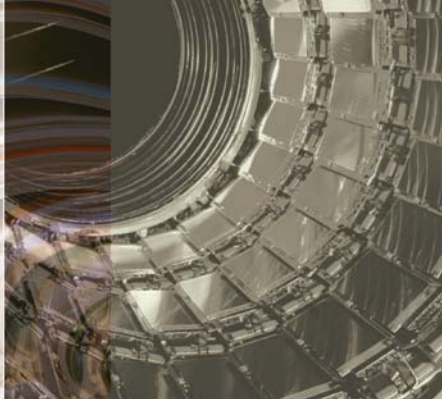


**Библиометрические индикаторы:  
практикум**

**В. В. Писляков**

**Национальный фонд подготовки кадров**



Результаты научной деятельности:  
ПОЛИТИКА. ОЦЕНКА. ВНЕДРЕНИЕ

---



**ИНФРА-М**



В. В. ПИСЛЯКОВ

**БИБЛИОМЕТРИЧЕСКИЕ ИНДИКАТОРЫ: ПРАКТИКУМ**

Москва  
ИНФРА-М, 2014

**Писляков, В.В.**  
ПЗ4 Библиометрические индикаторы: практикум [Текст] / В.В. Писляков; Редактор серии М.Ю. Барышникова; Ответственный редактор П.Г. Арефьев; Национальный фонд подготовки кадров. – М.: НФПК, Инфра-М, 2014. - 60 с. – (Серия «Результаты научной деятельности: Политика. Оценка. Внедрение»).

ISBN 978-5-16-010696-0 (print)  
ISBN 978-5-16-102714-1 (online)

В практикуме рассматривается ряд библиометрических индикаторов, связанных с цитируемостью научных публикаций: импакт-фактор (двухлетний, пятилетний, совокупный, средневзвешенный, относительный), средняя цитируемость статей организации/автора, средняя нормализованная цитируемость (с нормализацией по областям науки и по журналам), SNIP, SJR, Eigenfactor, индекс Хирша и др. Для каждого показателя дается определение и демонстрируется его расчет или поиск готового значения по базам данных (Web of Science, Scopus). Цель практикума – в доступной форме представить алгоритм расчета индикаторов и, что еще более важно, донести до читателя их «физический смысл». Демонстрируются особенности различных показателей, типичные ошибки при их расчете, даются практические рекомендации к их применению и корректной интерпретации. Значительное место занимают и подробные «кнопочные» указания — как в той или иной базе данных найти или самостоятельно рассчитать соответствующий индикатор

Пособие адресовано всем, кто профессионально занимается наукой, публикует результаты своей деятельности и оценивает научную результативность, – российским преподавателям и научным сотрудникам, специалистам в области управления наукой, учета и анализа научной деятельности, а также аспирантам.

*Писляков В.В. – заместитель директора библиотеки, Национальный исследовательский университет «Высшая школа экономики»;  
pislyakov@hse.ru.*

ББК 78.606

ISBN 978-5-16-010696-0 (print)  
ISBN 978-5-16-102714-1 (online)

© Писляков В.В., 2014.  
© НФПК, 2014.  
© Инфра-М, 2014.  
© Оформление и дизайн, О.А. Иванов, 2014.

# Содержание

1. Введение	7
2. Библиометрические инструменты Базы данных Web of Science и Scopus	8
2.1. Принципы организации библиометрических баз данных; методологические замечания	8
2.2. База данных Web of Science и аналитические надстройки	12
2.3. База данных Scopus	13
3. Индикаторы влияния статей (импакт-индикаторы)	14
3.1. Импакт-фактор журнала	14
3.2. Пятилетний импакт-фактор журнала, индекс оперативности	17
3.3. Аналоги журнального импакт-фактора в Scopus	18
3.4. Средняя цитируемость статей автора/организации	19
3.5. Совокупный и средневзвешенный импакт-фактор	22
4. Индикаторы относительной влияния статей (относительные, нормализованные импакт-индикаторы)	25
4.1. Относительный импакт-фактор, относительная цитируемость	25
4.2. Относительная цитируемость разнородного потока публикаций («crown indicator»)	30
4.3. Нормализация по источникам ссылок и индикатор SNIP	40
4.4. Ранговые индикаторы	45
5. «Взвешенные» индикаторы	46
5.1. SCImago Journal Rank (SJR)	47
5.2. Собственный фактор (Eigenfactor) и индекс влияния статьи (Article Influence)	48
6. Индекс Хирша (h-index)	49
7. Список литературы	54

## Сокращения:

---

ESI — Essential Science Indicators

JCR — Journal Citation Reports

SNIP — Source normalized impact per paper

WoS — Web of Science

КПС — коэффициент публикационной стратегии

ЦНЖ — цитируемость, нормализованная по журналам

ЦНОН — цитируемость, нормализованная по областям науки

## 1. Введение

В настоящем пособии рассматривается ряд библиометрических индикаторов, связанных с цитируемостью научных публикаций. Дается определение и демонстрируется расчет таких показателей, как импакт-фактор (двухлетний, пятилетний, совокупный, средневзвешенный, относительный), средняя цитируемость статей организации/автора, средняя нормализованная цитируемость (с нормализацией по областям науки и по журналам), SNIP, индекс Хирша и др.

Мы ставили себе целью в доступной форме представить алгоритм расчета индикаторов и, что еще более важно, донести до читателя их «физический смысл». Продемонстрировать особенности различных показателей, типичные ошибки при их расчете, дать практические рекомендации к их применению и корректной интерпретации. Значительное место занимают и подробные «кнопочные» указания — как в той или иной базе данных найти или самостоятельно рассчитать соответствующий индикатор.

### Примечание

Мы отдаем себе отчет в том, что интерфейс любой современной информационной системы претерпевает косметические или даже более серьезные изменения 1–2 раза в год, из-за чего вскоре данная составляющая нашего пособия неизбежно окажется устаревшей. Более того, базы данных пополняются новой информацией и, иногда, новыми индексруемыми источниками, что приводит к изменению чисел, фигурирующих в заданиях. Тем не менее ряд примеров по работе с ресурсами может преодолеть будущие изменения в онлайн-интерфейсах, а кроме того мы надеемся, что аккуратный читатель сумеет, имея под рукой подробные инструкции, сообразить, как надо выполнить то или иное действие в новом, изменившемся контексте.

Упомянем о некоторых соглашениях, которых мы собираемся придерживаться на протяжении данного текста. Во-первых, мы оставляем за рамками пособия вопрос о степени отражения научной значимости той или иной работы полученными ею библиографическими ссылками.



Эти споры ведутся не одно десятилетие, у них свое дискуссионное пространство, на которое не хотелось бы вторгаться в рамках данного, ставящего более технические вопросы, пособия. Если говорить несколько заостряя, в настоящем издании понятия «цитируемость» и «научный уровень» будут считаться синонимами.

Во-вторых, английское «citation», как справедливо указывалось, например, А. В. Полетаевым, означает не «цитирование» или тем более «цитату», а лишь библиографическую ссылку. Строго говоря, 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 настоящего издания «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.

## 2. Библиометрические инструменты. Базы данных Web of Science и Scopus

---

В настоящем пособии будут использоваться две международные библиометрические базы данных — Web of Science (WoS) и Scopus. Обе они *подписные* — они не размещены в открытом доступе, и для доступа к ним необходима платная подписка той организации, к которой вы относитесь (работаете, учитесь).

### **2.1. Принципы организации библиометрических баз данных; методологические замечания**

Что включает в себя «библиометрическая база данных» (или, как еще говорят, «база данных (научного) цитирования»), из чего она состоит и что именно в ней можно найти? Основной контент библиометрических баз

данных (как WoS, так и Scopus) — это научные журналы. При этом в базе данных *не содержится полного текста статей* этих журналов. О каждой статье хранится (и может быть выдана пользователю) следующая информация:

— 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 имелась в исходном тексте публикации; базы данных не составляют самостоятельно аннотации тех произведений, в которых рефераты отсутствуют;

— ключевые слова; иногда это несколько наборов ключевых слов — например, слова, приписанные статье в оригинальном тексте публикации («авторские ключевые слова»), и ключевые слова, «назначенные» базой данных исходя из внутреннего тезауруса ключевых слов и автоматических алгоритмов;

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

— организации, в которых работают авторы, т. н. «аффилиации», с почтовыми 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 (ISSN, язык оригинального документа, название и адрес издательства и др.), которые не важны для круга рассматриваемых нами вопросов.

Итак, если мы говорим, что «журнал расписывается (индексируется) библиометрической базой данных», это означает, что о каждой публикации данного журнала мы сможем узнать из базы данных сведения, перечисленные выше. Полного текста статей при этом, повторим, в базе не будет.

Помимо журнального контента, обе ведущие международные библиометрические базы включают также некоторое количество трудов конференций и, с недавнего времени, книг. При этом содержание полей, попадающих в базу данных, аналогично тому, что мы указали для журналов. Только в случае трудов конференций структурной единицей будет не статья, а доклад на конференции; в случае книг — глава из книги.

### Примечание

В случае сборников трудов конференций речь идет именно о «текстах докладов», а не о 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.

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 настоящем пособии, тематическая рубрика *статьи* определяется 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.

Кроме дисциплины каждой публикации в библиометрических базах данных присваивается *тип документа*. Это может быть научная статья (Article), научный обзор (Review), письмо (Letter), книжная рецензия (Book Review) и т. д. При расчете ряда библиометрических индикаторов могут учитываться лишь некоторые типы публикаций — чаще всего это Article и Review. Более 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 в данном пособии 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»), читатель с легкостью поймет это различие из контекста. Кроме того, для обозначения типа документа «научная статья» мы часто будем прибегать к английскому «Article», это дополнительно избавит от двусмысленностей.

В библиометрии статью считают публикацией той или иной лаборатории/организации/страны на основании институциональных адресов, которые указаны авторами в статье и, соответственно, перене-

сены в поле «affiliation» (или «address») базы данных. В данном случае при анализе не интересуются деталями авторской биографии. Аргументы «на самом деле этот ученый приехал из другой страны», «данный автор работал во время написания статьи еще в одной организации», «рукопись долго рассматривали и на момент публикации сотрудник уже давно уволился» не принимаются во внимание в ходе стандартного библиометрического исследования. «То, что указано в статье» — это окончательный ответ на вопрос, в какой организации работал автор, ее написавший.

### Примечание

Строго говоря, то же самое верно и для авторства. На самом деле мы не можем знать, действительно автор участвовал в написании статьи — или был добавлен в список авторов из почтения, в результате административного давления, по взаимному соглашению и т. д. Для библиометрии он — автор статьи. Добавим также, что автор, конечно, может указывать в статьях сразу несколько мест работы, и именно в таком виде 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 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.

### Внимание!

В рамках настоящего пособия будет использоваться только метод whole counting. Его отличает простота, однако следует помнить, что он не обеспечивает аддитивность большинства показателей. Например, если мы используем этот способ счета, то, отвечая на вопрос «сколько всего статей написали авторы организации», нельзя приводить арифметическую сумму числа статей, написанных каждым автором, — ввиду возможного (и вероятного) соавторства ученых организации при создании некоторых публикаций.

