

УДК 002.52+02:004
ББК 78.30+78.34(3)

БАЗЫ ДАННЫХ STN INTERNATIONAL В БИБЛИОТЕЧНОЙ РАБОТЕ: СИСТЕМНЫЙ ПОДХОД

© И. В. Зибарева, 2010

*Новосибирский институт органической химии им. Н. Н. Ворожцова
Сибирского отделения Российской академии наук,
630090, г. Новосибирск, пр. Лаврентьева, 9*

Рассмотрено системное использование баз данных международной научно-технической сети STN International, позволяющее научно-техническим библиотекам эффективно решать комплексные задачи, относящиеся к информационному обеспечению научных исследований.

Ключевые слова: базы данных, международная научно-техническая сеть STN International, Chemical Abstracts, Science Citation Index, библиометрические исследования, информационное обеспечение научных исследований.

The system use of STN International databases, that allows scientific and technical libraries to solve their complex tasks on information provision of scientific research, is shown.

Key words: databases, Scientific and Technical Network International, Chemical Abstracts, Science Citation Index, bibliometric studies, information provision of scientific researches.

Общемировая тенденция последнего времени – превращение научно-технических библиотек из упорядоченных хранилищ сведений в высокотехнологичные информационные центры, способные помимо традиционного обслуживания читателей решать комплексные задачи, связанные с информационным обеспечением научных исследований. Сюда относится, в частности библиометрическое изучение науки с применением библиографических баз данных (БД) [1]. Для естественных наук исследования в основном выполняются с помощью политематической БД Science Citation Index (SCI) и иных информационных ресурсов на ее базе, до недавнего времени являвшихся основным источником сведений о цитировании. В последние годы появились работы по применению, наряду с БД SCI, специализированных БД, в частности БД Chemical Abstracts (CA), для анализа цитирования научных публикаций по химии [2, 3].

Данная тенденция охватывает, естественно, и Россию, встречаясь, однако, с некоторыми специфическими трудностями. Так, БД производства ВИНТИ РАН не содержат встроенных компьютерных программ обработки данных, что затрудняет их использование, а создание БД Российский индекс научного цитирования (РИНЦ) еще не завершено. Тем не менее, отечественные специалисты, имеющие доступ к глобальным библиографи-

ческим БД, успешно решают разнообразные и весьма сложные библиометрические задачи [4–10], несмотря на то что российская наука представлена в таких БД заведомо неполно и, из-за языковых различий, с неизбежными ошибками. Основные усилия направлены как на информационную поддержку отечественной науки, так и на ее комплексный мониторинг (в целом или отдельных предметных областей). Как правило, исследования проводятся в крупных научно-технических библиотеках (БЕН РАН, ГПНТБ СО РАН) и научных информационных центрах (ВИНИТИ РАН). В частности в ГПНТБ СО РАН выполнены разнообразные библиометрические исследования различных проблемных направлений региональной науки. Основные результаты представлены в коллективной монографии [11], сборнике научных трудов [12] и отдельных обобщающих публикациях (например, [13]).

Дальнейшее развитие библиометрических исследований отечественной науки в научно-технических библиотеках будет, безусловно, характеризоваться все более широким использованием БД, их кластеров и сетей, и возникающих при этом новых возможностей. Необходимо отметить, что БД, наряду с читателями и библиотечно-информационными специалистами, – один из элементов системы документальных коммуникаций, соединенных как прямыми, так и обратными связями.

Подход к БД как элементу системы документальных коммуникаций предполагает не только их утилитарное использование для получения требуемых сведений, но и изучение особенностей представления информации и поисково-аналитического аппарата БД, что, несомненно, будет способствовать повышению их эффективности и совершенствованию работы библиотек.

В этом контексте особый интерес вызывает международная научно-техническая сеть Scientific and Technical Network (STN) International, объединяющая более 200 БД с едиными поисково-аналитическими возможностями [14]. В СО РАН эта сеть уже активно используется для комплексного библиометрического исследования отечественной науки, в первую очередь химической. Работы проводились, в основном, по трем взаимосвязанным направлениям: 1) получение величин библиометрических индикаторов, характеризующих различные аспекты отечественной химической науки; 2) изучение особенностей БД сети STN в ее контексте; 3) выявление возможностей сети в информационном обеспечении отечественных научных исследований. Цель данной статьи – обсуждение полученных результатов работы.

