-

шинство из них не будут процитированы в течение нескольких ближайших лет или более, в зависимости от темпов развития предметной области, в которой они проводятся исследования.

Таким образом, импакт-фактор используется для оценки ожидаемого влияния отдельных статей, и это довольно сомнительная практика с учетом вышеупомянутых отклонений, наблюдаемых для большинства журналов. Сегодня «Webometrics» все чаще используется, хотя существует мало свидетельств, что это лучше, чем традиционный анализ цитируемости. «Веб-citation» может произойти немного раньше, но это не то же самое, что цитируемость. Следует различать чтение и выгрузку статей, а также и фактическое цитирование новых научно-исследовательских работ. Тем не менее некоторые исследования показывают, что веб-citation является предвестником будущего цитирования.

Предположение, что импакт недавних статей не может быть оценен по SCI, не является абсолютно правильным. Хотя задержка в несколько лет по некоторой тематике возможна, статьи, которые достигают высокого импакта, как правило, цитируются в течение нескольких месяцев с момента публикации и, конечно, в течение года или около того. Этот образец «оперативности» позволяет ISI идентифицировать «горячие публикации» (“hot papers”) раз в два месяца для публикации в Science Watch. Тем не менее полное подтверждение высокого импакта, как правило, следует через два года. The Scientist ждет до двух лет, чтобы выбрать «горячие статьи» для комментариев авторов, но большинство из этих работ в конечном счете продолжают цитироваться, чтобы стать классикой цитирования (<http://www.citationclassics.org>).

Резюмируя множество противоречивых мнений о факторах воздействия, К. Хоффел (Hoeffel C.) так описал ситуацию: «Импакт-фактор не является идеальным инструментом для измерения качества статей, но нет ничего лучше. ИФ имеет то преимущество, что он уже существует и, следовательно, является хорошим инструментом для научной оценки. Опыт показал, что в каждой специальности лучшие журналы это те, в которых труднее всего опубликоваться, и эти журналы имеют более высокий импакт-фактор. Большинство из этих журналов существовали задолго до того, как импакт-фактор был разработан. Использование импакт-фактора как показателя качества широко распространено, потому что это хорошо совпадает с мнени-

ем, которое есть у нас о лучших журналах в нашей специальности» [15].

Очевидно, что было бы лучше, если бы система оценки качества основывалась на прочтении каждой статьи, но даже тогда у экспертов возникли бы трудности с принятием общего решения по ее оценке. Когда речь идет о такой сфере как оценивание, большинство людей не имеет достаточно времени или не хочет его тратить на это. Даже если они провели оценку, на их суждения, конечно, будут влиять комментарии тех, кто цитировал работу, а это, как известно, называется контекстный анализ цитирования. К счастью, в ближайшем будущем, возможности полнотекстового доступа в Сети сделают эту практическую задачу более выполнимой.