## 2.2. База данных Web of Science и аналитические надстройки

Web of Science (производитель: компания Thomson Reuters) охватывает более 12000 научных журналов, также содержит труды конференций и незначительное на данный момент количество книг. Вход в WoS:

<http://isiknowledge.com/wos>

WoS является ядром библиометрических продуктов Thomson Reuters, их главной частью. Помимо WoS также доступен (по подписке) ряд аналитических инструментов, в частности базы данных Journal Citation Reports (JCR) и Essential Science Indicators (ESI). Все эти продукты используют WoS как свою основу и лишь собирают, агрегируют из WoS данные по различным «информационным единицам» (журналам, ученым 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) и др. Вход в JCR:

<http://isiknowledge.com/jcr>

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 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 цити-

руемых среди тех, которые вышли в заданном году в заданной научной области. Вход в ESI:

<http://isiknowledge.com/esi>

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

### 2.3. База данных Scopus

Scopus (производитель: компания Elsevier) охватывает более 20000 научных и отраслевых журналов, также содержит труды конференций, книжные серии, в настоящее время добавляются монографии. Вход в Scopus:

<http://scopus.com>

Элементарные аналитические функции, связанные с журнальными библиометрическими индикаторами, реализованы внутри Scopus. Но основной блок аналитики делегирован компанией Elsevier порталу SCImago Journal & Country Rank, поддерживаемому исследовательской группой SCImago из университета Гранады. Адрес портала:

<http://scimagojr.com/>

Здесь все индикаторы публикуются в открытом доступе.

#### **Внимание!**

При определении *любого* библиометрического индикатора обязательным условием является указание базы данных, по которой ведется (должен вестись) расчет. Это неременное условие. Любой показатель, рассчитанный с использованием одной и той же методики, но по разным библиометрическим базам данных, будет принимать разные значения (а возможное совпадение будет случайным).

Теперь мы готовы к тому, чтобы перейти к подробному рассмотрению библиометрических индикаторов.

### 3. 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 вузе  $U$  за пятилетний промежуток 2008–2012 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 нашего импакт-индикатора. Например, это могут быть статьи, вышедшие в 2012 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 2012 г. выхода, на работы, опубликованные в 2008–2012 гг. учеными, работающими в вузе  $U$ . «Среднее» обозначает здесь усреднение по числу статей: найденное абсолютное число ссылок должно быть поделено на число статей в публикационном окне (т. е. на число статей, опубликованных учеными вуза  $U$  в 2008–2012 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.

#### 3.1. 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  $Y$  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.

### Пример

Рассчитаем импакт-фактор 2012 г. для журнала «Успехи химии» (англоязычная версия выходит под заглавием «Russian Chemical Reviews»).

В 2012 г. во всех журналах, охваченных WoS, было сделано 89 ссылок на статьи журнала «Успехи химии», опубликованные в 2011 г., и 157 ссылок на его статьи, опубликованные в 2010 г. Итого статьи «Успехов», вышедшие в 2010–2011 гг., получили в 2012 г.  $157+89=246$  цитирований. При этом в 2011 г. журнал опубликовал 54 статьи, в 2010 г. — 53, итого за промежуток 2010–2011 гг. («публикационное окно») вышло 107 статей.

Таким образом, импакт-фактор журнала за 2012 г. равен отношению  $246/107=2,30$ .

### Где найти?

Импакт-факторы публикуются ежегодно в базе данных JCR. В зависимости от подписки, у вас может быть доступ к данным за разное количество лет. Данные за очередной год обычно появляются в июне следующего года (т.е. импакт-факторы 2013 г. можно узнать в июне 2014 г. и т. д.). Надо иметь в виду, что JCR не публикует импакт-факторы для гуманитарных журналов.

### Пример

Найдем импакт-факторы 2012 г. журналов по экономике и упорядочим издания в порядке убывания импактов.

Зайти в **JCR** → в секции **Select a JCR edition and year** выбрать **JCR Social Sciences Edition** и в выпадающем меню год **2012** → в секции **Select an option** выбрать **View a group of journals by** и в выпадающем меню опцию **Subject Category** (эти позиции могут стоять по умолчанию) → кнопка **SUBMIT** → в секции **1) Select one or more categories from the list** выбрать в списке рубрику **ECONOMICS** → кнопка **SUBMIT** → вверху в меню **Sorted by** выбрать **Impact Factor** → кнопка **SORT AGAIN**



Получим таблицу, показанную на рис. 1. Вторая колонка содержит сокращенные названия журналов. Импакт-факторы журналов выделены цветом (как столбец, по которому проводится сортировка).

### Внимание!

Нередко издательства размещают текущие значения импакт-факторов своих журналов на их веб-страницах, в открытом доступе. Например: [http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)1532-2890](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1532-2890)

Rank	Abbreviated Journal Title <i>(linked to journal information)</i>	ISSN	Total Cites	Impact Factor
1	<a href="#">J ECON LIT</a>	0022-0515	5012	6.667
2	<a href="#">Q J ECON</a>	0033-5533	15000	5.278
3	<a href="#">J FINANC</a>	0022-1082	18729	4.333
4	<a href="#">J ACCOUNT ECON</a>	0165-4101	3494	3.912
5	<a href="#">ECONOMETRICA</a>	0012-9682	21481	3.823
6	<a href="#">BROOKINGS PAP ECO AC</a>	0007-2303	1776	3.680
7	<a href="#">AM ECON J-APPL ECON</a>	1945-7782	481	3.539
8	<a href="#">J ECON PERSPECT</a>	0895-3309	6047	3.489
9	<a href="#">J POLIT ECON</a>	0022-3808	15801	3.483
10	<a href="#">J FINANC ECON</a>	0304-405X	13075	3.424

Рис. 1. Журналы по экономике, в порядке убывания импакт-фактора. Скриншот интерфейса базы данных JCR (Thomson Reuters).

### Примечание

Обратите внимание, что на импакт-фактор также оказывают влияние и ссылки, полученные журналом из самого себя. В своем стандартном определении импакт-фактор включает самоцитирование журнала. Для того чтобы посмотреть, как изменился бы импакт, если бы самоцитирования не учитывались, необходимо открыть подробную информацию о журнале (получив таблицу рис. 1, нажать на название интересующего журнала). Там, в разделе Journal Self Cites, будет указан импакт-фактор без учета самоцитирования — Impact Factor without Self Cites.

### 3.2. Пятилетний импакт-фактор журнала, индекс оперативности

Импакт-фактор 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.

В базе данных JCR 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.

#### Где найти?

С 2007 г. пятилетние импакт-факторы публикуются ежегодно в базе данных JCR. Охват данных и метод их поиска те же, что и для импакт-фактора (колонка со значениями называется «5-Year Impact Factor»). Для журналов, недавно включенных в WoS, пятилетний импакт-фактор появится на несколько лет позже обычного двухлетнего (обычно на три года).

Заметим, 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 может иметь тенденцию к увеличению, по причине все более частого размещения препринтов в Интернете, открытия специальных разделов на сайтах издательств, где публикуются предварительные версии статей, принятых в печать, а также общего ускорения производственного 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 однолетнее публикационное окно и однолетнее окно цитирования, причём 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 быстро воспринимает его тексты и использует их при воспроизводстве научного знания.

#### **Где найти?**

Индекс оперативности публикуется ежегодно в базе данных JCR. Охват данных и метод их поиска те же, что и для импакт-фактора (колонка со значениями называется «*Immediacy Index*»).

### **3.3. Аналоги журнального импакт-фактора в Scopus**

По базе данных Scopus также вычисляется показатель, методика подсчета которого аналогична определению импакт-фактора. Сам термин «импакт-фактор» при этом не используется (ввиду зарегистрированных прав на него), индикатор назван более конкретно, «*Cites per Document (2 years)*». Кроме того, рассчитываются аналоги импакт-фактора с трех- и четырехлетним публикационным окном («*Cites per Doc (3 years)*»), («*Cites per Doc (4 years)*»). Окно цитирования при этом остается постоянным — один (отчетный) год.

#### **Где найти?**

Показатель публикуется на портале SCImago Journal & Country Rank (см. п. 2.3).

## Пример

Найдем индикатор Cites per Document (2 years) для журнала «Петрология» в 2011 г.

Зайти на портал **SCImago Journal & Country Rank** → в меню слева выбрать **Journal Search** → в поисковое окно ввести **Petrology** → проверить, что в выпадающем меню **in** стоит **Journal Title** → кнопка **Search** → в результатах поиска найдите позицию **Petrology. Russian Federation**, нажмите на эту ссылку

Можно прокрутить экран до диаграммы «Cites per Document in 2, 3 and 4 years windows» и посмотреть графическую зависимость этих показателей (рис. 2), а можно нажать вкладку «Data» справа от первой (самой верхней) диаграммы и получить численные значения всех индикаторов журнала. Искомый двухлетний показатель в 2011 г. равен 0,87.

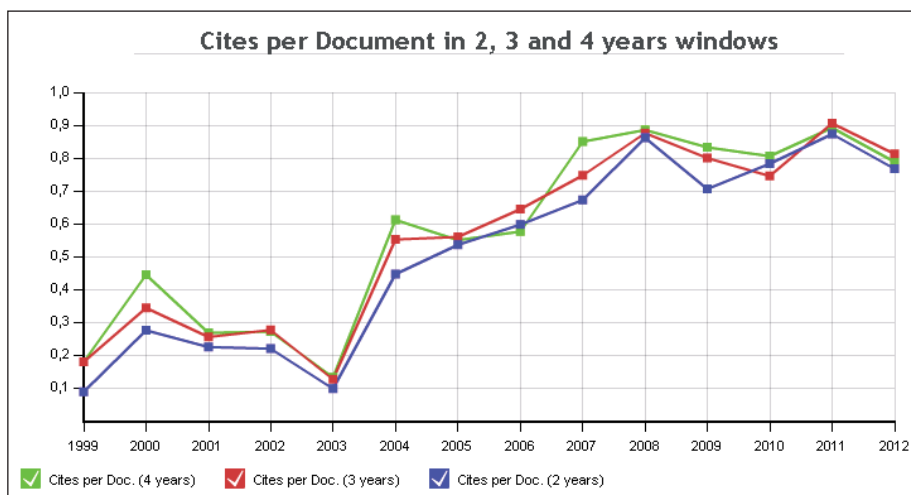


Рис. 2. Средняя цитируемость статей журнала «Петрология» при двух-, трех- и четырехлетнем публикационном окне. Скриншот интерфейса портала SCImago Journal & Country Rank.

### 3.4. Средняя цитируемость статей автора/организации

Наряду с импакт-фактором, рассчитываемым для журналов, можно ввести аналогичный по сути показатель для оценки средней влиятельности статей отдельного ученого или целой организации. 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). Заметим, что для бурно развивающихся областей, где знание быстро устаревает (таких как нанотехнологии, онкология, прикладная физика), можно выбрать интервал поменьше; для областей где старение знания происходит особенно долго (математика, зоология, социология), имеет смысл увеличить оба окна.

### Где найти?

Средняя цитируемость авторов и организаций, наряду с числом опубликованных статей и числом полученных ими цитирований, публикуются в базе данных ESI. Информация дается по последним десяти годам, плюс прошедшая часть текущего года (с задержкой в 2 месяца), с возможностью просмотра отдельно пятилетних интервалов. Обновление — раз в 2 месяца.

Однако в готовом виде показатель можно посмотреть лишь для небольшого числа отечественных ученых/организаций: база ESI, как упоминалось в п. 2.2, включает в себя только тех авторов и те организации, которые хотя бы в одной из научных областей попали в 1% наиболее цитируемых. Поэтому, например, организация, английское название которой начинается с «Moscow», в ESI на данный момент всего одна (это МИФИ — у МГУ принятый в ESI английский вариант начинается с «Lomonosov»). В связи с этим при анализе российской науки чаще всего приходится считать данный показатель вручную.

