

РОССИЙСКИЙ ИНДЕКС НАУЧНОГО ЦИТИРОВАНИЯ

Российский индекс научного цитирования (РИНЦ) - это национальная информационно-аналитическая система, аккумулирующая более 2 миллионов публикаций российских авторов, а также информацию о цитировании этих публикаций из более 3200 российских журналов. Она предназначена не только для оперативного обеспечения научных исследований актуальной справочно-библиографической информацией, но является также и мощным инструментом, позволяющим осуществлять оценку результативности и эффективности деятельности научно-исследовательских организаций, ученых, уровень научных журналов и т.д.

Ссылка на интернет-ресурс

http://elibrary.ru/projects/science_index/author_tutorial.asp

Проверить свой индекс цитирования <http://elibrary.ru/authors.asp>

ТЕКУЩЕЕ СОСТОЯНИЕ

Показатель	Значение
Число наименований журналов:	31 449
Из них российских журналов:	6980
Число журналов с полными текстами:	6595
Из них российских журналов:	2366
Из них в открытом доступе:	1393
Общее число выпусков:	976 086
Общее число статей:	15 390 045
Общее число пристатейных ссылок:	102 168 946
Дата последнего обновления:	18.03.12
Общее число организаций:	6079
Из них зарегистрированных:	1594
Число зарегистрированных читателей:	788 521

В данном разделе приведен ряд статей из академической периодики, посвященных практике создания индексов цитирования, а также методологии их использования при оценке научной и публикационной деятельности. Надеемся, эта литература поможет усвоить прогрессивный зарубежный опыт, а также создать "дискуссионное поле" вокруг разработки и применения РИНЦ, чтобы объединить усилия всего отечественного научного сообщества перед этой задачей.

Статьи, в основном англоязычные, будут снабжены краткой аннотацией или более полным обзором на русском языке от ведущего данного раздела.

I. НАЦИОНАЛЬНЫЕ ИНДЕКСЫ ЦИТИРОВАНИЯ

L i a n s h e n g M. Document database construction in China in the 1990s: A review of developments // The Electronic Library. 2000. Vol. 18, No. 3. P. 210–215. [\[pdf\]](#)

Обзор развития рынка реферативных баз данных в Китае. В последнем разделе автор описывает возникновение в 90-е годы новых ресурсов, предоставляющих возможность анализа и оценки научных работ, — т. е. цитатных баз. Фигурируют два основных конкурента на поле китайских национальных баз данных цитирования — "Chinese Science Citation Database" и "Chinese Sci-tech Paper and Citation Statistical Database" (см. далее более распространенное название последнего ресурса).

W u Y., P a n Y., Z h a n g Y., M a Z., P a n g J., G u o H., X u B., Y a n g Z. China Scientific and Technical Papers and Citations (CSTPC): History, impact and outlook // Scientometrics. 2004. Vol. 60, Iss. 3. P. 385–397. [\[pdf\]](#)

Подробное описание создания одного из двух конкурирующих цитатных продуктов в Китае — базы данных China Scientific and Technical Papers and Citations (CSTPC), разработанной и поддерживаемой Китайским институтом научной и технической информации (фактически, китайским ВИНТИ). В базу включаются только STM-журналы.

К началу 2000-х годов CSTPC охватила более 1400 журналов из 4600 STM-изданий, выходящих в Китае. При отборе журналов учитывались следующие критерии: роспись издания в ведущих мировых и отечественных реферативных службах, мнение экспертов-предметников, наличие у издающей организации статуса национального научного общества; кроме того, использовался опыт предыдущей (продлившейся недолго) попытки создания китайского аналога SCI в университете Ланьчжоу.

За время существования CSTPC с 1988 г. число журналов в базе увеличилось более чем на 200 изданий. При этом в качестве критериев для включения новых журналов бралось качество публикуемых статей по оценкам экспертов в соответствующей области, наличие научного рецензирования в редакторском цикле, тираж издания, способность его привлечь на свои страницы видных авторов, представленность его в международных индексирующих службах и наличие в редколлегии ученых с мировым именем. Кроме того, ставилась задача

поддерживать развитие журналов в зарождающихся отраслях, особенно на стыке наук, а также включать разумное число изданий из менее развитых провинций Китая.

Для каждого журнала, включенного в CSTPC, рассчитывается общее число цитирований, импакт-фактор, индекс оперативности, доля статей, проспонсированных крупными научными фондами, процент иностранных авторов и др. Если у издания оказывались очень плохие показатели, оно могло быть исключено из базы.