1. Brodman E. Choosing physiology journals // Bull. Med. Library Association. 1944. Vol. 32. P. 479–483.
2. Garfield E. Citation analysis as a tool in journal evaluation // Science. 1972. Vol. 178. P. 471–479. [Reprinted in: Current Contents. 1973. № 6. P. 5–6; Essays of an Information Scientist. 1977. Vol. 1. P. 527–544.]
3. Lock S. P. Journalology: Are quotes needed? // CBE Views. 1989. Vol. 12. P. 57–59.
4. Pudovkin A. I., Garfield E. Rank-normalized Impact Factor: A way to compare journal performance across subject categories. Annual Meeting, Providence, RI. November 17, 2004 // Proceedings of the 67th Annual Meeting of the American Society for Information Science & Technology. 2004. Vol. 41. P. 507–515.
5. Garfield E. Long-term vs. short-term journal impact: Does it matter? (part II) // The Scientist. 1998. Vol. 12. P. 12–13.
6. Garfield E. Is the ration between number of citations and publications cited a true constant? // Current Contents. 1976. № 6. P. 5–7. [Reprinted in: Essays of an Information Scientist. 1977. Vol. 2. P. 419–421.]
7. Kresge N., Simoni R. D., Hill R. L. The most highly cited paper in publishing history: Protein determination by Oliver H. Lowry // J. Biol. Chem. 2005. Vol. 280. P. e25.
8. Lowry O. H. Protein measurement with folin phenol reagent // Current Contents / Life Sciences. 1977. Vol. 1. P. 5–7.
9. Yu G., Wang X.-Y., Yu D.-R. The influence of publication delays on impact factors // Scientometrics. 2005. Vol. 64 P. 235–246.
10. Garfield E. Long-term vs. short-term journal impact: Does it matter? // The Scientist. 1998. Vol. 12. P. 10–12.
11. Garfield E. Which medical journals have the greatest impact? // Annals of Internal Medicine. 1986. Vol. 105. № 2. P. 313–320 [Available in: <http://garfield.library.upenn.edu/essasys/v10p007y1987.pdf>]
12. Pudovkin A. I., Garfield E. Algorithmic procedure for finding semantically related journals // Journal of the American Society for Information Science & Technology (JAS-IST). 2002. Vol. 53. P. 1113–1119.
13. González L., Campanario M. Structure of the Impact Factor of journals included in the Social Sciences Citation Index: Citations from documents labeled “Editorial Material” // J. Am. Soc. Infor. Sci Technol. 2006. Vol. 58. P. 252–262.
14. Bensman S. J., Wilder S. J. Scientific and technical serials holdings optimization in an inefficient market: a LSU serials redesign project exercise // Library Resources & Technical Services. 1998. Vol. 42. P. 147–242.
15. Hoeffel C. Journal Impact Factors (letter) // Allergy. Vol. 53. P. 1225.

СПИСОК СОКРАЩЕНИЙ

ABBREVIATIONS

DOI 10.15826/B978-5-7996-1352-5.0008

ВИНИТИ – Всесоюзный институт научной и технической информации РАН.

ГРНТИ – Государственный рубрикатор научно-технической информации.

ИИЕТ АН СССР – Институт истории естествознания и техники АН СССР.

ИПКИР – Институт повышения квалификации информационных работников Госкомитета СССР по науке и технике.

ЮНЕСКО – Международная организация Организации Объединенных Наций по вопросам образования, науки и культуры.

НИОКР – научно-исследовательские и опытно-конструкторские разработки.

НТИ – научно-техническая информация.

РЖ – реферативные журналы.

РИНЦ – Российский индекс научного цитирования.

УЦЛ – указатель цитированной литературы.

ЦЕРН – Европейская организация по ядерным исследованиям (Conseil Européen pour la Recherche Nucléaire, сокращение названия на французском языке).

A&HCI – Arts & Humanities Citation Index (указатель цитируемой литературы по гуманитарным наукам и искусству).

ASCA – Automatic Subject Citation Alert (автоматизированная служба сигнального оповещения абонентов по интересующей их тематике, цитирующих или цитирующим авторам).

BKCI-S – Book Citation Index Science.

BKCI-SSH – Book Citation Index Social Sciences & Humanities.

CA – Chemical Abstracts (реферативный журнал, охватывающий все направления химии, биохимии и химической технологии).

CAS – Chemical Abstracts Service (химическая реферативная служба, подразделение Американского химического общества).

CC – Current Contents (еженедельные бюллетени сигнальной информации, содержащие оглавления научных журналов по различным направлениям науки).

CPCI-S – Conference Proceedings Citation Index Science.

CPCI-SSH – Conference Proceedings Citation Index Social Science & Humanities.

CWTS – Centre for Science and Technology Studies (Центр исследования науки и технологии Лейденского университета, сокращение названия на голландском языке).

ESI – Essential Science Indicators.

FTE – Full-Time Equivalent (среднесписочная численность персонала).

GCI – Genetic Citation Index (указатель цитируемой литературы по генетике).

IF – Impact Factor (импакт-фактор журнала).

ISI – Institute for Scientific Information (Институт научной информации).

ISSI – International Society for Scientometrics and Informetrics (Международное общество по наукометрии и информетрии).

JCR – Journal Citation Reports.

JNCI – Journal Normalized Citation Impact.

NASA – National Aeronautics and Space Administration (Национальное управление по воздухоплаванию и исследованию космического пространства США).

