

СПРАВОЧНО-ИНФОРМАЦИОННЫЙ РАЗДЕЛ

УДК 929:002

Р. С. Гиляревский, А. И. Чёрный



32-35

Доктор Юджин Гарфилд: научно-информационная деятельность

Приводятся основные сведения о созданной Ю. Гарфилдом крупной информационной системе — Указателе библиографических ссылок по естественным наукам (ныне — *Web of Science* и *Web of Knowledge*).

Ключевые слова: указатель цитирования, импакт-фактор, *Thomson-Reuters*, *E. Garfield*

Сегодня имя Юджина Гарфилда знакомо каждому серьезному ученому. Значимость его научных трудов оценивается по показателям, которые именно он ввел в научный оборот. Для этого он создал и наладил регулярный выпуск *Указателей научного цитирования (Science Citation Index — SCI)*, которые позволяют оценивать труды любого ученого по степени их цитирования другими учёными, т. е. по фактическому вкладу ученого в мировую научную литературу. И тем более имя Ю. Гарфилда знакомо каждому, кто работает в области научной коммуникации, информатики и науковедения, так как без использования его идей и трудов эти области науки уже не могут развиваться.

Ю. Гарфилд является автором более 200 научных статей, нескольких книг и патентов. Он член более 20 американских и зарубежных научных и профессиональных академий и ассоциаций, член редколлегий ряда научных журналов, почетный профессор нескольких американских и зарубежных университетов и институтов, награжден медалями Дж. П. Ветерилла и Д. С. Прайса. В России Ю. Гарфилд является почетным доктором наук Дальневосточного государственного университета (с 1997 г.) и иностранным членом Российской академии естественных наук (с 1996 г.).

Первоначально Ю. Гарфилд намеревался стать химиком и поэтому начал изучать эту науку в Колорадском университете (г. Боулдер) и в Калифорнийском университете (г. Беркли). В 1948 г. он получил диплом бакалавра по химии в Колумбийском университете (г. Нью-Йорк). Там же он получил в 1954 г. диплом магистра по библиотекведению и стал работать химиком и консультантом по химической информации в фармацевтической фирме *Evans Research and Development Corp.* Свое академическое образование Ю. Гарфилд завершил в Пенсильванском университете (г. Филадельфия), где в 1962 г. ему была присвоена ученая степень доктора наук по структурной лингвистике. Название его докторской диссертации — «Алгоритм перевода химических наименований в молекулярные формулы».

В область информатики Ю. Гарфилд вступил в 1951 г., когда стал работать над Уэлчским проектом по индексированию медицинской литературы (*Welch Medical Indexing Project*) в Университете им. Джона Хопкинса (г. Балтимор, шт. Мэриленд). Заказчиком этого проекта была Армейская медицинская библиотека США, которая позже была переименована в Национальную медицинскую библиотеку. Армейская библиотека выпускала печатный библиографический указатель *Current List of Medical Literature*. Задача состояла в том, чтобы значительно ускорить составление и выпуск этого указателя посредством применения перфокарт машинной сортировки. Для этого нужно было переработать список применявшихся предметных заголовков, а их было более 30 тысяч!

Работа над Уэлчским проектом позволила создать теоретическую и практическую основу для разработки первой в США комплексной Системы анализа и поиска медицинской литературы (*Medical Literature Analysis and Retrieval System — MEDLARS*) в Национальной медицинской библиотеке США (1964 г.), которая стала выпускать ежемесячный указатель *Index Medicus*.

Тогда же Ю. Гарфилд обратил внимание на огромнейшую роль периодических изданий в распространении научной информации и понял, что для информационной поддержки ученых необходимо быстро доводить до них оглавления тщательно подобранных журналов, что весьма ограниченный по численности круг журналов обычно содержит самые важные публикации, что и в сфере научных коммуникаций также действует правило 80/20, по которому 4/5 всех важных научных статей публикуются в весьма ограниченном списке журналов, составляющих около 1/5 их общего количества. Для описания такого распределения статей в научных журналах Ю. Гарфилд выдвинул гипотезу, названную им *законом концентрации*. Подобные распределения называют также *принципом* или *законом Парето*, *законом Ципфа*, а в библиотечном деле — *законом рассеяния Брэдфорда* (1933 г.).