Библиометрические исследования

Результаты библиометрических исследований во многом зависят от используемых источников первичной библиографической информации и критериев отбора релевантных публикаций. Наиболее сложным объектом являются, по-видимому, химические науки с их огромным и продолжающим быстро возрастать объемом библиографических данных [15] – так, в 2006 г. в БД СА было учтено около 1 000 000 новых документов [16]. В случае отечественной науки основная проблема при использовании зарубежных БД состоит в достижении максимально возможной полноты исходной информации. Сеть STN позволяет проводить поиск одновременно в нескольких БД. Их совместное применение с различной номенклатурой реферируемых источников и способами индексирования обеспечивает всесторонний охват информационных массивов и, следовательно, получение достоверных величин библиометрических индикаторов. Кроме того, поисково-аналитические возможности сети STN позволяют легко получить ранговые распределения, например, по авторам, типам публикаций, ключевым словам и др.

Показательным примером могут служить библиометрические исследования отечественных химических наук, основанные на совместном применении БД SCISearch и БД СА сети STN [13, 17]. Информационные массивы, представляющие различные иерархические уровни организации науки –

отдельные ученые, исследовательские институты, научные журналы – позволили выработать стратегию получения достоверных значений разнообразных библиометрических индикаторов, характеризующих отечественные научные исследования по ряду аспектов, в том числе по продуктивности и цитируемости (причем как с учетом самоцитирования, так и без него). Индикаторы соответствовали системе государственной оценки результативности научных организаций и системе показателей результативности научной деятельности сотрудников РАН, а также принципам формирования и критериям отбора научных журналов в «Перечень» ВАК. Полученные библиометрические данные, таким образом, могут служить информационной базой – как для оценки научного потенциала, так и для принятия решений, связанных с планированием и финансированием исследований.

При изучении цитируемости российских публикаций по химии показано, что хотя количественные данные по цитируемости с 1996 г. в БД СА и БД SCISearch согласуются, их совместное использование обеспечивает более полную и объективную картину. Исходя из данных по цитируемости, вычислены импакт-факторы (ИФ) ряда отечественных журналов, причем для изданий, не учитываемых БД Journal Citation Reports (JCR), впервые. Обнаружено влияние на величины библиометрических показателей журналов запаздывания реферирования в БД их переводных версий по сравнению с оригинальными. Для ведущего отечественного журнала «Успехи химии» результаты, полученные при совместном использовании БД СА и БД SCISearch, позволили заключить, что его показатели периода 1996–2006 гг., представленные в БД Essential Science Indicators (ESI), существенно занижены. Это свидетельствует о целесообразности независимого вычисления значений величин библиометрических индикаторов, характеризующих отечественную науку, особенно для сравнения и принятия административных решений.

Важно, что поисково-аналитический аппарат сети STN позволяет оперативно обрабатывать большие тематические информационные массивы. Примером может служить проводимое в настоящее время при поддержке Российского фонда фундаментальных исследований (РФФИ) комплексное библиометрическое изучение российской химической науки начала XXI в. Так, в БД СА найдено около 133 000 российских публикаций 2001–2005 гг., в среднем более 26 500 в год; в БД SCISearch за этот же период – около 47 000 тыс. публикаций, в среднем около 9400 в год. Неидентичными (уникальными) по двум БД являются более 143 000 российских публикаций, в среднем более 28 650 в год. Из полученных данных следует, что среднегодовое число уникальных публика-

ций, найденных в обеих БД, почти на 2000 превышает число, основанное лишь на одной БД СА.