NCI – Normalized Citation Impact.

NIH – National Institutes of Health (Национальный институт здравоохранения США).

NSF – National Science Foundation (Национальный научный фонд США).

OECD – Organization for Economic Co-operation and Development (Организация по экономическому сотрудничеству и развитию).

PNAS – Proceedings of the National Academy of Sciences.

PPP\$ – Purchasing power parity US dollars (паритет покупательной способности в долларах США).

S&EI – Science & Engineering Indicators (Отчет NSF, публикуемый каждые два года, содержит библиометрическую статистику Web of Science CC).

SCI – Science Citation Index (указатель цитируемой литературы).

SCIE – Science Citation Index Expanded (указатель цитируемой литературы по естественным наукам и технике).

SDI – Selective Dissemination of Information (Служба избирательного распределения информации).

SJR – SCImago Journal Rank.

SNA – Social Network Analysis.

SNIP – Source Normalized Impact per Paper (цитируемость, нормализованная по источникам ссылок).

SSCI – Social Science Citation Index (указатель цитируемой литературы по общественным наукам).

TRL – Technology readiness level (уровни готовности технологии к использованию на практике).

Web of Science CC – Web of Science Core Collection.

WoS – Web of Science (информационная платформа компании Thomson Reuters, на которой размещены ее основные информационные продукты).

UCSD – University of California, San Diego.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

INDEX

DOI 10.15826/B978-5-7996-1352-5.0009

а-индекс	128	готовность технологии к использованию на практике – Technology readiness level (TRL)	64
h _a -индекс	104	дробный счет – fractional counting	80
λ-знания	56	Европейская ассоциация наукометрических институтов – European Network of Indicators Designers	45
Ω-знания	56	Еженедельный предметный индекс – Weekly Subject Index	201
агрегированный импакт-фактор – aggregate impact factor	90	журналология – Journalology	232
Американская ассоциация по развитию науки – American Association for the Advancement of Science (AAAS)	204	закон Парето	112, 125
Американское общество генетиков – American Society on Human Genetics (ASHG)	20	заметка редактора (тип документа) – Editorial	79
Американское химическое общество – American Chemical Society (ACS)	16, 122, 175	затраты валового национального продукта на одного человека – wealth intensity	42
аннотация статьи (реферат)	77	избирательное распределение информации – Selective Dissemination of Information (SDI)	27
Атлас науки – Atlas of Science	38	импакт-фактор журнала – journal impact factor (классический, двухлетний)	82
аффилиация	18, 77, 117, 168	импакт-фактор журнала, пятилетний	84
библиографические сведения	77	индекс Хирша	43, 50, 102, 127, 144, 151
ВИНИТИ	17, 27, 33, 174, 205	Индекс цитируемости по генетике – Genetic Citation Index (GCI)	21
время полужизни полученных ссылок – cited half-life	107	Институт истории естествознания и техники (ИИЕТ) АН СССР	30
время полужизни сделанных ссылок – citing half-life	107	Институт научной информации США – Institute for Scientific Information (ISI)	15, 201, 205
геоинформационные системы (ГИС)	169		
геокодирование	168		
гипотеза Пуанкаре	51		
горячие публикации – hot papers	37, 237		
государственный рубрикатор научно-технической информации (ГРНТИ)	131		