Ещё работая над Уэлчским проектом, Ю. Гарфилд подготовил несколько пробных выпусков

Current Contents (CC), охватывавших научную литературу по наиболее важным темам. Это были *Current Contents / Pharmaco-Medical & Life Sciences* и *Contents in Advance* (по библиотечному делу). Эти пробные выпуски показали, что Ю. Гарфилд нашел перспективное направление информационного обслуживания ученых, у него появился первый корпоративный подписчик на *CC*. Это была фирма Bell Telephone Labs, которая заказала 500 экземпляров издания по фармако-медицине и биологическим наукам.

Для выполнения работ, связанных с подготовкой и рассылкой *CC*, нужны были помощники. В 1956 г. Ю. Гарфилд создал фирму, которая первоначально называлась Documation, Inc. Но это название почти совпадало с названием другой фирмы — Documentation, Inc., созданной М. Таубом ещё в 1952 г. Поэтому новая фирма стала называться Eugene Garfield Associates — Information Engineers. Позже она была переименована в Institute for Scientific Information (ISI) и независимо существовала до 1992 г. На выбор такого названия для своей фирмы, по признанию самого Ю. Гарфилда, повлияло создание в Москве Института научной информации (ИНИ), с которым мировое научное сообщество связывало большие надежды.

В 1958 г. у Ю. Гарфилда появилась первая платная помощница — выпускница средней школы Беверли Бартоломео, которая стала его секретарем и работала в ISI 43 года. Она контролировала подготовку информационных изданий ISI и в 1995 г. ушла в отставку с должности заместителя директора ISI по изданию баз данных. К этому времени численность персонала ISI уже превысила 500 человек!

Ю. Гарфилду удалось создать в ISI творческую обстановку, которая позволила ему не только наладить оперативный выпуск сложных информационных продуктов, но и развернуть обширную научно-информационную работу.

Издание *Current Contents* — благодаря невиданной оперативности его подготовки (30–40 дней с момента выхода в свет журналов-первоисточников) — быстро получило признание мирового научного сообщества. Дальнейшее развитие бюллетеня *CC* осуществлялось за счёт всё более тщательного отбора журналов-первоисточников, более глубокой их обработки, сокращения ее сроков, а также сроков удовлетворения запросов подписчиков на копии статей.

В настоящее время *CC* вошел в состав многоотраслевой базы данных (БД) *Current Contents Connect (CCC)*, которая стала частью информационного ресурса *Web of Knowledge*. В этом ресурсе отражается содержание 8,8 тыс. журналов и около 2 тыс. книг (ежегодно). Обновление этой БД происходит ежедневно. Возможно получение *CC* в печатном виде, а также на электронных носителях и в онлайн-режиме. Эта БД может быть получена в виде подборки по следующим разделам науки и техники: биология и сельское хозяйство; искусство и гуманитарные науки; клиническая практика; инженерное дело, технология, прикладные науки; биологические науки; физика, химия и науки о Земле; общественные и поведенческие науки.

Другое выдающееся достижение Ю. Гарфилда — разработка и организация регулярного выпуска *Указателя научного цитирования (Science*

Citation Index — SCI), который очень скоро не только стал важным средством информационной поддержки ученых, но и создал фактологическую основу для развития библиографических и наукометрических исследований. Под цитированием здесь понимается не приведение дословных выдержек из текста того или иного автора, как в русском языке, а лишь библиографическая ссылка на его труд в пристатейной библиографии. Зная частоту такого цитирования опубликованных автором работ, можно определить подлинное влияние цитированного автора на развитие той или иной науки или научной дисциплины, а также вычислить ряд наукометрических и библиометрических показателей. По *SCI* можно также определить научную продуктивность ученого: круг связанных с ним других ученых; журналы, в которых он обычно публикует свои работы; оперативность доведения новой информации до широкого круга ученых; число ученых, занятых данным исследованием в той или иной стране или в мире; и еще много других показателей.

Первой публикацией Ю. Гарфилда на эту тему была статья “Указатели цитирования для науки: новое измерение в документации через ассоциации идей” в журнале *Science* 15 июля 1955 г. Но для организации регулярного выпуска *SCI* Ю. Гарфилду потребовалось еще 9 лет. За это время он создал фирму ISI, наладил выпуск бюллетеней *CC* и хорошо изучил научные журналы.