В CSTPC включаются **не все** статьи из журналов. За бортом остаются такие материалы, как лекционные заметки, рефераты докладов, речи и обращения, статьи, касающиеся практического "know-how".

Статья в CSTPC будет отнесена к той организации, чей адрес в сведениях об авторах указан первым. Присвоение тематической рубрики проводится **вручную**, в соответствии со специальным государственным рубрикатором. Кроме того, публикации присваиваются отдельные атрибуты: "статья написана в рамках проекта, финансируемого крупным научным фондом", "у статьи иностранные авторы", "соавторство обусловлено международным/межрегиональным/межинститутским сотрудничеством/сотрудничеством в рамках одной организации" и др. В некоторые годы собирается также информация о задержке между подачей рукописи и выходом журнала, о возрасте, поле и занимаемой должности первого автора и т. д.

Также приведены данные по дисциплинарному составу журналов в CSTPC, количеству ежегодно поступающих статей в базу (ок. 180 тыс. в 2000 г.), ежегодному количеству цитирований, получаемых всеми статьями в базе (более 550 тыс. в 2000 г.). Размещены несколько отчетов из публикуемого на основе CSTPC "Chinese S&T Journal Citation Report".

Основная работа по конкурирующей базе Chinese Science Citation Database (CSCD) нам осталась пока недоступной: **Jin B., Wang B. Chinese Science Citation Database: its construction and application // Scientometrics. 1999. Vol. 45, Iss. 2. P. 325–332.** Однако о разнице между ресурсами можно составить представление из [комментария](#) проф. Liang. В CSCD, которую разрабатывает Центр документации и информации Китайской академии наук, меньше журналов (в 2,2 раза к 1998 г.), чем в CSTPC. Кроме того, в относительных величинах CSCD фокусируется в первую очередь на фундаментальных науках (34,5% контента), а CSTPC — на прикладных (75,2%). Приводится также таблица с точной раскладкой по дисциплинам для обеих баз.

Xin - ning S., Xin - ming H., Xin - ning H. Developing the Chinese Social Science Citation Index // Online Information Review. 2001. Vol. 25, No. 6. P. 365–369. [\[pdf\]](#)

Изложение истории создания и принципов устройства Китайского общественнонаучного индекса цитирования (CSSCI). Индекс создают и поддерживают Нанкинский университет и Гонг-конгский научно-технический университет, год первого релиза — 2000. В CSSCI

включены более 500 китайских периодических изданий общественнонаучного и гуманитарного профилей. В качестве приложения к базе публикуется отчет со статистическим анализом состояния общественных наук в Китае. Статья изобилует техническими подробностями, касающимися внутренней структуры индекса.

Jin B., Zhang J., Chen D., Zhu X. Development of the "Chinese Scientometric Indicators" (CSI) // Scientometrics. 2002. Vol. 54, Iss. 1. P. 145–154. [\[pdf\]](#)

На основе базы данных Chinese Science Citation Database (CSCD) создан продукт, позволяющий проводить статистический анализ китайской науки. Причина создания: при недостаточном финансировании китайской науки было решено активно вводить количественные оценки в процесс управления научными исследованиями. В результате к моменту написания данной статьи с библиометрией и наукометрией вынуждены считаться все, кто вовлечен в научную и научно-административную деятельность, а количественные методы "из легкого интереса превратились практически в ежедневную практику".

Китай выпускает ок. 440 тыс. статей в год (1998), из них лишь 4,5% попадает в Science Citation Index. При этом существует более 4200 журналов, из которых в SCI расписывается лишь 31 (0,7%). Поэтому национальная база цитирования была совершенно необходима, и к 1999 г. в нее уже включили 1064 китайских журналов. При этом в упомянутой [выше](#) статье Jin B. et al. "Chinese Science Citation Database: its construction and application" предложен **единый метод для оценки статей из SCI и CSCD**.

Индикаторы, предоставляемые CSI, разделены на 8 групп, и к 2001 г. их суммарное число достигло 191. Приводится описание каждой группы индикаторов и некоторые количественные данные по Китаю (например, общее число представленных организаций, общее число опубликованных ими статей и др.). Китайскими специалистами применяется подход Брэдфорда при выделении "ядра", включающего научные институты, которые печатаются активнее всех. В приближении выполняется правило Парето (80/20), а в "ядро" входит 3,4% от всех организаций, публикующих при этом более 70% от общего числа статей.

Имеется группа параметров, относящаяся к структуре соавторства, например распределение по странам среди соавторов работ китайских ученых (с разбивкой первый автор из Китая/первый автор из-за рубежа). Также в ресурсе представлены возрастные характеристики авторов китайских публикаций (автор приводит соответствующие диаграммы).