Рассмотренный комплексный подход к использованию БД сети STN в библиотечно-информационных исследованиях применим не только к химическим дисциплинам, но и ко всем другим областям науки и техники, охватываемым сетью. Так, для библиометрического изучения физики может быть использована БД INSPEC; биологии – БД BIOSIS, биомедицины – БД MEDLINE, техники – БД COMPENDEX, и т. д. Патентные БД сети STN, в том числе БД российских патентов RUSSIPAT, позволяют изучать инновационный потенциал исследований.

Следует подчеркнуть, что такой подход абсолютно необходим при библиометрическом изучении нанонауки и нанотехнологии – недавно возникшей мультидисциплинарной области (ННТ). Ее изучение потребовало совместного применения трех БД – SCISearch, СА и INSPEC. Одной из них, даже политематической БД SCISearch, явно недостаточно. В случае отечественных исследований лишь 26,5% (т. е. меньше трети) неидентичных публикаций по ННТ найдено во всех указанных БД или в любых двух из них.

В целом можно заключить, что при корректном использовании БД сети STN International – надежный источник достоверных значений различных библиометрических индикаторов, характеризующих текущее состояние, тенденции, динамику и перспективы развития отечественной науки на различных уровнях ее организации [18].

Исследование особенностей баз данных

При использовании библиографических БД для библиометрических целей необходимо иметь в виду, что большинство БД создавалось лишь для упорядоченного хранения и быстрого поиска информации – их применение в библиометрических исследованиях не предполагалось. Связанные с этим основные проблемы – отсутствие стандартизации библиографических данных и недостаточная полнота информации в некоторых полях [17]. В ряде случаев это требует тщательного ручного редактирования первичных результатов, в том числе с использованием независимой информации. Библиометрические исследования отечественной науки с помощью БД сети STN International позволили, среди прочего, всесторонне изучить используемые БД и выявить особенности представления в них отечественных публикаций.

На представительных примерах отдельных ученых, научных институтов и журналов показано, что совместное применение БД SCISearch и БД СА сети позволяет составить практически полные списки оригинальных научных публикаций, про-

веренных с помощью официальных библиографических указателей трудов ученых, годовых отчетов институтов и данных редакций журналов. Важно, что кроме монографий и статей в так называемых центральных журналах найдены публикации в региональных (местных) изданиях, сборниках трудов молодых ученых, а также депонированные в ВИНТИ рукописи, тезисы докладов на российских конференциях и отчеты.

При геолокации отечественных публикаций обнаружено, что поле «адрес» систематически используется для них только с 1960-х гг., охватывая 50–60% документов. В ряде случаев в качестве адреса указана только страна – СССР.

Изменение названий приводило в ряде случаев к прекращению учета цитируемости отечественных журналов в БД SCISearch, и без того сравнительно немногочисленных в ней. Это показывает, насколько продуманной должна быть издательская политика.

Там, где возможно прямое сравнение, показано, что сведения БД SCISearch и БД СА по цитируемости отечественных публикаций количественно согласуются. В изученных примерах лишь менее чем для 5% публикаций ссылки найдены только в одной из БД. В целом в БД SCISearch обнаружено больше ссылок из журналов на русском языке, чем в БД СА, а также найдены ссылки из журналов на китайском языке, отсутствующие в БД СА. При этом в ней отсутствовали ссылки из отечественных журналов, для которых реферируются оригинальные версии. Таким образом, совместное использование БД обеспечивает более полную и объективную картину цитирования отечественных публикаций.

Определенной проблемой в контексте российской науки являются многочисленные и разнообразные технические ошибки в БД. Часть из них – неизбежное следствие языковых различий: русскоязычные публикации реферируются в англоязычных БД. Часть – результат ошибок в первичной литературе, внесенных авторами публикаций, например, в виде библиографически некорректных ссылок. Затем следуют технические ошибки, сделанные либо издателями журналов, например, в виде типографских опечаток, либо производителями БД, например, в виде ошибочной геолокации (так, институт СО РАН, находящийся в Новосибирске, может быть приписан к Москве) и т. д. [17].

Интересной специфической проблемой, связанной с отечественной наукой, с недавних пор стало двойное цитирование в БД российских публикаций [13]. Ведущие российские журналы издаются, как правило, в двух версиях – оригинальной и переводной. Некоторые международные журналы при цитировании публикаций в таких изданиях требуют давать, как одну, ссылку на обе версии.