### Пример

Рассчитаем по базе Web of Science среднюю цитируемость статей, написанных в Тверском государственном техническом университете, учитывая ссылки, полученные в 2008–2012 гг. статьями сотрудников университета, вышедшими за тот же промежуток времени, в 2008–2012 гг.

Зайти в БД **Web of Science** → вверху около надписи **Basic Search** нажать стрелку и выбрать **Advanced Search** → проверить, что в двух меню настроек **Restrict results by languages and document types** выбрано **All languages** и **All document types**, а ниже в меню **TIMESPAN** выбрано **All years** → раскрыть меню **MORE SETTINGS** и в списке **Web of Science Core Collection: Citation Indexes** оставить

галочки только у журнальных баз **Science Citation Index Expanded (SCI-EXPANDED)**; **Social Sciences Citation Index (SSCI)**; **Arts & Humanities Citation Index (A&HCI)** → в поисковое окно ввести запрос **(AD=Tver) AND (PY=2008-2012)** (т. к. англоязычных вариантов написания названия вуза может быть несколько, будем отталкиваться в поиске от названия города) → кнопка **Search** → найдено 415 статей, надо перейти в поисковые результаты, нажав ссылку с числом результатов (**415**) в нижней части страницы **Search History** → для выбора статей, написанных в ТвГТУ, нажать справа над результатами поиска ссылку **Analyze Results** → в разделе **Rank the records by this field** выбрать **Organizations**, в разделе **Set display options** выбрать **Show the top 500 Results**, а **Minimum record count (threshold)** установить на **0** → кнопка **Analyze** → в полученном списке ТвГТУ соответствуют следующие позиции (их необходимо отметить галочками, см. рис. 3):

**TVER STATE TECH UNIV**  
**TVER TECH UNIV**  
**TVER STATE ENGN UNIV**  
**TVER STATE TECHNOL UNIV**  
**TVER STATE TECH UNIV TVER RUSSIA**

→ кнопка **View Records** вверху слева над списком организаций → найдено 48 статей, для анализа их цитируемости нажать справа на экране ссылку **Create Citation Report**

**Results Analysis**  
[<<Back to previous page](#)

415 records. (AD=Tver) AND (PY=2008-2012)

Rank the records by this field:	Set display options:	Sort by:
<input type="text" value="Grant Numbers"/> <input type="text" value="Group Authors"/> <input type="text" value="Languages"/> <input checked="" type="text" value="Organizations"/>	Show the top <input type="text" value="500"/> Results. Minimum record count (threshold): <input type="text" value="0"/>	<input checked="" type="radio"/> Record count <input type="radio"/> Selected field

---

Use the checkboxes below to view the records. You can choose to view those selected records, or you can exclude them (and view the others).

	Field: Organizations	Record Count	% of 415	Bar Chart	
<input type="checkbox"/>	TVER STATE UNIV	226	54.458 %	<div style="width: 54.458%;"></div>	<input type="checkbox"/>
<input type="checkbox"/>	TVER STATE MED ACAD	71	17.108 %	<div style="width: 17.108%;"></div>	<input type="checkbox"/>
<input type="checkbox"/>	RUSSIAN ACAD SCI	66	15.904 %	<div style="width: 15.904%;"></div>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	TVER STATE TECH UNIV	35	8.434 %	<div style="width: 8.434%;"></div>	<input checked="" type="checkbox"/>
<input type="checkbox"/>	MOSCOW MV LOMONOSOV STATE UNIV	28	6.747 %	<div style="width: 6.747%;"></div>	<input type="checkbox"/>
<input type="checkbox"/>	UNIV ULM	13	3.133 %	<div style="width: 3.133%;"></div>	<input type="checkbox"/>
<input type="checkbox"/>	INDIANA UNIV	10	2.410 %	<div style="width: 2.410%;"></div>	<input type="checkbox"/>
<input type="checkbox"/>	SCI RES INST SYNTHET FIBERS	8	1.928 %	<div style="width: 1.928%;"></div>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	TVER TECH UNIV	8	1.928 %	<div style="width: 1.928%;"></div>	<input checked="" type="checkbox"/>
<input type="checkbox"/>	TVER MED ACAD	7	1.687 %	<div style="width: 1.687%;"></div>	<input type="checkbox"/>
<input type="checkbox"/>	STATE MED ACAD	6	1.446 %	<div style="width: 1.446%;"></div>	<input type="checkbox"/>

Data rows displayed in table  
 All data rows (up to 200,000)

Рис. 3. Распределение статей по организациям. Скриншот интерфейса базы данных WoS (Thomson Reuters).

Получен отчет по цитируемости статей ТвГТУ, вышедших в 2008–2012 гг. В правой части экрана WoS сам считает показатель Average Citations per Item, однако нас он не устраивает, поскольку учитывает все ссылки, полученные статьями, в то время как наша цель — найти число ссылок, полученных в 2008–2012 гг. В начале таблицы со списком статей (прокрутите экран ниже графиков) считается суммарное число цитирований, полученных статьями в каждом году (2010, 2011 и т. д.). Нам необходимо сложить показатели 2008–2012 гг., для чего один раз придется нажать синий треугольничек под цифрой 2010, показывающий влево, чтобы увидеть показатели за 2008–2009 гг. Текущие данные по годам таковы: в 2008 г. исследуемые статьи не получили ссылок, в 2009 г. — 2 ссылки, в 2010 г. — 18 ссылок, в 2011 г. — 32 ссылки, в 2012 г. — 24 ссылки. Итого 48 статей ТвГТУ получили за 2008–2012 гг. 76 ссылок. Таким образом, средняя цитируемость статьи равна  $76 / 48 = 1,58$ .

Аналогичным образом можно рассчитывать среднюю цитируемость любого набора статей — от отдельного автора до целой страны или какой-либо научной области в целом.

### **3.5. Совокупный и средневзвешенный импакт-фактор**

Как следует из определения, импакт-фактор является не характеристикой журнала как целого, а показателем средней влиятельности, 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 влиятельность, а импакт-фактор издания, в котором она опубликована, делает это лишь косвенно и недостоверно. Поэтому в п. 3.4 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. 3.4 мы брали 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» (иногда его называют «совокупный импакт-фактор (статей) организации/ученого», что неверно: этим выражением скорее можно называть индикатор, введенный в п. 3.4, если понимать «импакт» в обобщенном смысле). Считается 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.

### Пример

Рассчитаем совокупный и средневзвешенный импакт-фактор журналов, в которых публиковались авторы из ТвГТУ в 2010 г.

Получить, как показано в предыдущем примере (п. 3.4), список всех статей ТвГТУ за 2008–2012 гг. Далее открыть пункт **Publication Years**



слева в фасеточном меню **Refine Results** → поставить галочку у 2010 г. → кнопка **Refine** → найдено 10 статей за 2010 г., открыть пункт **Source Titles** слева в фасеточном меню **Refine Results** → нажать ссылку **more options / values** внизу этого раздела меню

Получен список всех журналов (8 штук), в которых публиковались ученые ТвГТУ в 2010 г. Далее в базе данных JCR можно найти их импакт-факторы за 2010 г. (в принципе, можно было бы брать импакт-факторы другого года, это зависит от поставленной задачи):

Fibre Chemistry — 2 статьи, IF=0,269  
Journal of Friction and Wear — 2 статьи, IF=0,204  
Applied Catalysis B Environmental — 1 статья, IF=4,749  
Ecological Indicators — 1 статья, IF=2,967  
Ferroelectrics — 1 статья, IF=0,512  
International Journal of Heat and Fluid Flow — 1 статья, IF=1,802  
Physics of the Solid State — 1 статья, IF=0,727  
Polymer Science Series B — 1 статья, IF=0,421

Для нахождения совокупного импакт-фактора надо для каждого журнала взять произведение числа статей в нем на импакт, после чего сложить результаты по всем журналам. Итог:

$$2 \cdot 0,269 + 2 \cdot 0,204 + 4,749 + 2,967 + 0,512 + 1,802 + 0,727 + 0,421 = 12,124$$

Обратите внимание, что более трети вклада в совокупный импакт-фактор принесла одна-единственная статья из высокоцитируемого журнала «Applied Catalysis B».

Средневзвешенный импакт-фактор журналов, в которых публиковались ученые ТвГТУ в 2010 г., равен отношению совокупного импакт-фактора к числу статей 2010 г., которых вышло 10. В итоге средневзвешенный импакт равен 1,212.

### Примечание

Можно заметить, что лишь 3 из 10 исследуемых статей опубликованы в журналах с импактом выше найденного нами средневзвешенного. Это обычная ситуация при рассмотрении распределений цитирования, которые чаще всего являются сильно асимметричными: много статей/журналов получают по небольшому количеству ссылок, и лишь небольшая фракция имеет высокую цитируемость. Часто в таких случаях более корректным показателем является медианный импакт-фактор найденных журналов, вместо среднего (средневзвешенного) импакт-фактора. В нашем случае он был бы равен 0,467

## 4. 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. Самое главное при этом то, что нас не интересовало, к какой научной области относятся оцениваемые статьи. Однако сравнение эффективности академической деятельности представителей различных наук невозможно осуществить с помощью абсолютных, «простых» индикаторов.

### Внимание!

Центральным фактом, который необходимо иметь в виду при оценке эффективности деятельности ученых или организаций, является сильная зависимость абсолютных библиометрических показателей от научной дисциплины, в которой публикуются изучаемый автор/организация.

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.

### 4.1. 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 дисципли-

не 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  $Y-2$ . Именно этот «агрегированный» («собранный») импакт-фактор дисциплины берется в качестве знаменателя, именно на него делится импакт каждого журнала, чтобы получился относительный импакт-фактор.

### Где найти?

Данные по относительному импакт-фактору нигде не публикуются в готовом виде. Однако значение абсолютного импакта журнала можно найти в базе JCR так, как это делалось в п. 3.1. Агрегированный импакт-фактор всей дисциплины также берется из JCR, с помощью следующей процедуры:

Зайти в **JCR** → в секции **Select a JCR edition and year** выбрать **JCR Science Edition** и в выпадающем меню год 2012 → в секции **Select an option** выбрать **View a group of journals by** и в выпадающем меню опцию **Subject Category** → кнопка **SUBMIT** → в секции **1) Select one or more categories from the list** выбрать в списке рубрику **ACOUSTICS**, а в секции **2) Select to view Journal data or aggregate Category data** переключиться в позицию **View Category Data – sort by**: → кнопка **SUBMIT**

Так мы нашли агрегированный импакт-фактор для дисциплины ACOUSTICS: он указан в столбце «Aggregate Impact Factor».

### Пример

Импакт-фактор журнала «ACS Catalysis» за 2012 г. равен 5,27, импакт-фактор журнала «SIAM Journal on Optimization» — 2,08. При этом первый журнал отнесен к тематической рубрике «физическая химия», агрегированный импакт-фактор которой 3,92. Второй журнал помещен в категорию «прикладная математика», для нее агрегированный импакт-фактор равен 1,08. Таким образом, относительный импакт-фактор журнала «ACS Catalysis» оказывается равным  $5,27 / 3,92 = 1,34$ , а журнала «SIAM Journal on Optimization» —  $2,08 / 1,08 = 1,93$ . По этому показателю математический журнал ощутимо превосходит журнал по химии, в то время как по абсолютному импакт-фактору он отстает более чем в два с половиной раза.

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. В таком случае описанный метод может помочь 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, например, 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 индикатора.