Стоит также выделить группу индикаторов, отражающих статистику, связанную с научными фондами (в первую очередь Natural Sciences Foundation of China, NSFC). Предоставляется статистика по статьям, вышедшим при поддержке того или иного фонда, рейтинги колледжей и исследовательских институтов в соответствии с числом публикаций, поддержанных NSFC, а также цитатные данные по таким публикациям (в частности выясняется, что статьи, поддержанные фондом, цитируются в среднем в 2 раза более активно, чем остальные).

Статья, содержащая больше вопросов, чем ответов, тем более что по [ссылке](#), на которой размещен описываемый ресурс, разобраться в чем-либо затруднительно. Продукт Taiwan Humanities Citation Index задуман как дополнение к упомянутым выше CSCD и CSSCI, которые фокусируются на исследователях с материка. Фокус THCI — гуманитарные исследования, опубликованные в журналах, издаваемых на острове Тайвань.

Окончательный список отобранных индексируемых изданий содержит 245 наименований (сейчас на сайте указана цифра 251), а к апрелю 2004 г. было обработано 3550 журнальных выпусков, более 35 тыс. исходных статей и 500 тыс. цитированных публикаций. При этом индексируются только исследовательские и научные статьи — за рамками индекса остаются короткие сообщения, редакторские заметки, проза, поэзия, письма, рецензии и др.

Отличительной и довольно яркой чертой работы над THCI является попытка росписи не только цитирований в "формальном" их понимании, но и просто цитат в тексте, не вынесенных в пристатейный список или в подстрочное примечание. Различные способы цитирований (их набралось шесть) расставлены разработчиками в порядке приоритета. Если в одной статье используется сразу несколько типов цитирования (например, подстрочные примечания и цитирования прямо в тексте), то в THCI фиксируются только те работы, при ссылке на которые использовался тип цитирования с более высоким приоритетом.

Кроме того, помимо обычных полей, при индексировании заполнялось поле "язык статьи" для каждой цитируемой публикации, а также поле "финансирование" (для идентификации работ, вышедших при поддержке таких источников, как Национальный совет по науке или министерство образования).

Интересный подход использован для учета перемещений авторов из организации в организацию в течение своей карьеры. Фиксируются различные записи на **автора статьи** и "соответствующего" ему **ученого**. Запись "автор статьи" хранит название организации, в которой работал автор, и должность, которую он занимал **на момент на написания работы**, а также связь с соответствующей записью ученого. Запись "ученый" хранит **текущие**, последние (судя по публикациям) данные о карьерной позиции исследователя.

В статье приводится не только структурная схема базы данных THCI, но и ряд практических правил индексации различных типов исходных и цитируемых работ. Оказывается, индексируемые поля иногда зависят от способа цитирования документа. Например, если это выноска в список литературы, то о цитируемой статье собирается начальная и конечная страницы. Если же это цитата прямо в тексте, то фиксируется только "стартовая" страница отрывка, на который идет ссылка. По всей видимости, это также означает, что несколько ссылок в тексте некоторой статьи на одну и ту же работу иногда должны быть обработаны как различные цитирования.

В завершение статьи в качестве основных сложностей, с которыми столкнулись создатели THCI, называются разнообразие стандартов цитирования в различных дисциплинах, журналах и даже статьях одного журнала; нехватка квалифицированных сотрудников, в т.ч. владеющих иностранными языками; неверные или неполные ссылки на источники; трудности рубрикации индексируемых работ.

Надо отметить, что анонсированный в статье инструмент для количественного анализа цитат [Citation Analysis Retrieval System](#) до сих пор находится в состоянии разработки и недоступен для пользователей. Также обратите внимание на ошибку в статье на рис. 1: стрелка, обозначающая цитирование, конечно, должна быть направлена в обратную сторону.

Negishi M., Sun Y., Shigi K. Citation Database for Japanese Papers: A new bibliometric tool for Japanese academic Society // Scientometrics. 2004. Vol. 60, Iss. 3. P. 333-351. [\[pdf\]](#)

В Японии к созданию собственного индекса цитирования приступили в 1995 г. Работу взял на себя Национальный институт информатики Японии, продукт в английском варианте был назван "Citation Database for Japanese Papers" (CJP). Данный ресурс индексирует только STM-издания, выходящие в Японии, и к концу 2003 г. включил в себя около 1000 наименований расписываемых журналов и ок. 620 тыс. проиндексированных статей.