В БД, однако, каждая подобная ссылка рассматривается как две, что завышает формальные библиометрические показатели отечественных авторов и организаций.

Безусловно, все отмеченные проблемы и ограничения необходимо учитывать при библиометрическом изучении отечественной науки с использованием зарубежных библиографических БД.

Новые аспекты информационной работы научных библиотек

Имеющийся опыт использования сети STN International позволяет сделать некоторые выводы о применимости БД и поисково-аналитических функций сети не только для библиометрических исследований, но и для решения разнообразных задач, направленных на повышение эффективности информационного обеспечения науки. Среди них – оптимизация поиска исходной библиографической информации, проверка ее на полноту и адекватность. Решение этих задач полезно для информационного сопровождения исследований на различных уровнях их организации.

Регулярно работая с БД, библиографы имеют возможность досконально изучить их свойства с тем, чтобы, рекомендуя БД пользователям (прямая связь), объективно характеризовать ее возможности, преимущества, недостатки и ограничения. При этом они могут сообщать о замеченных недостатках, ограничениях и ошибках производителям БД (обратная связь), работая, тем самым, на ее улучшение. Контакт с издателями (производителями первичной информации), также важен, поскольку библиографические БД создаются на основе их продукции.

Поисково-аналитические возможности современных БД позволяют получать различную дополнительную информацию: распределение публикаций по видам; геолокация авторов; списки наиболее активных и (или) цитируемых авторов и (или) институтов и т. д. Такие распределения представляют собой аналитико-синтетическую обработку результатов поиска. При этом происходит прибавление ценности («value-adding») к информации, причем в удобном виде для дальнейшего анализа полезного как для потребителей информации (читателей, ученых), так и для самих библиотечных специалистов – при подборе литературы для выставок, создании библиографических пособий, тематических справок и т. д. У этой работы есть и чисто библиотечный аспект – например, ее результаты позволяют принимать решения о комплектовании научно-технических библиотек, не полагаясь исключительно на ИФ журналов из БД JCR, тем более что многие российские (особенно региональные) издания в ней не учитываются.

Результаты работы могут представлять интерес не только для библиотечно-информационного общества. Так, библиометрический анализ журналов – часть библиографического источниковедения. Выработка рекомендаций по повышению профессионального престижа журналов полезна для их издателей; анализ публикаций (цитируемости) ученых и научных коллективов – для соответствующих администраций и органов, формирующих научную политику, в том числе региональную.

В заключение уместно кратко коснуться общих перспектив. Можно надеяться, что комплексные библиометрические исследования науки в российских научно-технических библиотеках с использованием зарубежных библиографических БД получат дальнейшее развитие. Этому, в частности, способствует возрастающая доступность таких БД – SCI (Web of Science), JCR, ESI, производимых Thompson Scientific (ранее Institute for Scientific Information); БД Scopus, производимой Elsevier, и некоторых других. Этим исследованиям также будет содействовать появление встроенных программ обработки данных в БД производства ВИНТИ, открытие общего доступа к БД РФФИ, а также завершение работ по БД РИНЦ. Научно-техническим библиотекам целесообразно проводить систематическую оценку новых электронных ресурсов, изучать возможности БД для выработки обоснованных рекомендаций по их использованию.