### Пример

Найдем цитируемость статей экономиста Екатерины Журавской относительно цитируемости всех российских статей по экономике. В качестве публикационного окна и окна цитирования возьмем интервал 2008–2012 гг., но на этот раз посчитаем этот показатель по базе данных Scopus. Сначала найдем все статьи 2008–2012 гг. выхода, в которых хотя бы один автор указал Россию как страну, где расположено его место работы, и выделим среди них статьи по экономическим наукам.

Зайти в **Scopus** → в поисковом окне набрать **Russia\*** (чтобы собрать значения **Russia** и **Russian Federation**: в отличие от WoS, в Scopus не проведена унификация этих вариантов), в выпадающем меню справа выбрать **Affiliation Country**, в разделе **Limit to: Date Range (inclusive)**

установить значения Published на 2008 и 2012 → кнопка поиска с лупой → слева в фасеточном меню навести мышку на раздел **Subject Area**, тогда появится ссылка **View more**, нажать ее (мы ищем экономику среди дисциплин, пока ее нет) → еще раз нажать **View more**, список еще раз расширится и появится позиция **Economics, Econometrics and Finance**, поставить рядом с ним галочку → кнопка **Limit to**

Получен список из 576 статей по экономике, написанных отечественными авторами с 2008 по 2012 г. Теперь необходимо посчитать, сколько ссылок они получили с 2008 по 2012 г. Для этого используем режим View citation overview (он полностью аналогичен функции Create Citation Report, которую мы использовали в WoS в п. 3.4)

Выделим все результаты поиска. Над результатами поиска, слева от функции **Export**, есть пустой квадрат со стрелкой справа от него. Нажать на стрелку, в выпавшем окошке выбрать **Select all** → нажать ссылку **View citation overview**, также размещенную над поисковыми результатами → для получения информации о цитируемости найденного массива статей за 2008–2012 г. в зоне **Overview options** установить значения **Date range** на **2008** и **2012** → кнопка **Update Overview**

В ячейке Subtotal показано число ссылок, полученных всеми вышедшими в 2008–2012 г. статьями российских экономистов за тот же период 2008–2012 г. Оно равно 709 (колонка Total содержит также ссылки из 2013 и последующих годов, поэтому нас не устраивает). Таким образом, средняя цитируемость статьи равна  $709 / 576 = 1,23$ .

Для нахождения аналогичного показателя для Е. Журавской используем специальный поиск авторов.

Зайти в **Scopus** → выбрать режим **Author search** (вкладка над поисковой строкой) → набрать в графе **Author Last Name** фамилию **Zhuravskaya**, в графе **Author Initials or First Name** указать **E** → кнопка поиска с лупой → из двух результатов нас интересует **Zhuravskaya, Ekaterina V.** (поскольку второй автор не экономист), навести на нее мышку, тогда появится надпись **Documents** под числом документов этого автора (**21** документ), нажать эту ссылку **Documents**

Все публикации Е. Журавской найдены, необходимо при помощи фасеточного меню слева выделить только те статьи, которые вышли в 2008–2012 г. (их будет 8), после этого исследовать их цитируемость точно так же, как это было сделано выше для всех экономистов. Выяснится, что 8 статей Журавской, вышедшие в 2008–2012 г., были процитированы за тот же период времени 42 раза (см. рис. 4). Таким образом, в среднем одна статья получила 5,25 ссылки, что более чем в 4 раза превосходит показатель для всего российского экономического сообщества. В данном случае высокий уровень публикаций исследуемого автора не вызывает сомнений.

## Примечание

Для строгости заметим, что мы не требовали, чтобы все анализируемые статьи Е. Журавской были подписаны российским местом работы. Наш поиск был задан так, что среди найденных 8 публикаций вполне могли фигурировать и те, в которых у автора указаны только зарубежные аффилиации (в данном случае это не так, но это, в известном смысле, случайность).

## Примечание

При выборе публикационного окна больше одного года (как было сделано в данном случае) проявляется, однако, некоторая слабость использованного метода. Предположим, что автор публиковал свои статьи не равномерно из года в год, а, например, по нарастающей — всё 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 (если предположить, что научный уровень всех статей в обоих журналах одинаков). Эти недостатки преодолеваются при использовании более тщательного подхода, описанного в п. 4.2.

Sort documents		Date range		Citations							
Year descending		2008	to	2012	Update Overview						
8 Cited Documents		Citations									
		<2008	2008	2009	2010	2011	2012	Subtotal	>2012	Total	
Total		0	0	1	8	11	22	42	32	74	
<input type="checkbox"/>	2012 Everyone hates privatization, bu...						1	1		1	
<input type="checkbox"/>	2011 M-form hierarchy with poorly-div...					1		1	2	3	
<input type="checkbox"/>	2011 Media and political persuasion: ...					2	2	4	6	10	
<input type="checkbox"/>	2010 Interest group politics in a fed...					1	1	2	2	4	
<input type="checkbox"/>	2010 Businessman candidates					1	3	4	3	7	
<input type="checkbox"/>	2010 What do Russians think about tra...					1	2	3	3	6	
<input type="checkbox"/>	2009 Who wants to revise privatizatio...			1	1	1	6	8	8	16	
<input type="checkbox"/>	2009 (Un)happiness in transition		1	7	4	7	7	19	8	27	

Display 25 documents 1 to 8

Рис. 4. Цитируемость публикаций Е. Журавской из года в год, функция View Citation Overview. Скриншот интерфейса базы данных Scopus (Elsevier).

Тот же метод легко обобщается на организации, где ученые 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. 4.2.

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 пункте 4.2.

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

#### **4.2. Относительная цитируемость разнородного потока публикаций («crown indicator»)**

Представим себе теперь ситуацию, максимально приближенную к реальной: необходимо оценить научный уровень статей, опубликованных организацией широкого научного профиля, например российским 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. 4.1. Здесь особенных 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. 4.1:

— 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*.

Если мы исследуем университет, то в качестве референтной группы логично выбрать все российские вузовские публикации (заметим: включая публикации и самого анализируемого университета) — 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 отечественной науки. Наконец, будет логичным и выбор всех мировых публикаций для анализа соответствия исследуемого вуза общемировому уровню.

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 данному моменту);

— тип публикации (как уже говорилось в п. 2.1, в среднем, например, научные обзоры цитируются чаще, чем оригинальные исследовательские статьи).

Для учета всех этих факторов и корректного определения относительной цитируемости всего публикационного потока организации предлагается процедура, суть которой в следующем. Определим, сколько получили бы ссылок «средние» публикации из референтной группы, имеющие *ту же дисциплинарную принадлежность, тот же год выхода и тот же тип документа*, что и публикации исследуемого вуза. Вводится показатель «ожидаемое число ссылок»: для каждой статьи вуза это среднее число ссылок, которое получили в референтной группе публикации с теми же «характеристиками», что у данной статьи (область науки, год выхода, тип документа). Если рассматриваемая статья вуза в действительности получила больше ссылок, чем ожидаемое их число, то она превосходит средний уровень такой же публикации в референтной группе, если меньше — уступает ему. Далее «ожидаемое число ссылок» суммируется по всем статьям вуза и получается «ожидаемое число ссылок» для всего университета.



Если переформулировать, то «ожидаемое число ссылок» для вуза — это число цитирований, которое получил бы вуз, если бы все его статьи строго держались среднего научного уровня референтной группы. «Средний уровень» при этом вычисляется отдельно по каждой «категории» публикаций группы-референта (т.е. для публикаций, обладающих идентичной дисциплинарной привязкой, годом выхода и типом).

Теперь осталось взять реальное число ссылок, полученных всеми статьями вуза (его часто называют «наблюдаемым числом ссылок»), и посчитать его отношение к вычисленному ожидаемому уровню. Этот индикатор показывает, получила ли вся совокупность статей исследуемого университета ссылок больше или меньше, чем если бы публикационный поток вуза обладал средней по референтной группе цитируемостью. (Повторим, при этом «среднее» дифференцируется по характеристикам дисциплина/год выхода/тип публикации.) Назовем этот показатель «цитируемость, нормализованная по областям науки» (ЦНОН).

### Примечание

Показатели «ожидаемое число ссылок» и «наблюдаемое число ссылок» в западной традиции называются, соответственно, «(mean) expected citation rate» и «(mean) observed citation rate». Их отношение, итоговый показатель, в традиции Центра исследования науки и технологий (CWTS) Лейденского университета называется, ни много ни мало, «главный (коронный) индикатор» (crown indicator). Более скромное название — «средняя нормализованная цитируемость» (mean normalized citation score). Почти такое же обозначение используют в Центре мониторинга исследований и разработок (г. Лёвен) — «нормализованная средняя цитируемость» (normalized mean citation rate). Стоит также заметить, что сейчас предлагается и альтернативный способ усреднения по всему массиву ссылок, полученных публикациями заданной организации. А именно, сначала для каждой отдельной ее статьи считается отношение числа полученных ссылок к ожидаемому их числу (со всеми условиями, упомянутыми выше), а потом берется среднее всех этих отношений.

### Пример

Ввиду сложности оперирования с реальными показателями вуза, поддерживающего многопрофильную публикационную активность, на этот раз продемонстрируем алгоритм определения цитируемости, нормализованной по областям науки, на модельных данных.

координаты ячеек			1	2	3	4
			физика		химия	
			Article	Review	Article	Review
			<b>2010</b>			
A	вуз	публикации	20	—	10	1
B		ссылки	66	—	43	9
C	Россия	публикации	1000	150	800	100
D		ссылки	3000	900	3200	900
			<b>2011</b>			
E	вуз	публикации	20	1	14	—
F		ссылки	52	5	35	—
G	Россия	публикации	1000	150	800	100
H		ссылки	2000	600	2400	600
			<b>2012</b>			
J	вуз	публикации	20	1	20	1
K		ссылки	26	2	44	3
L	Россия	публикации	1100	150	900	100
M		ссылки	1100	300	1800	300

Таблица 1. Публикации и полученные ими ссылки: анализируемый вуз и Россия (условные данные).

В Таблице 1 представлены статистические данные о публикациях некоторого (условного) вуза, специализирующегося на физике и химии. Для простоты ограничимся только двумя типами документов — Article (исследовательская статья) и Review (научный обзор) — и тремя годами выхода оцениваемых статей, с 2010 по 2012. Для демонстрации методики расчета ЦНОН этого будет достаточно. В качестве референтной группы выбраны все организации России, поэтому также приведены данные (тоже условные) по всем российским публикациям в соответствующей дисциплине. Итак, в строках А, Е, J — число публикаций исследуемого вуза по физике и химии в 2010–2012 гг. с разбиением по типам документов (Article и Review). В строках В, F, K — число ссылок, которые получили соответствующие работы за период 2010–2012 гг. Строки С, G, L и D, H, M содержат ту же информацию по всем российским публикациям в тех дисциплинах, в которых публикуются сотрудники вуза.

### Примечание

В данном случае нам было необязательно ограничивать окно цитирования сверху 2012 годом. Можно было считать, сколько указанные документы получили ссылок «на текущий момент», «до настоящего времени». Ввиду строгой корректности алгоритма подсчета ЦНОН это не оказало бы негативного влияния на методику. Единственное, что требуется в таком случае, — зафиксировать момент, в который берутся все данные о полученных документами ссылках — как для публикаций вуза, так и для всей референтной группы.