При старте работ журналы выбирались по следующим критериям: издания, публикуемые обществами, зарегистрированными в Научном совете Японии (причем предпочтение отдавалось более крупным обществам); издания, публикующие оригинальные статьи; поддержка баланса между дисциплинарными направлениями, чтобы более равномерно покрыть фундаментальные и прикладные науки.

Ввод ссылок в базу осуществлялся вручную, причем ему предшествовала работа по выделению цветным маркером (!) различных полей в библиографическом описании (автор, заглавие, том и т. д.).

Приводятся две сложности, с которыми столкнулись создатели индекса: 1) Авторы статей нередко указываются в публикациях в различных форматах — либо только фамилия, либо фамилия и имя (кроме того, в японском нет пробела, отделяющего одну составляющую от другой). Ввод имен осуществлялся **в том виде, в каком они приведены в статье**, что усложняет работу пользователя по поиску. 2) Аналогичная сложность с сокращениями наименований цитируемых журналов, которые также варьируются в зависимости от стандартов, принятых в каждом издании. Ввод наименований в базу, насколько можно понять, также осуществляется "as is", однако различным вариантам сокращений присваивается идентификатор на основе "Объединенного каталога периодических изданий", что позволяет пользователям отождествлять журналы.

Отметим, уже вне рамок статьи, что в настоящий момент [поиск по японскому индексу цитирования](#) находится в свободном доступе, однако для просмотра библиографической или цитатной информации требуется подписка (индивидуальная или организации). Объем базы превысил 920 тыс. записей и 10 млн. цитируемых публикаций. Уникальной возможностью является предоставляемый **совместный поиск по CJP и базам данных ISI** (возможный для подписчиков обоих продуктов). При этом, однако, ISI не позволил создателям настроить ссылки из своих записей на статьи в журналах, индексируемых в CJP. Сейчас разработчики РИНЦ также ищут возможность интеграции сведений из международных индексов цитирования с создаваемой отечественной базой.

II. МЕТОДЫ ИСПОЛЬЗОВАНИЯ БАЗ ДАННЫХ ЦИТИРОВАНИЯ

Rousseau R. L. Journal Evaluation: Technical and Practical Issues // Library Trends. 2002. Vol. 50, Iss. 3. P. 418–439. [\[pdf\]](#)

Статья живого классика наукометрии прекрасно подходит для начального ознакомления с методами оценки качества научных журналов и научной деятельности. Имеются словарь терминов и приложение, куда вынесены наиболее сложные уравнения.

Вводятся показатели и методы:

1. "Классический" (Гарфильдовский) импакт-фактор.

$$[CIT(Y, Y-1) + CIT(Y, Y-2)] / [PUB(Y-1) + PUB(Y-2)],$$

где CIT(X, Y) — суммарное число цитирований, полученных в году X всеми статьями, вышедшими в определенном журнале в году Y; PUB(X) — суммарное число статей, опубликованных в данном журнале за год X.

2. "Индекс оперативности" (Immediacy index).

$$CIT(Y, Y) / PUB(Y)$$

Как много цитирований получает журнал в том же году, за который мы рассматриваем опубликованные статьи (в расчете на одну статью).

3. "Диахронный" импакт-фактор.

$$[CIT(Y+2, Y) + CIT(Y+1, Y)] / PUB(Y)$$

В отличие от пункта 1, который определяет "синхронный" импакт-фактор, диахронный показатель учитывает цитирования, которые получают статьи журнала, вышедшие в **фиксированном году**. Такой показатель, пишет автор, "явно предпочтительнее". Как пункт 1, так и пункт 3 могут быть с легкостью распространены на большее "окно цитат": часто двухлетний интервала недостаточно. Кроме того, исходный год, за который считается импакт фактор, может учитываться или (как в приведенных формулах) не учитываться. В первом случае добавилось бы слагаемое CIT(Y, Y).

4. Коэффициенты самоцитируемости и самоцитирования.

Коэффициент самоцитируемости журнала = [число цитирований в журнале на статьи, опубликованные в нем же] / [общее число цитирований данного журнала всеми изданиями]

Коэффициент самоцитирования журнала = [число цитирований в журнале на статьи, опубликованные в нем же] / [общее число цитирований в данном журнале всех изданий]

(все цитирования в обоих случаях берутся на определенном временном интервале)
Высокий коэффициент самоцитируемости журнала говорит о его малой заметности.
Высокий коэффициент самоцитирования журнала говорит об изолированности, обособленности покрываемой им дисциплины.