Институт повышения квалификации информационных работников (ИПКИР) Госкомитета СССР по науке и технике	35	модель векторного пространства – Vector space model (VSM)	173
Институт проблем развития науки РАН	62	научная статья (тип документа) – Article	79, 92, 98, 122, 156, 160
Казахстан	173	Научная электронная библиотека eLIBRARY.RU	117
каиновая кислота	61	научный обзор (тип документа) – Review	79, 92, 98, 122, 160
картирование – mapping	164	Национальный институт здравоохранения – National Institutes of Health (NIH)	20
классификация University of California, San Diego (UCSD)	177	Национальный научный фонд США – National Science Foundation (NSF)	15, 19, 38, 165
ключевые слова	77, 172	негативное цитирование	26, 120
книжная рецензия (тип документа) – Book Review	79	номенклатура ВАК	132
концепция гандикапа	119	нормативная теория цитирования	118
Корпус экспертов по естественным наукам	53	окно цитирования – citation window	82
ко-цитирование	29, 37, 71, 108, 121, 171, 176, 205, 225	Организация экономического сотрудничества и развития – Organization of Economic, Cooperation and Development (OECD)	41, 55, 62, 113
кошелек Миллера	165	относительная цитируемость разнородного потока публикаций – normalized citation impact, crown indicator	92
коэффициент нецитируемости	99	относительный импакт-фактор	90
Кто есть кто в российской науке	127	Офис по науке и технике – Office Science and Technology (OST)	42
Кто публикуется в науке – Who is Publishing in Science (WIPIS)	203	паритет покупательной способности в долларах США – Purchasing power parity US dollars (PPP\$)	65
Манхэттенский проект	72	патентный ландшафт	71, 173
Матфея, индекс	94, 120	письмо (тип документа) – letter	79
Матфея, эффект	73, 119	плотность цитирования	232
международная классификация болезней – International Statistical Classification of Diseases and Related Health Problems	175	полный счет – total counting, whole counting	80
Международное общество по наукометрии и информетрии – International Society on Scientometrics and Informetrics (ISSI)	25, 38, 232	публикационное окно – publication window	82
Международный математический союз – International Mathematical Union	44	ранговые индикаторы	96
метод библиографического сочетания – Bibliographic coupling	121, 176	реферативные журналы	16, 113, 116, 129, 175
МИФИ	153		

рефлексивная теория цитирования	119
решения при полной информации – informed decision	108
Российский индекс научного цитирования (РИНЦ)	124, 146
самоцитирование	26, 87, 101, 121, 157, 167
Северо-Западный научно-методический центр (СЗНМЦ)	62
совокупный импакт-фактор	86
социальная конструктивистская теория цитирования	119
– список Белла	115
– средневзвешенный импакт-фактор	86
Среднее цитирование статей по обла- стям знаний – Average Citation Rates for papers published by field	125
среднесписочная численность персона- ла Full-Time Equivalent (FTE)	66, 115
средняя частота цитирования, также см. цитируемость в расчете на одну статью	125
Статистический институт ЮНЕСКО	41
– тематика (рубрика)	77, 94, 129,
– тепловая карта – heatmap	171
– тип публикации	79
– указатели цитируемой литературы (УЦЛ)	36
Указатель ссылок Шепарда – Shepard's Citations	16, 20, 116
универсальная десятичная классифи- кация	130
фасетная классификация	130
Центр по изучению науки и техни- ки – Observatoire des Sciences et des Techniques (Франция)	39
ЦЕРН	148, 153
цитируемость в расчете на одну статью – cites per document	27, 31, 85, 152, 154, 206, 216, 223, 230

цитируемые авторы – cited authors	21, 205
цитируемые статьи – cited articles	21, 205
Academic Ranking of World Universities (ARWU, Шанхайский рейтинг)	43, 156, 161
ACM Transactions on Database Systems	150, 168
AGRIS	142
Altmetrics (альтметрики)	149, 155
Altmetrics for Institutions	151
Arts & Humanities Citation Index (A&HCI)	28, 38, 80, 112
article influence score (индекс влияния статьи)	44, 100
arXiv	115, 151
Automatic Subject Citation Alert (ASCA)	27
Baselines (expected citation rate, ожидае- мое число ссылок)	81, 93, 126
biases	57
Book Citation Index – Science	81
Book Citation Index – Social Sciences & Humanities	81
Centre for Science and Technology Studies (CWTS)	39, 43
Chemical Abstracts	175
Chinese Science Citation Database	43, 117
Citation Classics	36, 215, 218, 225
CitNetExplorer	170
COLLNET	38
Conference proceedings (тип документа)	124
Conference Proceedings Citation Index – Science	43, 81
Conference Proceedings Citation Index – Social Science & Humanities	43, 81
Current Contents	17, 21, 27, 35, 37, 134, 202, 227
Cyberinfrastructure for Network Science Center	165