Определим «ожидаемое число ссылок», которые «должны» получить публикации рассматриваемого вуза. Например, в ячейке A1 указано, что в 2010 г. сотрудники учебного заведения опубликовали 20 документов типа «Article» по физике. При этом всех российских публикаций этого типа по этой дисциплине в 2010 г. вышло 1000 штук (C1), и они в сумме получили за 2010–2012 гг. 3000 ссылок (D1). Средняя цитируемость отечественных работ по физике типа «Article», вышедших в 2010 г., составила, таким образом, 3 ссылки на статью (D1 / C1). 20 публикаций вуза из ячейки A1, если бы они в точности соответствовали среднероссийскому уровню, должны были бы получить 60 ссылок. Для A1 «ожидаемое число ссылок» равно 60. Для A2 этот показатель, естественно, равен нулю (публикаций типа «Review» в 2010 г. по физике у авторов вуза не выходило); для A3:  $D3 / C3 * A3 = 3200 / 800 * 10 = 40$ ; для A4:  $900 / 100 * 1 = 9$ . Итого публикации 2010 г. «должны» получить  $60 + 40 + 9 = 109$  ссылок. Аналогичным образом рассчитываются значения ожидаемого числа ссылок для строк E и J, которые будут равны  $40 + 4 + 42 = 86$  для публикаций 2011 г. и  $20 + 2 + 40 + 3 = 65$  для 2012 г. Итого ожидаемое число ссылок для всего массива публикаций исследуемого вуза равно  $109 + 86 + 65 = 260$ .

### Примечание

Обратите внимание, что несмотря на искусственное происхождение приведенных в Таблице 1 данных, они наделены некоторыми свойствами из реальной жизни. Например, число полученных публикациями ссылок уменьшается от работ 2010 г. к работам 2012 г.: более свежие статьи еще не успели получить значительное число цитирований. Поэтому хотя в целом публикационная активность анализируемого вуза из года в год возрастает, ожидаемое число полученных ссылок уменьшается.

Наблюдаемое, реальное число ссылок, полученных вузом, посчитать еще проще — достаточно сложить показатели в строках B, F и K. Итоговая сумма — 285. Наблюдаемое число ссылок больше ожидаемого, а это значит, что публикации вуза имеют в среднем более высокий научный уровень, чем среднестатистические российские работы, имеющие те же характеристики (тип/год выхода/область науки). Коэффициент ЦНОН =  $285 / 260 = 1,10$ , больше единицы.

Обратите внимание на примечательное обстоятельство. Если бы мы использовали обычный «огрубленный» подход, введенный в п. 4.1, мы бы получили обратный результат. Средняя цитируемость работ вуза, взятых как целое, без дифференциации по типам/годам выхода/дисциплинам, равно отношению суммы строк B, F, K к сумме строк A, E, J =  $285 / 108 = 2,64$ . При этом средняя цитируемость всех российских работ равна отношению суммы строк D, H, M к сумме строк C, G, L =  $17100 / 6350 = 2,69$ . Средняя цитируемость публикаций вуза, взятых как целое, получилась бы ниже среднего уровня по России. В данном случае только тщательный подход с разделением по «характеристикам» публикаций и сопоставлением подобного с подобным дало возможность сделать правильный вывод об эффективности научной деятельности вуза на фоне среднероссийского уровня.

Очевидно, что аналогичный подход 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 исследуемой организации будет равно среднему числу цитирований, полученных публикациями данного журнала, которые вышли в том же году и относятся к тому же типу документа. Далее вновь проводится суммирование ожидаемого числа ссылок по всем статьям организации. Наблюдаемое число ссылок считается без изменений. Отношение будет искомым показателем (можно его назвать «цитируемость, нормализованная по журналам», ЦНЖ). Этот индикатор характеризует, 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.

### Примечание

Данный показатель, посчитанный для целых стран, также называют 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.

### Где найти?

К сожалению, показатели нормализованной цитируемости (как по научным областям, так и по журналам) не публикуются в готовом виде в «первичных» библиометрических базах данных (WoS, Scopus). Их можно найти, например, в аналитическом инструменте InCites компании Thomson Reuters (здесь наблюдаемая/ожидаемая цитируемость называются, соответственно, actual/expected citations). Кроме того, эти или очень похожие показатели часто вычисляются в различных рейтингах, например в SCImago Institutions Rankings (базируется на данных из Scopus) или CWTS Leiden Ranking (базируется на данных из WoS):

<http://www.scimagoir.com/>

<http://www.leidenranking.com/ranking>

В первом из них показатель называется «normalized impact», во втором — «mean normalized citation score»; в обоих случаях это нормализация по областям науки. Можно взять МНОН из этих рейтингов, однако в данном случае вся ответственность по корректному его расчету ложится на производителей рейтинга, которые не всегда удачно справляются с данными по российским организациям. С другой стороны, самостоятельное вычисление по WoS или Scopus ЦНЖ или, тем более, ЦНОН — весьма трудоемкий процесс.

Поскольку указанные инструменты и рейтинги имеют международный статус, в качестве референтной группы всегда выступают все мировые публикации. Для простоты далее мы также будем считать референтной группой весь мир. (Вычисление знаменателя ЦНЖ по всем, а не по избранным статьям журнала это уже неявно предполагало.) Для определения показателей организации на фоне какой-либо другой группы (например, страны), достаточно найти индикаторы этой группы относительно всего мира и сравнить с показателями исследуемой организации относительно всего мира.

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 скорее всего будут в среднем меньше, чем погрешность при любом другом автоматическом определении тематики публикации. Более того, если мы — как обычно делается в библиометрических базах данных — сами дисциплины приписываем статьям на основе тематической рубрикации журналов, в которых они опубликованы, то получается, что нормализация по журналам дает более точный, детальный взгляд на сопоставление публикаций, чем нормализация по областям науки.

Здесь, однако, есть ключевой момент. Предположим, что некоторый университет получает высокий показатель нормализованной по журналам цитируемости, ЦНЖ. Но при этом если мы посчитаем его цитируемость, нормализованную по областям науки, ЦНОН, может неожиданно выясниться, что она меньше единицы, что она низкая. Как такое может быть? Объяснение существует: это означает, что *авторы данного вуза выбирают для публикации своих статей слабые журналы*. А уже там, в рамках слабых журналов, они могут превзойти средний уровень, так что цитируемость их статей превысит средние показатели этих непрестижных изданий. Ситуации «первый парень на деревне» противостоит обратная, «последний в городе», когда авторы организации подают свои статьи в ведущие журналы (заметим: и их там публикуют), однако, попав туда, публикации не достигают среднего уровня цитируемости этих изданий. В таком случае мы можем получить высокую цитируемость, нормализованную по областям науки, и низкую — при нормализации по журналам.

Так или иначе очевидно, что для всестороннего анализа научной деятельности организации или автора следует использовать оба показателя — и цитируемость, нормализованную по областям науки, и цитируемость, нормализованную по журналам. При этом отношение первого показателя ко второму называют коэффициентом публикационной стратегии (КПС, англ. Relative Publication Strategy Index). Он, как мы видели, действительно характеризует стратегию авторов, используемую при публикации своих статей. Если коэффициент больше единицы, автор(ы) стремятся публиковать свои работы в высокоцитируемых изданиях, в журналах, чьи показатели выше среднего по области науки, к которой они относятся. Если коэффициент меньше единицы — в среднем автор(ы) публикуют свои статьи в изданиях ниже среднего уровня.

### Примечание

Можно заметить, что у ЦНОН и ЦНЖ одинаковый числитель — это «наблюдаемое число ссылок», т. е. суммарное число цитирований, полученных организацией/автором. Таким образом, КПС равен отношению двух «ожидаемых цитирований»: посчитанных по средней цитируемости журналов, в которых публиковалась организация, и по средней цитируемости всех областей науки, где выходили ее статьи. Обратите внимание, что КПС не зависит от реального («наблюдаемого») числа ссылок, полученных статьями организации: он характеризует не полученные цитирования, а исключительно уровень тех журналов, где авторы публикуют свои работы.

Суммируем сведения об индикаторах цитируемости, нормализованной по журналам и по областям науки (ЦНОН и ЦНЖ), при помощи рис. 5. На данной диаграмме ось абсцисс обозначает цитируемость, нормализованную по журналам (ЦНЖ), а ось ординат — цитируемость, нормализованную по областям науки (ЦНОН). Каждая организация, для которой посчитаны ЦНЖ и ЦНОН, представляется в виде точки на данном графике, как показаны в качестве примера точки для организаций  $o$ ,  $o_1$  и  $o_2$ . Такие графики, по осям которых отложены нормализованные величины, обычно называют «диаграммами относительных величин» (relational charts). Они дают возможность наглядно представить показатели одной или нескольких организаций и прекрасно подходят для аналитической работы.

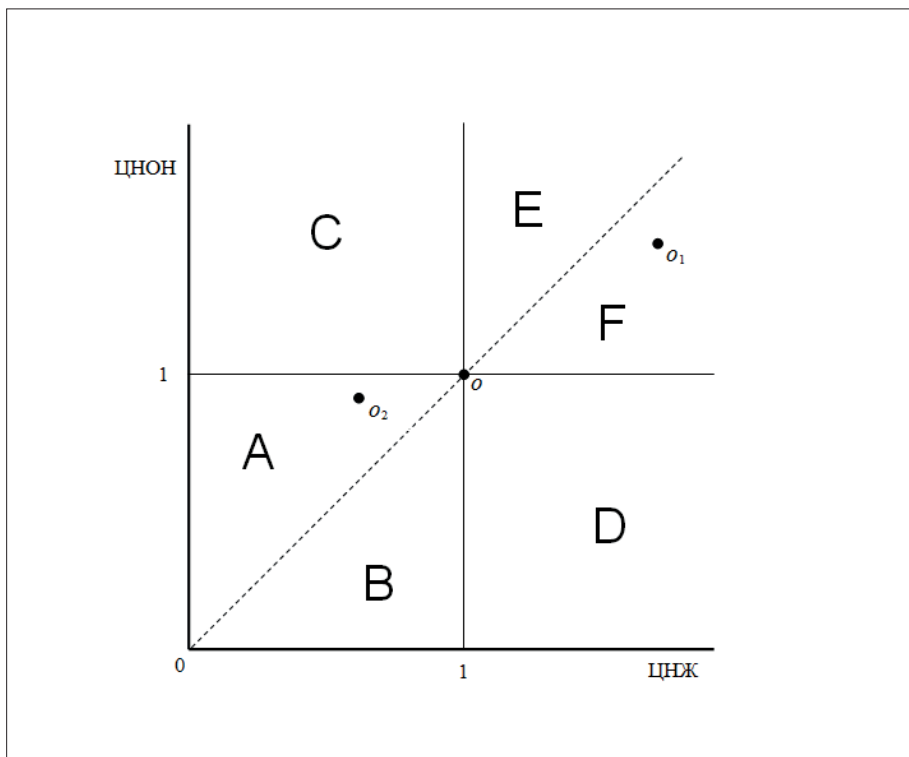


Рис. 5. Диаграмма относительных величин для индикаторов ЦНЖ и ЦНОН.