5. По поводу дисциплинарного импакт-фактора сделано важное замечание. Существуют две методики его подсчета: одна учитывает цитирования из всех журналов (как делает ISI), другая — только из журналов, отнесенных к той же дисциплине. Рекомендуется иметь в виду оба метода. Недостаток первого — существует возможность, что журнал получает больше ссылок из изданий не своей тематики. Недостаток второго — поражение в правах тех изданий, которые стремятся играть роль связок между различными дисциплинами.

6. При оценке эффективности работы научных коллективов рекомендуется обратить внимание на два показателя:

отношение среднего числа цитирований на одну статью у авторов группы к среднему числу цитирований тех журналов, в которых они публиковались (также при расчете на одну статью) отношение среднего числа цитирований на одну статью у авторов группы к среднему числу цитирований по всей дисциплине, в которой они работают (опять же при расчете на одну статью)

7. При вычислении импакт-фактора целой дисциплины признается более корректным считать не средний импакт-фактор всех журналов данного направления, а импакт-фактор этих журналов взятых как целое (т.е. брать не среднюю величину отношений, а отношение средних величин): всего получено цитат / всего вышло публикаций во всем массиве журналов.

8. Упоминается "эффект Матфея", благодаря которому, в частности, "многие страны, от которых не ожидают хороших показателей цитирования (на основании импакт-факторов публикующих их журналов), получают цитирований даже меньше, чем можно было бы ожидать". Сильнее всего страдают от такого эффекта, например, Китай, **постсоветские страны**, Нигерия. Обратите на опечатку в формуле 7 в приложении: в числителе вместо $Y-1$ следует читать $Y-i$.

Bordons M., Fernandez M. T., Gomez I. Advantages and limitations in the use of impact factor measures for the assessment of research performance in a peripheral country // Scientometrics. 2002. Vol. 53, Iss. 2. P. 195-206. [\[pdf\]](#)

Важное достоинство статьи — представленный взгляд из периферийной (в научном плане) страны и обсуждение тех же проблем, которые встанут при использовании РИНЦ. Авторы пишут об испанском опыте.

Приведен конкретный пример "эффекта Матфея" (ср. его [упоминание](#) в статье Р. Руссо): "Число цитирований, которое получает статья, написанная автором из периферийной страны, часто ниже, чем можно было бы ожидать на основании импакт-фактора опубликовавшего ее журнала, как Ками установил для испанской биомедицины. На промежутке 1995–1999 гг. Испания превышала средний импакт-фактор только для 5 дисциплин из 20. Вероятно <...> вмешивается социологический фактор: ученые предпочитают ссылаться на работы авторов из ведущих стран, которые считают более надежными". Авторы делают вывод: при оценке заметности периферийной страны в мировой науке, возможно, **предпочтительнее считать "ожидаемый" импакт-фактор** (на основе журналов, публиковавших авторов данной страны), **а не подсчитывать сами цитаты, реально полученные публикациями**. В первом случае мы менее зависимы от "социологической" компоненты.

Строятся два показателя для составления картины научной деятельности регионов страны:

индекс активности = отношение доли определенной дисциплины в общем массиве публикаций региона к доле этой же дисциплины в в общем массиве публикаций всей страны

относительный импакт-фактор = отношение среднего ожидаемого импакт-фактора публикаций региона по определенной дисциплине к среднему ожидаемому импакт-фактору публикаций всей страны по данной дисциплине

Если первый показатель больше единицы, то активность данного региона в данной дисциплине выше, чем в среднем по стране. Если второй показатель больше единицы, то данный регион печатает публикации по соответствующей дисциплине в «более авторитетных» изданиях, чем это делается в среднем по стране.

Для оценки деятельности мультидисциплинарных научных центров вводится "нормированная позиция среди журналов" (normalised journal position), своего рода упрощенный вариант нормализованного по дисциплине импакт-фактора (нормализуются не значения импактов, а места в рейтинге). Для этого берется список журналов ISI по определенной дисциплине, выстроенный в порядке убывания импакт-фактора. Определяется положение, которое занимал бы в этом рейтинге вымышленный журнал с импакт-фактором, равным среднему импакт-фактору публикаций данного центра по данной дисциплине. Номер этой позиции делится на суммарное число журналов в списке по дисциплине и результат вычитается из единицы. Полученный нормированный показатель меняется от 1 до 0 и, чем выше он, тем солиднее деятельность центра в данной дисциплине на фоне всего мира. Этот параметр — один из способов сравнивать между собой результаты деятельности организации в различных отраслях (для мультидисциплинарных научных центров).

При оценке центров рекомендуется совмещать показатели, построенные на основе импакт-факторов, с основывающимися просто на цитированиях — например, суммарное число ссылок на публикации центра или процент цитировавшихся хотя бы раз статей.