Cybermetrics	44	Korean Citation Index (KCI)	43, 117
Derwent Innovations Index	117, 134, 144	Leiden Ranking	43, 160
Derwent World Patents Index	173	MERIL	72
DIALOG	43	Microsoft Academic Search	141, 144
DORA	45	NASA	50
Eigenfactor	44, 100	ORCID	143
Essays of an Information Scientist	23, 30	Organization Enhanced	18
Essential Science Indicators (ESI)	38, 42, 81, 85, 98, 134, 152	PageRank	45, 100, 123
Eugene Garfield Associates, Information Engineers	17	Pajek	171
Frascati Fields of Science	132	Percentiles for papers published by field	152
Geocoding API Google	168	Permuterm Subject Index	17
GIS-Lab	169	PLoS ONE	135, 141, 150
Google Scholar	141, 144, 155	Public Library of Science (PLOS)	114
g-индекс	105, 129	PubMed	115
HEEACT	161	QS World University Ranking	157, 161
Highly Cited Papers	98	Quacquarelli Symonds (QS)	157
HistCite	169, 229	R&D 100 Awards	180
HITS	45	references	23
immediacy index (индекс оперативности, или индекс немедленного цитирования)	85, 232	reprint author	154
Impact Factor Without Journal Self Cites	87	Research Excellence Framework	41
InCites	44, 81, 122, 152,	Research Front	37, 71, 81, 122, 224
Index Medicus	116	ResearcherID	18, 143
Individual h-index (original)	129	Royal Society Publishing	45
Individual h-index (PoP variation)	129	Russian Science Citation Index	117, 144
I-индекс	129	Sci ² Tool	165, 170
Journal Citation Reports (JCR)	28, 35, 43, 81, 90, 101, 135, 227	SciELO Citation Index	43, 117
Journal Normalized Citation Impact (JNCI)	94	Science and Engineering Indicators	45
KeyWords Plus	173	Science Citation Index (SCI)	15, 30, 38, 43, 80, 201, 227
		Science Index	142
		Science Indicators (в настоящее время Science & Engineering Indicators)	38
		ScienceWatch	136

SCImago	44, 159
SCImago Institutional Ranking	159
SCImago Journal Rank (SJR)	44, 123, 155, 160
Scopus	44, 124, 157, 177
Snowball Metrics	154
Social Network Analysis (SNA)	170
Social Sciences Citation Index — SSCI	28, 38, 43, 80, 100, 148
source article	21, 205
Source Normalized Impact per Paper (SNIP)	123, 155
super-cited	233

Thomson Reuters	28, 35, 40, 45, 76, 80, 84, 117, 157, 173
Times Higher Education World University Ranking (THEWUR)	43, 154, 156, 161
U.S. News	161
VOSviewer	175
Web of Science (WoS)	18, 26, 44, 76, 117, 120, 132, 152, 157, 228
Webometrics	44, 237
Welsh Machine Project	17, 29
Wordle	173