Так, на рис. 5 выделены шесть областей (A, B, C, D, E, F), попадание в каждую из которых может определенным образом охарактеризовать научную деятельность исследуемой организации. Области разделяют прямые  $ЦНЖ=1$ ,  $ЦНОН=1$  и  $ЦНОН=ЦНЖ$ . Обратите внимание на биссектрису  $ЦНОН=ЦНЖ$  (показана пунктиром) — поскольку  $КПС=ЦНОН / ЦНЖ$ , выше этой прямой находятся организации с коэффициентом публикационной стратегии больше 1, ниже — меньше 1. Таким образом, попадание точки, соответствующей той или иной организации, в одну из данных областей может быть охарактеризовано так:

**А:** сотрудники таких организаций выбирают для своих публикаций журналы выше среднего уровня, однако цитируемость их работ оказывается настолько ниже среднего уровня этих журналов, что в итоге и цитируемость относительно областей науки оказывается ниже среднего уровня

**В:** сотрудники таких организаций выбирают для своих публикаций журналы ниже среднего уровня, при этом цитируемость их работ оказывается ниже даже среднего уровня этих журналов, поэтому цитируемость относительно областей науки, конечно, оказывается ниже среднего уровня

**С:** сотрудники таких организаций выбирают для своих публикаций журналы выше среднего уровня, поэтому хотя цитируемость их работ и оказывается ниже среднего уровня этих журналов, в итоге цитируемость относительно областей науки оказывается выше среднего уровня

**D:** сотрудники таких организаций выбирают для своих публикаций журналы ниже среднего уровня, поэтому хотя цитируемость их работ и оказывается выше среднего уровня этих журналов, в итоге цитируемость относительно областей науки оказывается ниже среднего уровня

**E:** сотрудники таких организаций выбирают для своих публикаций журналы выше среднего уровня, при этом цитируемость их работ оказывается еще выше среднего уровня этих журналов, в итоге цитируемость относительно областей науки, конечно, оказывается выше среднего уровня

**F:** сотрудники таких организаций выбирают для своих публикаций журналы ниже среднего уровня, однако цитируемость их работ оказывается настолько выше среднего уровня этих журналов, что в итоге и цитируемость относительно областей науки оказывается выше среднего уровня

### **Примечание**

В приведенных характеристиках, являющихся, по сути, легендой к рис. 5, для краткости говорится о «журналах выше/ниже среднего уровня». Речь идет, конечно, о сравнении со средней цитируемостью дисциплин, к которым эти журналы относятся, причем берется оценка по всей совокупности журналов и дисциплин, где имеются публикации организации. Точная формулировка: сотрудники организации публикуются в таком наборе журналов, что взвешенная сумма их средних цитируемостей выше/ниже взвешенной суммы средних цитируемостей соответствующих дисциплин (веса берутся по числу статей организации в каждом из журналов/дисциплин). Из этого определения становится ясно, что ряд журналов, в которых опубликованы статьи организации, могут занимать, наоборот, отстающие/ведущие позиции в своих дисциплинах, но их вклад в КПС нивелируется большим числом статей в высокоцитируемых/низкоцитируемых изданиях.



Диаграмма относительных величин, подобная представленной на рис. 5, может быть использована как для анализа публикационной стратегии и цитируемости статей организаций в фиксированный момент времени, так и в динамике. В последнем случае траектории точек, соответствующие меняющимся показателям организаций, выявят направление и темп эволюции научной деятельности ученых соответствующих вузов/институтов/лабораторий и т. д.

Наконец, обратим внимание на точку *o*. Она соответствует организации, статьи которой в среднем совпадают по цитируемости и со средней цитируемостью журналов, где они опубликованы, и со средней цитируемостью областей науки, к которым они относятся (это, конечно, не означает, что *каждая отдельная статья* обязательно обладает этими свойствами); в среднем сотрудники такой организации выбирают для опубликования своих работ журналы среднего уровня цитируемости. Если бы мы рассмотрели совокупно сразу все публикации («референтной группы»), относительно которой оцениваются организации (напомним, что сами организации также входят в эту группу), такой «метаорганизации» на рис. 5 также соответствовала бы точка *o*.

### 4.3. Нормализация по источникам ссылок и индикатор SNIP

Показатели, рассмотренные нами в пп. 4.1 и 4.2, имеют одну общую особенность: при нормализации, взятии частного, в знаменателе стоит среднее число ссылок, *полученных* теми или иными публикациями — например, публикациями заданного типа/года выхода из некоторой области науки или публикациями заданного типа/года выхода из некоторого журнала. В зарубежной литературе это называется «cited-side normalization» — нормализация по цитируемости, нормализация по адресатам ссылок (т. е. цитируемым статьям).

При решении задачи междисциплинарного сопоставления, учета индивидуальных особенностей разных областей науки, возможен и иной подход. Можно в качестве базы нормализации взять не число ссылок, *полученных* теми или иными публикациями, а число ссылок, *сделанных* в некотором массиве статей (например, в рамках всей научной дисциплины). Такой метод нормирования называется «нормализация по источникам ссылок» («citing-side normalization») или «source normalization»). Идея данного метода — в расчете своеобразного «потенциала цитируемости» публикаций, относящихся к той или иной области науки, или, можно сказать, нахождение «плотности ссылочного поля» дисциплины. Достигается это ответом на вопрос: а сколько в данной научной области в среднем делается ссылок в одной статье? Это и задает, так сказать, «плотность ссылочного поля» и определяет среднюю веро-

ятность получения ссылок статьями из данной дисциплины. По таким образом определенному «потенциалу цитируемости» нормируется реальное число ссылок, полученное публикациями. В разных областях науки будут разные потенциалы цитируемости, что даст корректную нормировку абсолютного числа полученных статьями ссылок в разных дисциплинах.

### **Пример**

Предположим, что в одной дисциплине существует 10 статей, каждая из которых сделала (за определенный период времени) по 5 ссылок на статьи из этой же дисциплины. Тогда всего существует 50 ссылок, распределенных по публикациям данной научной области, т. е. в среднем по 5 на статью. В другой области науки также есть 10 статей, но каждая из них сделала всего по одной ссылке. Тогда среднее число ссылок на статью — 1 и вероятность получения статьей цитирования в 5 раз меньше. Поэтому и потенциал цитируемости будет в 5 раз меньше, чем в первом случае.

Нормировка на «потенциал цитируемости», можно сказать, учитывает «ценность» каждой ссылки. В областях, где в расчете на одну публикацию было сделано мало ссылок, ценность каждой будет больше, поэтому и больший вклад в показатель даст каждая полученная ссылка. В областях, где «ссылочное поле» плотное и ссылок сделано много, надо получить больше ссылок для достижения эквивалентного уровня «влиятельности» журнала. Итак, «потенциал цитируемости» отражает среднюю длину списка цитируемой литературы в статьях той или иной области.

### **Примечание**

При определении потенциала цитируемости речь идет именно о среднем числе сделанных ссылок в расчете на одну статью, а не о суммарном количестве ссылок в данной научной области. Здесь в методике заложено то, что большое число статей в дисциплинах приводит к увеличению общего количества сделанных ссылок, но одновременно — к росту возможных «адресатов» ссылок, тех статей, которые могут быть процитированы в данной области. Поэтому если говорить о «вероятности» получения ссылки, то следует считать именно среднюю, а не суммарную длину списков цитируемой литературы в соответствующей научной дисциплине. Схематически говоря, если статей в дисциплине мало, получение ссылки имеет меньшую ценность: публикациям просто «нечего цитировать», кроме как друг друга.

В 2009 г. был введен первый журнальный индикатор рассматриваемого типа, получивший широкое распространение, — «source normalized impact per paper» (SNIP). В развернутой форме название можно перевести как «нормализованная по источникам ссылок цитируемость в расчете на одну статью». Он был рассчитан для всех изданий, входящих в базу данных Scopus.

### Где найти?

SNIP публикуется с различной периодичностью (ранее — раз в полгода, сейчас — в год) в базе данных Scopus. Рассчитывается SNIP для журналов, ряда трудов конференций и даже для некоторых книжных серий. Чтобы найти издание и значение индикатора для него, используйте в Scopus функцию поиска по источникам Browse Sources. Кроме того, значения SNIP можно найти в открытом доступе на сайте JournalMetrics, также поддерживаемом компанией Elsevier:

<http://www.journalmetrics.com/values.php>

В действительности с 2012 г. в этих источниках публикуется «улучшенный» индикатор, т. н. SNIP2, см. далее.

В методике расчета SNIP прежде всего для каждого журнала нестандартным образом определяется «индивидуальная дисциплинарная область». Здесь, в отличие от практически всех остальных методик, научную (дисциплинарную) область формирует не набор журналов, а набор *статей*. А именно, индивидуальная дисциплинарная область журнала — это все статьи в Scopus, которые вышли в отчетном году (год, за который мы рассчитываем SNIP) и *цитировали хотя бы один раз выпуски данного журнала, опубликованные за последние 10 лет (в SNIP2 — за последние 3 года)*.

Таким образом авторам индикатора удастся уйти от привычного деления журналов как целого по тематическим рубрикам (как, например, в JCR), — деления неизбежно негибкого и шаблонного. Теперь каждый журнал, во-первых, имеет свою собственную дисциплинарную область; во-вторых, эту область составляют не целые журналы, а отдельные публикации.

Расчет SNIP учитывает только публикации типов «Article», «Conference paper» и «Review» (исследовательская статья, доклад из сборника трудов конференции и научный обзор). Причем это ограничение дейст-

вует при рассмотрении как вопроса «откуда приходит ссылка», так и «куда ведет ссылка». Учитываются только ссылки, сделанные из документов одного из трех указанных типов *на* документы одного из трех указанных типов. Кроме того, при делении (нормировании) на число публикаций также учитываются только документы данных типов: можно сказать, что публикации других типов «не существуют» для SNIP.

Хронологические интервалы при расчете SNIP: публикационное окно — три года, предшествующих отчетному (не включая самого отчетного года); окно цитирования — только отчетный год. Анализируется цитируемость в отчетном году статей, вышедших в течение трех предыдущих лет. Таким образом, публикационное окно на один год больше, чем при расчете импакт-фактора.

SNIP определяется как отношение средней цитируемости статей журнала к потенциалу цитируемости его индивидуальной дисциплинарной области. Средняя цитируемость — это, разумеется, число полученных журналом (точнее, его публикациями трех перечисленных выше типов) ссылок деленное на число этих публикаций. Что касается потенциала цитируемости, то он определяется особым образом. Как говорилось выше, потенциал цитируемости — это среднее число ссылок в списках цитируемой литературы у статей заданной научной области (в нашем случае — «индивидуальной дисциплинарной области» журнала). Однако в SNIP при расчете потенциала цитируемости учитываются только ссылки, ведущие на документы, которые а) также содержатся в базе данных Scopus и б) изданы в течение трех лет, предшествующих отчетному году. Можно сказать, что на потенциал цитируемости влияют только такие ссылки, которые могли бы повлиять на числитель SNIP данного журнала — т. е. на среднюю цитируемость его статей. Они удовлетворяют всем условиям: ведут из документов трех учитываемых в SNIP типов на документы тех же трех типов; сделаны из отчетного года на документы, вышедшие в течение трех предыдущих лет; ведут на документы, имеющиеся в Scopus, а не «за пределы» базы данных. Среднему количеству *именно таких* ссылок в статьях индивидуальной области журнала равен потенциал цитируемости, помещенный в знаменатель SNIP.