Список литературы

1. *Wormell I.* Databases as analytical tools // *Encyclopedia of Library and Information Science.* – N. Y., 2000. – Vol. 70, suppl. 33. – P. 77–92.
2. *Ridley D. D.* Citation searches in on-line databases: possibilities and pitfalls // *Trends in Analyt. Chemistry.* 2001. – Vol. 20. – P. 1–10.
3. *Whitley K. M.* Analysis of Scifinder Scholar and Web of Science citation searches // *J. of Amer. Soc. fo Inform. Science Technology.* – 2002. – Vol. 53, N 14. – P. 1210–1215.
4. *Wilson C. S., Markusova V. A.* Changes in the scientific output of Russia from 1980 to 2000, as reflected in the Science Citation Index, in relation to national politico-economic changes // *Scientometrics.* – 2004. – Vol. 59, N 3. – P. 345–389.
5. *Маршакова-Шайкевич И. В.* Вклад России в развитие мировой науки: библиометрическая оценка [Электронный ресурс] // *Отч. зап.* – 2002. – № 7. – URL: http://magazines.russ.ru/oz/2002/7/2002_07_39.html
6. *Ефременкова В. М., Севастьянов В. Г.* Информационное сопровождение научных исследований по карбиду кремния // *Науч.-техн. информ. Сер. 1.* – 2004. – № 9. – С. 16–27.
7. История, состояние и прогноз работ по нейросетевым исследованиям на основе статистического анализа баз данных INSPEC, SCI, ВИНТИ и CAS / В. М. Ефременкова [и др.] // *Науч.-техн. информ. Сер. 1.* – 2003. – № 9. – С. 21–30.

БИБЛИОГРАФОВЕДЕНИЕ

8. Анализ некоторых информационных ресурсов по астрономии и исследованию космического пространства / В. М. Ефременкова [и др.] // НТИ-2007 : 7 Междунар. конф., посвящ. 55-летию ВИНТИ, «Информ. о-во. Интеллектуал. обраб. информ. Информ. технологии», Москва, 24–26 окт., 2007 : материалы конф. М., 2007. – С. 122–126.
9. *Слащева Н. А.* Наукометрические исследования в библиотеке (на примере Центральной библиотеки ПНЦ РАН) // Наукоеведение. – 2002. – № 3. – С. 147–154.
10. *Слащева Н. А., Мохначева Ю. В.* Электронная информация в наукометрических исследованиях // Науч.-техн. информ. Сер. 1. – 2003. – № 5. – С. 21–27.
11. Информационные ресурсы региона: принципы формирования, организационная, видовая структура / Отв. ред. Е. Б. Соболева ; Гос. публич. науч.-техн. б-ка Сиб. отд-ния Рос. акад. наук. – Новосибирск, 2000. – 167 с.
12. Состояние и развитие основных направлений научных исследований: библиометрический анализ : сб. науч. тр. / Гос. публич. науч.-техн. б-ка Сиб. отд-ния Рос. акад. наук ; редкол.: Е. Б. Соболева (отв. ред.) и др. – Новосибирск, 2003. – 183 с.
13. Некоторые аспекты библиометрических исследований в Сибирском отделении Российской академии наук / И. В. Зибарева [и др.] // Библиотековедение. – 2008. – № 3. – С. 39–45.
14. STN International [Electronic resource]. – URL: <http://www.stn-international.de>
15. Поиск химической информации : справ. рук. по использованию традиц. и компьютер. средств / В. М. Потапов [и др.]. – М. : Изд-во МГУ, 1990. – 174 с.
16. CAS Statistical Summary [Electronic resource]: 1907–2007. – URL: <http://www.cas.org/ASSETS/836E3804111B49BFA28B95BD1B40CD0F/casstats.pdf>
17. *Зибарева И. В., Круковская Н. В.* Опыт использования библиографических БД для наукометрических исследований российской химической науки [Электронный ресурс] // Образоват. технологии и о-во. – 2007. – № 1. – С. 297–303. – URL: http://ifets.ieee.org/russian/depository/v10_i1/pdf/30.pdf
18. *Зибарева И. В.* Разработка модели использования баз данных сети «STN International» в библиометрических исследованиях отечественной науки : дис. канд. пед. наук / Гос. публич. науч.-техн. б-ка Сиб. отд-ния Рос. акад. наук. – Новосибирск, 2009. – 201 с.

Автор благодарит РФФИ за финансовую поддержку (проект № 08-06-00337).

Материал поступил в редакцию 16.11.2009 г.

Сведения об авторе: *Зибарева Инна Владимировна* – руководитель Центра международной научно-технической сети STN International, тел.: (383) 330-96-62, e-mail: zib@nioch.nsc.ru