ИМЕННОЙ УКАЗАТЕЛЬ¹

PERSONS

DOI 10.15826/B978-5-7996-1352-5.0010

Абрикосов, А. А.	37	Васерман, Стенли (S. Wasserman)	171, 190
Адейр, У. (William Adair)	16	Васильев, Р. Ф.	125
Адамс, К.	30	Васьковский, В. Е.	28
Акоев, М. А.	40	Вернадский, Владимир Иванович	32
Алахвердян, А. А.	40	Веспиньяни, Алесандро (A. Vespignani)	171, 190
Аллен, Гордон (Gordon Allen)	20	Влахи, Я. (Y. Vlachu)	40
Алфимов, М. В.	36	Владуц, Г.	33
Атвуд, К. (K. C. Atwood)	22	Воутерс, П. (P. Wouters)	11, 32, 39
Барабаши, Альберт-Ласло (A.-L. Barabasi)	171, 190	Гарвей, У. (William (Bill) Garvey)	29
Баринова, З.	33	Гарфилд, Юджин (Eugene Garfield)	15, 19, 23, 41, 45, 116, 122, 169, 171, 190, 198, 199
Батлер, Л. (Linda Butler)	39	Гессен, Борис Михайлович	32
Бедфорд, Г. (G. Bedford)	20	Гильберт, Давид	57
Белинский, В. Г.	111	Гиляревский, Руджерио Сергеевич	35
Белл, Джеффри	115	Глушков, В. М.	34
Бергстром, Карл (Carl Bergstrom)	100	Глэнцел, В. (W. Glänzel)	39, 104
Берельсон, Б. (B. Berelson)	146	Горькавый, Н. Н.	50
Бернал, Дж. (John Desmond Bernal)	24, 32	Грановский, Ю. В.	10, 33, 35, 40
Бониц, М. (M. Bonitz)	120	Гриффит, Белвер (Belver C. Griffith)	25, 29
Браун, Тибор (Tibor Braun)	34, 40	Грэхам, Л. (L. R. Graham)	32
Браун, Эндрю	24	Добров, Г. М.	10, 34
Бреховских, Л. М.	37	Дьюи, М.	130
Броад, В. (W. J. Broad)	18	Зитт, М. (M. Zitt)	39
Брэдфорд, С. (Bradford, S.)	18	Ицковиц, Г. (H. Etzkowitz)	170, 181
Бухарин, Николай Иванович	32	Капица, А. П.	28
Буш, Ванневар (Vannevar Bush)	15		
Бэр, К. Э.	130		
Вайнберг, А. (Alvin Weinberg)	25, 41		
Валтман, Л. (L. Waltman)	104		

¹Из именного указателя исключены персоналии, упомянутые в переводах статей Ю. Гарфилда.

Каплан, Н. (N. Kaplan)	26	Ориент, И. М.	33
Катц, С. (Sylvan Katz)	39	Перельман, Г. Я.	51
Кедров, Бонифатий Михайлович	32	Перри, Дж. (J. Perry)	16
Кинг, Д. (D. King)	42	Писляков, В. В.	40, 120
Кожевников, В. Л.	52	Прайс, Дерек де Солла (Derek de Solla Price)	22, 24, 29, 40, 55, 58, 111, 118, 124
Коллинз, Рэндалл (Randall Collins)	58, 172	Пудовкин, А. И.	40
Коренной, А. А.	27	Раан, А. Ван (Antony (Ton) Van Raan)	39, 45
Кронин, Б. (B. Cronin)	31	Ракитов, А. И.	40
Латур, Бруно (Bruno Latour)	55	Ранганатан, Ш. Р.	131
Ле Пеир, Кейс (C. Le Pair)	39	Рид, Л. (L. Reed)	16
Ледерберг, Дж. (Joshua Lederberg)	15, 20, 26, 30	Роджер, Уильям Бартон	182
Лейдесдорф, Л. (L. Leidesdorff)	39	Роус, П. (Peyton Rous)	30
Либкинд, А. Н.	36, 40	Руссо, Р. (R. Rousseau)	87
Лотка, Альфред Джеймс (Alfred James Lotka)	170	Славинский, Б. В.	28
Лоури, Оливер Х. (Oliver H. Lowry)	36	Смолл, Г. (Henry Small)	25, 29, 38, 41
Лысенко, Т. Д.	26	Соколов, М. М.	53
Маркусова, В. А.	40, 199	Стяжкин, Н. И.	33
Мартин, Б. (B. Martin)	39	Сэлтон, Г. (G. Salton)	170
Маршакова, И. В.	38, 40	Терещенко, В. И.	34
Медведев, Дмитрий Анатольевич	35	Терман, Фредерик	182
Мертон, Р. К. (Robert K. Merton)	25, 30, 118	Уэллс, Г.	17
Мехтиев, А.	36	Хайтун, С. Д.	146
Микулинский, Семен Романович	32	Хамфри, Х. (Hubert Humphrey)	19
Миндели, Л. Э.	40	Хирш, Хорхе (J. Hirsch)	43, 50, 102, 127
Минин, В. А.	36	Цукерман, Г. (Harriet Zuckerman)	30
Михайлов, А. И.	33	Червон, Х. Ю.	40
Мокир, Джоэль (Joel Mokyr)	56	Чжан, Итан (Yitang Zhang)	52
Москалева, О. В.	40, 96	Шер, И. (Irving Sher)	28, 31
Муд, Хэнк (Henk Moed)	149	Шуберт, А. (A. Schubert)	40
Мульченко, З. М.	33, 125	Эйзенхауэр, Д. (Dwight Eisenhower)	19
Мэй, Р. (R. May)	41	Эйк, Н. Ван (N. J. Van Eck)	104, 170
Налимов, Василий Васильевич	31, 34, 125, 198	Ярошевский, Михаил Григорьевич	40
Николайсен, Ж. (Nicolaisen, J.)	119		
Ольденбург, С. Ф.	32		