### Примечание

Все кажущиеся «усложнения», которые мы находим в методике SNIP, служат одной цели: созданию «равных условий» для различных научных дисциплин, что сделает возможным междисциплинарное сравнение индексов. Так, учет в потенциале цитируемости только ссылок, ведущих «внутрь» базы данных, позволяет уравнивать области, где цитируемая литература хорошо представлена в Scopus, и те, где встречается много ссылок на материалы, не охваченные базой (в частности, где больше ссылок на книжные, а не журнальные источники; на диссертации, кодексы и т. д.). Ограничение при расчете потенциала «возраста» ссылок тремя предшествующими годами сглаживает разницу между дисциплинами, где основной массив цитирований ведет на свежую литературу (попадающую в «зону влияния на SNIP»), и областями науки, где чаще цитируются источники, вышедшие давно (в последних и потенциал цитируемости, знаменатель, будет меньше — ввиду исключения вклада в него «древних» ссылок). Еще раз: только те ссылки, которые теоретически могут повлиять на числитель SNIP, влияют на его знаменатель.

Итак, показатель SNIP выражает отношение числа полученных журналом цитирований в расчете на одну статью к специальному образом вычисленному потенциалу цитируемости индивидуальной дисциплинарной области данного журнала. Являясь в известном смысле «усовершенствованием» импакт-фактора, он, вероятно, лишен его наглядности и простоты, однако приближает нас к возможности кросс-дисциплинарного сравнения *абсолютных* значений индикатора.

В заключение заметим, что, несмотря на свою молодость, SNIP прошел первичную апробацию в научном сообществе. В частности это привело к тому, что с 2012 г. публикуются значения «оптимизированного» SNIP, т. н. «SNIP2», рассчитываемого по скорректированной методике. Изменения преследовали цель избавиться от «контр-интуитивных» свойств исходного показателя — например, добавление еще одной ссылки на журнал никогда не должно вести к уменьшению SNIP (при некоторых условиях исходный показатель мог уменьшаться), а слияние двух журналов с разными SNIP не должно приводить к тому, что итоговое «объединенное» издание будет иметь SNIP, выходящий за пределы интервала между двумя SNIP исходных журналов (это также было возможно).

#### 4.4. Ранговые индикаторы

Упомянем еще об одном подходе, применяемом при проведении кросс-дисциплинарных сопоставлений журналов. Можно принципиально отказаться от использования абсолютных значений 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 (импакт-фактора, SNIP, SJR и т. д.) проводится анализ только занятых в этой иерархии изданиями порядковых мест.

Суть рангового метода — разбиение полученного упорядоченного списка 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.

#### Пример

Вернемся к примеру п. 4.1 и сравним ранговыми методами журналы «ACS Catalysis» (дисциплина «физическая химия») и «SIAM Journal on Optimization» (категория «прикладная математика»), используя упорядочивание по импакт-фактору. Первый журнал занимает 24-е место в своей научной области по импакт-фактору среди 135 журналов и поэтому относится к первому квартилю. Второй занимает 10-е место из 247 и также относится к первому квартилю. При таком делении оба журнала равны. Однако если бы вместо квартилей мы использовали деление рейтинга на 10 частей (децили), то первый журнал оказался бы во втором дециле, а второй — в первом. При таком «более подробном» подходе «ACS Catalysis» уступает журналу «SIAM Journal on Optimization».

#### Примечание

Возможен предельный случай таких «делений», 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 показать, что при этом показатель «относительной позиции» журнала, занимающего  $n$ -ное

место в рейтинге, равен  $(N - n) / (N - 1)$ . Для рассмотренных выше изданий этот индикатор равен 0,83 (ACS Catalysis) и 0,96 (SIAM Journal on Optimization). У более «грубого» подхода, у деления на те или иные более крупные «квантили», есть то преимущество, что некоторые группы журналов попадают в одну и ту же категорию, что интуитивно соответствует представлению о том, что существуют «примерно одинаковые» издания.

Аналогичный подход может быть применен 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 в случае с журналами). 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. примечание из п. 3.5, где ранговый подход, а именно использование медианы, предлагался для анализа импакт-факторов). В этом случае отказ от учета значений индикаторов и исследование лишь мест в рейтингах может оказаться более эффективным.

## 5. «Взвешенные» индикаторы

Все изученные ранее индикаторы имели одну общую черту: ссылки, которые получали журнал/автор/организация, были для них «одинаковыми», показатели «не замечали», откуда именно получено то или иное цитирование. Замена одного цитирующего журнала на другой, который дает столько же ссылок, не приводило к изменению индикатора. Некоторым исключением здесь можно считать SNIP, но и в его случае косвенное влияние оказывает лишь то, какая статья цитирует журнал

или, вернее, статья с каким количеством ссылок в списке цитируемой литературы: опосредованно, через влияние на «потенциал цитируемости», это могло приводить к разному вкладу в SNIP.

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. Аналогичная идея лежит в основе расчета ранга «авторитетности» веб-страниц некоторыми поисковыми машинами, например Google (т.н. «page rank»).

### 5.1. SCImago Journal Rank (SJR)

Показатель предложен испанской исследовательской группой SCImago. Так же, как и SNIP, он рассчитывается для всех журналов, входящих в базу данных Scopus. Индекс SJR вычисляется 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 них. Таким образом рассчитывается «престиж», полученный каждым журналом из других журналов, входящих в Scopus, и из самого себя. 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. (Точный алгоритм устроен таким образом, что данное состояние будет обязательно достигнуто.)

#### Примечание

Можно показать, что описанная процедура эквивалентна подсчету числа посещений каждого журнала неким условным читателем в ходе случайного блуждания по ссылкам. Это бесконечный процесс, при котором читатель переходит от журнала по случайно выбранной в нем ссылке к другому журналу (или к тому же самому, если ссылка является самоцитированием), а 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 «престиж» будет пропорционален числу заходов читателя в соответствующий журнал.



При расчете SJR в процедуре перераспределения престижа участвуют ссылки только на те статьи, которые вышли за три последних года. Кроме того, для уменьшения влияния самоцитирования журналов, его величина для каждого издания искусственно ограничивается максимумом в 33% от всех сделанных журналом цитирований. На финальном этапе для определения значения индекса SJR производится нормировка престижа на число статей в журнале, таким образом индекс SJR не зависит от объема издания.

### Где найти?

SJR публикуется вместе со SNIP, поэтому также может быть найден или в подписной базе данных Scopus, или в ресурсе открытого доступа JournalMetrics (см. п. 4.3). Точно так же, как в случае SNIP, с 2012 г. в этих источниках публикуется «улучшенный» индикатор, т. н. SJR2.

## 5.2. Собственный фактор (Eigenfactor) и индекс влияния статьи (Article Influence)

Взвешенный журнальный индикатор, получивший название «собственный фактор» (Eigenfactor), был предложен в 2007 г. специалистами из лаборатории Карла Бергстрёма (Университет Джорджа Вашингтона). Его смысл очень близок к смыслу SJR. Так, разработчики данного показателя интерпретируют его с помощью процесса случайного перехода по ссылкам, т. е. так же, как и в одном из способов иллюстрации физического смысла индекса SJR. Согласно определению, для вычисления значения собственного фактора решается матричное уравнение, однако когда этот метод применяется на практике, показатель определяется с помощью итеративной процедуры, что похоже на способ вычисления индекса SJR. 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»).

Индекс влияния во многом похож на индекс SJR, отличия носят скорее технический, нежели принципиальный характер. В частности, при его расчете учитываются ссылки на статьи, вышедшие за 5 последних лет (а не за три года), а ссылки журнала на свои же статьи исключаются целиком (тогда как SJR ограничивает самоцитирование уровнем 33%).

При этом наиболее серьезным отличием является то, что индекс влияния и индекс SJR рассчитываются по разным базам цитирования: индекс SJR — по базе Scopus, а индекс влияния (и собственный фактор) — по базе WoS.

### Где найти

С 2007 г. Eigenfactor и Article Influence журналов публикуются в JCR. Для их нахождения в базе используется та же последовательность действий, что и при поиске импакт-фактора (см. п. 3.1): значения индикаторов будут приведены в таблице, которая получалась при нахождении импакта, в соответствующих колонках Eigenfactor® Score и Article Influence® Score. Кроме того, с некоторым запозданием (в настоящий момент — год) эти индикаторы размещаются в свободном доступе на сайте <http://eigenfactor.org/>

В заключение данного раздела отметим, что и индекс SJR и индекс влияния сглаживают разницу уровней активности цитирования в различных научных дисциплинах, поскольку «престиж» журнала распределяется поровну между всеми исходящими из него ссылками.

## 6. Индекс Хирша ( $h$ -index)

Индекс Хирша был предложен физиком из Калифорнийского университета в Сан-Диего Х. Хиршем. Этот показатель кардинальным образом отличается от всех рассмотренных ранее, поскольку он пытается дать комплексную оценку одновременно числу публикаций ученого и их цитируемости. Кроме того, методика подсчета индекса Хирша намеренно уходит от определения средних величин, средней цитируемости статей. Это имеет свой смысл, поскольку усреднение зачастую не дает полной картины об исследуемом множестве статей и не позволяет производить корректные сравнения эффективности научной деятельности авторов и организаций. При усреднении сильные искажения могут быть вызваны отдельными публикациями, которые получают очень много ссылок, являясь, фактически, «выбросами».

Согласно определению, индекс Хирша массива публикаций ученого равен  $h$ , если  $h$  статей из этого массива получили не менее  $h$  цитирований, а остальные, соответственно, не более  $h$  цитирований. Графически это проиллюстрировано на рис. 6. Каждая квадратная 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. 6 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. На этом рисунке индексу Хирша соответствует такое максимальное целое число  $h$ , при котором мы еще можем разместить показанный квадрат  $h \times h$  под точками данных. 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 (такой квадрат уже «не помещается»).

Иной способ интерпретации того же графика рис. 6 — поиск пересечения множества точек-публикаций биссектрисой  $y = x$ , показанной на рисунке пунктиром. Индекс Хирша будет соответствовать максимальному номеру статьи, находящейся *не ниже* прямой  $y = x$  (выше или на этой прямой).

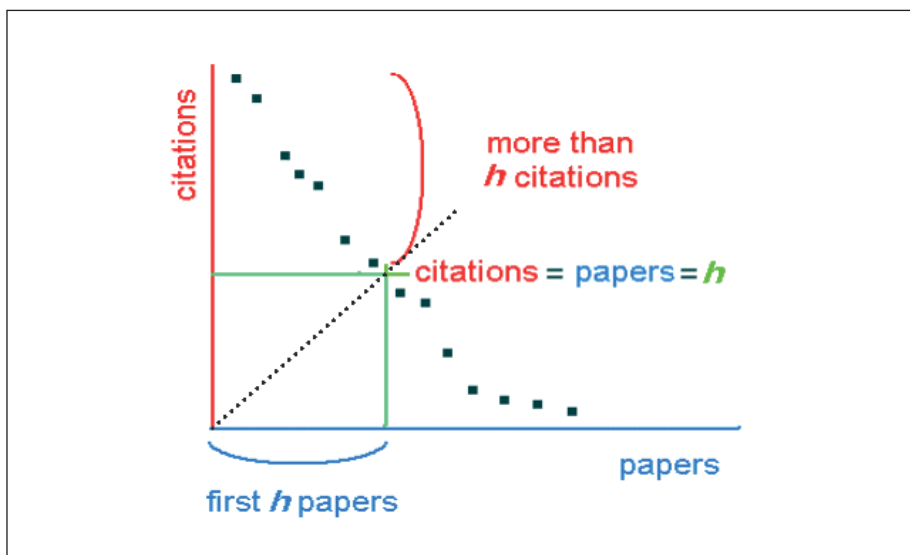


Рис. 6. Индекс Хирша (основа для диаграммы взята с сайта wikipedia.com).