Руководство обобщает и систематизирует знания о наукометрии для читателей, имеющих начальные представления и поставивших перед собой цель разобраться в предмете. Материал, представленный в книге, должен послужить для читателя основой для продолжения изучения практического применения методов наукометрии при оценке результативности и эффективности исследователей и научных коллективов. В книге представлена история возникновения идеи Юджина Гарфилда по использованию библиографических ссылок как средства научного поиска, борьбы за воплощение этой идеи в жизнь и создания универсального инструмента для поиска научной информации Science Citation Index. Рассмотрены возможности и ограничения наукометрии при принятии решений о целесообразности выделения ресурсов для поддержки научной работы и показана важность привлечения экспертов в предметной области к проведению оценок. В книге рассматриваются библиометрические индикаторы, оценивающие цитируемость журналов, авторов, научных коллективов, организаций и целых стран. Подчеркивается необходимость грамотной

и аккуратной трактовки наукометрических индикаторов при принятии административных решений, распределении грантов, осуществлении кадровой политики и т. д. Рассмотрена связь наукометрических показателей с природой научной коммуникации. В книге описаны альтернативные способы оценки публикаций и использование библиометрических показателей при построении рейтингов университетов и научных организаций. Дана характеристика методов визуального представления наукометрической информации. Рассмотрены ограничения в прогнозировании научно-технического развития и вопросы повышения качества существующих научных направлений и развития новых направлений в организации. В книгу включены три статьи Ю. Гарфилда, с любезного согласия автора впервые издаваемые на русском языке.

Руководство предназначено для научных работников, руководителей и администраторов вузов, научных организаций и проектов, библиотек и информационных центров, аспирантов и студентов соответствующих специальностей и всех интересующихся вопросами измерения и оценки развития науки и технологии.

ABSTRACT

The Russian Scientometric Handbook is designed to provide an overview of the field of scientometrics. The Handbook describes the history of creation of the breakthrough concept of citation indexing by Dr. Eugene Garfield, and development of the first multidisciplinary scholarly citation index, the Science Citation Index. Application of scientometric tools and methods in research management and resource allocation is discussed. Authors survey various scientometric indicators relevant to individual researchers, journals, research institutions and whole countries. Authors explore new types of indicators, such as altmetrics, relationship between scientometric indicators

and the nature of scientific communication, and various methods of visualizing scientometric information. Possibilities and limitations of various scientometric techniques are examined. Authors highlight the need for an informed and reasonable approach to the use of quantitative indicators for research assessment. The Handbook includes the first Russian translations of three articles by Dr. Eugene Garfield.

The Handbook is intended for use by researchers, science analysts, universities and research institutions administrators, libraries and information centers staff, graduate students, and the general public interested in scientometrics and research evaluation.

Научное издание

М. А. Акоев, В. А. Маркусова, О. В. Москалева, В. В. Писляков

РУКОВОДСТВО ПО НАУКОМЕТРИИ: ИНДИКАТОРЫ РАЗВИТИЯ НАУКИ И ТЕХНОЛОГИИ

Редактор М. А. Акоев Корректор
Н. В. Шевченко Дизайн обложки
С. Г. Слюсарев
Компьютерная верстка В. В. Таскаев

Подписано в печать 16.12.2014.
Бумага офсетная. Формат 70×90/16. Уч.-изд. л. 19,37. Усл. печ. л. 18,28
Заказ № 1756. Тираж 1600 экз.

Отпечатано в типографии ИПЦ УрФУ
620000, г. Екатеринбург, ул. Тургенева, 4
Тел./факс: +7 (343) 358-93-06
e-mail: press-urfu@mail.ru