Заметим 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. Для достижения высокого значения индекса Хирша необходимо, чтобы автор писал *много работ, каждая из которых получила бы много цитирований*. Именно в этом смысле *h*-индекс пытается отразить сбалансированную оценку сразу и публикационной активности ученого/организации и цитируемости их работ.

### Пример

Если автор написал 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.

### Где найти?

Индекс Хирша может быть рассчитан по любой базе данных, фиксирующей ссылки, соответственно, его значение может быть разным в зависимости от выбранной базы. Готовое значение индекса Хирша для ученых можно найти, например, в авторских профилях базы данных Scopus. В этом случае, однако, необходимо следить, чтобы база данных корректно распознала работы рассматриваемого ученого, не приписала ему чужие статьи и, наоборот, в полной мере собрала в его профиль его публикации. Если этого достичь не удается, лучше произвести расчет *h*-индекса вручную, контролируя корректность приписанных автору статей. Кроме того, при необходимости это позволит задать требуемые временные рамки (в профиле указывается только *h*-индекс для всей совокупности публикаций автора). В WoS индексы посчитаны для тех ученых, которые завели себе уникальный авторский идентификатор ResearcherID (и разрешают выставлять свои метрики на обзор другим пользователям).

### Пример

Найдем *h*-индекс в базе данных Scopus специалиста по библиометрии и истории науки, сотрудника Университета Квебека в Монреале Ива Жингра (Yves Gingras). Сделаем это из профиля автора.

Зайти в **Scopus** → выбрать режим **Author Search** (вкладка над поисковой строкой) → набрать в графе **Author Last Name** фамилию **Gingras**, в графе **Author Initials or First Name** указать **Y** → кнопка поиска с лупой → из результатов нас интересует первая позиция (согласно месту работы, **Affiliation**), нажать на фамилию **Gingras**, **Yves** → откроется профиль автора

В списке показателей под фамилией искомое значение **h Index**, 17.

## Пример

Найдем  $h$ -индекс в базе данных WoS вуза «Высшая школа экономики» (ВШЭ) для журнальных публикаций, вышедших с 2004 г. по настоящее время.

Зайти в WoS → в выпадающем меню справа от поисковой строки выбрать поисковое поле **Organization-Enhanced** (на будущее имейте в виду: не для всех организаций режим поиска Organization-Enhanced в WoS пока работает идеально и распознает все варианты написания названия организации) → нажать появившуюся ссылку **Select from Index** (для более эффективного поиска будем выбирать вуз из индекса, списка всех организаций) → в появившемся поисковом окне **Enter text to find organizations containing or related to the text** набрать **higher school of economics** → кнопка **Find** → получены три организации, нажать кнопку **Add** около позиции **National Research University - Higher School of Economics** → внизу страницы, в окне **Transfer your selected organization(s) below to the Organizations - Enhanced field on the search page**, появилось выбранное название вуза, нажать в этой секции страницы большую кнопку **OK** → произошло возвращение к поиску, в разделе **TIMESPAN** установить границы поиска **2004 и 2014**, а также раскрыть меню **MORE SETTINGS** и в списке **Web of Science Core Collection: Citation Indexes** оставить галочки только у журнальных баз **Science Citation Index Expanded (SCI-EXPANDED); Social Sciences Citation Index (SSCI); Arts & Humanities Citation Index (A&HCI)** → кнопка **Search** справа от поисковой строки → упорядочить результаты поиска по цитируемости, выбрав над поисковыми результатами в разделе **Sort by**: позицию **Times Cited - highest to lowest**.

Мы получили список работ авторов из ВШЭ, упорядоченный по убыванию цитируемости. При этом слева у каждой позиции в WoS указан ее порядковый номер в этом списке, а справа приводится число ссылок, полученных работой к настоящему моменту (рис. 7). Спускаясь вниз по данному списку, нам необходимо найти позицию, у которой значения слева (порядковый номер) и справа (полученные цитирования) *совпадают* (или, вернее, максимальную позицию, где число справа не меньше числа слева). В нашем случае это позиция №13. Это и есть искомый индекс Хирша: как видим, в поисковых результатах есть 13 статей, каждая из которых цитируется не менее 13 раз (это все работы с 1-го по 13-й номер), но при этом нет 14 статей, каждая из которых цитировалась бы не менее 14 раз (14-я статья цитируется, как видим, 13 раз).  $h$ -индекс = 13.

Аналогичным образом необходимо проводить расчеты  $h$ -индекса для любого множества статей:

- получить все анализируемые статьи в виде результата поиска
- упорядочить их в порядке убывания цитируемости

— исследовать, где, на какой статье, число полученных ссылок станет меньше порядкового номера публикации в упорядоченном списке  
— номер предыдущей статьи равен индексу Хирша для найденного массива статей

При этом исследуемым массивом могут быть любые документы: публикации автора, группы авторов, организации, страны, дисциплины как целого и т. д. Всё, что нужно, — это знать число полученных ссылок каждым из исследуемых документов.

<input type="checkbox"/>	11. <b>Lusztig limit of quantum sl(2) at root of unity and fusion of (1, p) Virasoro logarithmic minimal models</b> By: Bushlanov, P. V.; Feigin, B. L.; Gainutdinov, A. M.; et al. NUCLEAR PHYSICS B Volume: 818 Issue: 3 Pages: 179-195 Published: SEP 11 2009 <a href="#">Full Text</a> <a href="#">View Abstract</a>	<b>Times Cited: 15</b> (from Web of Science Core Collection)
<input type="checkbox"/>	12. <b>Refining the Theory of Basic Individual Values</b> By: Schwartz, Shalom H.; Cieciuch, Jan; Vecchione, Michele; et al. JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY Volume: 103 Issue: 4 Pages: 663-688 Published: OCT 2012 <a href="#">Full Text</a> <a href="#">View Abstract</a>	<b>Times Cited: 14</b> (from Web of Science Core Collection)
<input checked="" type="checkbox"/>	13. <b>Theory and Validity of Life Satisfaction Scales</b> By: Diener, Ed; Inglehart, Ronald; Tay, Louis SOCIAL INDICATORS RESEARCH Volume: 112 Issue: 3 Pages: 497-527 Published: JUL 2013 <a href="#">Full Text</a> <a href="#">View Abstract</a>	<b>Times Cited: 13</b> (from Web of Science Core Collection)
<input type="checkbox"/>	14. <b>Modular metric spaces, I: Basic concepts</b> By: Chistyakov, Vyacheslav V. NONLINEAR ANALYSIS-THEORY METHODS & APPLICATIONS Volume: 72 Issue: 1 Pages: 1-14 Published: JAN 1 2010 <a href="#">Full Text</a> <a href="#">View Abstract</a>	<b>Times Cited: 13</b> (from Web of Science Core Collection)
<input type="checkbox"/>	15. <b>Algebra of differential operators associated with Young diagrams</b>	<b>Times Cited: 11</b>

Рис. 7. Список статей Высшей школы экономики, упорядоченный по убыванию цитируемости (интерфейс базы данных WoS; фрагмент).

Как у любого библиометрического индикатора, у  $h$ -индекса есть и недостатки. Вот некоторые из них:

— если совершается расчет  $h$ -индекса за всё время жизни ученого, то у более возрастных авторов будет преимущество перед молодыми; такой индекс Хирша не может уменьшаться у ученого с течением времени;

—  $h$ -индекс не предполагает поправки на дисциплинарную область: сравнение абсолютных значений индекса у ученых, работающих в разных областях науки, невозможно;

— 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 13 первых статей из предыдущего примера получила бы по 1000 ссылок (а цитируемость остальных осталась прежней), на значении  $h$ -индекса это никак бы не отразилось. 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.

## 7. Список литературы

---

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

Кузнецов А. Ю., Бредихин С. В. Методы библиометрии и рынок электронной научной периодики. Новосибирск, Москва: 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, No. 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, No. 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-Anegon 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 / H. F. Moed, W. Glänzel, U. Schmoch (eds.). Dordrecht: Kluwer Academic Publishers, 2004.

Hirsch J. E. An index to quantify an individual's scientific research output // Proceedings of the National Academy of Sciences. 2005. Vol. 102. N. 46. P. 16569-16572.

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, No. 3. P. 541–550.

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

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

## Благодарность

Автор выражает признательность Павлу Арефьеву и Николаю Мазову за идею данной работы и Елене Диесперовой за плодотворное обсуждение рукописи.



## Об авторе

---

**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.

Исправление

Утверждение на с. 36

«Для определения показателей организации на фоне какой-либо другой группы (например, страны), достаточно найти индикаторы этой группы относительно всего мира и сравнить с показателями исследуемой организации относительно всего мира»

в общем случае неверно. Значения ЦНОН, определенные относительно мирового уровня, не всегда дают возможность сравнивать исследуемые научные единицы между собой. Предположим, у страны есть только два химических института (больше никаких публикаций по химии нет), при этом структура публикаций по годам и типам документов у них одинакова, но один институт имеет ЦНОН (относительно всего мира) 0,8, а второй — 0,4. Пусть ученые других научных направлений в данной стране публикуют хорошие работы, так что ЦНОН всей страны оказывается равным 1. Прямолинейное сравнение ЦНОН первого химического института (относительно всего мира) с ЦНОН всей страны ведет к предположению, что институт неэффективен в рамках государства. Однако при выборе в качестве референтной группы только данной страны, химии этой страны должны сравниваться только с химиками той же страны, что ведет к обратному выводу: на фоне однотипных публикаций данного государства первый институт, безусловно, «сильный», выше среднего уровня.

В общем случае необходим комплексный подход с учетом и дисциплин, и годов выхода публикаций, и типов документа — «ожидаемое число ссылок», знаменатель ЦНОН, следует считать по средним показателям публикаций только референтной группы.



Учебно-методическое издание

**Писляков Владимир Владимирович**  
Библиометрические индикаторы: практикум

Редактор П. Г. Арефьев  
Дизайн, верстка О. А. Иванов

Подписано в печать 20.12.2014.  
Формат 60х90/16. Бумага офсетная. Печать офсетная.  
Гарнитура Newton. Усл. печ. л. 3,75. Уч.-изд. л. 3,28.  
Тираж 150 экз. Заказ № 1109  
ТК 336600-500813-201214

ООО «Научно-издательский центр ИНФРА-М»  
127282, Москва, ул. Полянная, д. 31 В, стр. 1  
тел.: (495)-15-96, 280-33-86. Факс: (495) 280-36-29  
e-mail: books@infra-m.ru <http://www.infra-m.ru>



В практикуме рассматривается ряд библиометрических индикаторов, связанных с цитируемостью научных публикаций: импакт-фактор (двухлетний, пятилетний, совокупный, средневзвешенный, относительный), средняя цитируемость статей организации/автора, средняя нормализованная цитируемость (с нормализацией по областям науки и по журналам), SNIP, SJR, Eigenfactor, индекс Хирша и др. Для каждого показателя дается определение и демонстрируется его расчет или поиск готового значения по базам данных (Web of Science, Scopus).

Цель практикума — в доступной форме представить алгоритм расчета индикаторов и, что еще более важно, донести до читателя их «физический смысл». Демонстрируются особенности различных показателей, типичные ошибки при их расчете, даются практические рекомендации к их применению и корректной интерпретации. Значительное место занимают подробные «кнопочные» указания — как в той или иной базе данных найти или самостоятельно рассчитать соответствующий индикатор.

### **Библиометрические индикаторы: практикум**

В. В. Писляков



**НАЦИОНАЛЬНЫЙ  
ФОНД ПОДГОТОВКИ КАДРОВ**