Первую серию *SCI* в 5 печатных томах — по естественным и техническим наукам — Ю. Гарфилд выпустил в 1964 г. Издание содержало 1,4 млн библиографических ссылок из 613 журналов. Уже через два года *SCI* стал доступен на магнитной ленте.

С 1976 г. ISI начал выпускать *Указатель цитирования по общественным наукам (Social Science Citation Index — SSCI)*, а с 1979 г. — *Указатель цитирования по искусству и гуманитарным наукам (Arts&Humanities Citation Index — A&HCI)*.

Библиографические ссылки на труды других авторов, которые ученые обычно приводят в своих опубликованных работах, представляют большой интерес и нуждаются в отдельном исследовании. Какие цели преследует ученый, приводящий такие ссылки? Связано ли это с желанием подчеркнуть преемственность научных исследований? Говорит ли это о большом научном значении процитированных работ? Свидетельствует ли это о повышенной популярности этих учёных? Таких вопросов можно задать множество. И на любой из них пока нет однозначного ответа.

В 2000 г. по *Arts&Humanities Citation Index* был составлен перечень фамилий из 10 авторов, работы которых за последние 7 лет (1993–2000 гг.) имели наибольшие показатели цитирования. В этот перечень вошли: К. Маркс, В. И. Ленин, У. Шекспир, Аристотель, Библия, Платон, З. Фрейд, Н. Хомский, Ф. Гегель, Цицерон.

Названные авторы попали в этот список не только потому, что они внесли наибольший вклад в развитие мировой науки или культуры, но еще по каким-то другим причинам.

Указатель цитирования как вид информационных изданий не был изобретен Ю. Гарфилдом. Издания такого типа использовались, например,

в американской юридической практике с 1873 г. Эта практика основана на принципе юридическо-го прецедента. Американским юристам необходимо быстро находить юридические решения, принятые судами по конкретным делам, — а их принимается великое множество. Чтобы облегчить поиск подходящих дел для юристов, был разработан и с 1873 г. выпускается специальный указатель, получивший название “Указатель Шепарда” — по имени его создателя. В настоящее время этот указатель, послуживший прототипом *SCI*, принадлежит фирме LexisNexis (г. Дейтон, шт. Огайо).

Об основных проблемах, которые возникали и решались при издании *CC*, а в дальнейшем — и при подготовке и издании *Указателей научного цитирования*, Ю. Гарфилд с 1962 г. и по 1972 г. писал и публиковал статьи в еженедельных выпусках *CC* под общим названием *Current Comments*. Эти статьи были опубликованы также в серии сборников (16 томов), которые названы им *Очерки информатика (Essays of an Information Scientist — Philadelphia, Penn.: ISI Press, 1962–1993)*. Сборники вошли в золотой фонд литературы по информатике и науковедению.

Кроме того, важнейшие идеи Ю. Гарфилда об указателях цитирования были опубликованы им в книге *Citation Indexing: Its Theory and Application in Science, Technology and Humanities* (Philadelphia, Penn.: ISI Press, 1979. — xiii, 274 pp.).

С 1986 г. Ю. Гарфилд стал выпускать двухнедельную газету *The Scientist*, предназначенную в основном для науковедов.

В 1992 г. Ю. Гарфилд продал ISI канадской информационно-издательской корпорации Thomson. После нескольких переименований ISI получил название Thomson Scientific. Используя информационные продукты ISI, многолетний опыт ее персонала по информационному обслуживанию ученых, а также учитывая развитие Web-технологии и электронных изданий, корпорация Thomson начала создание новой информационной системы, для чего она приобрела права на использование ряда наиболее популярных БД.

В настоящее время эта новая информационная система имеет следующую конфигурацию. Ее основу составляет компьютерная платформа Web of Science (WoS), которая представляет собой интегрированную совокупность аппаратных, программных и технологических средств, обеспечивающих более эффективное выполнение определенных функций, чем при использовании этих средств отдельно. WoS предназначена для междисциплинарного поиска текущей и ретроспективной информации, содержащейся примерно в 8,7 тыс. тщательно отобранных научных журналах. Информационный поиск можно производить по разным признакам — по библиографическим ссылкам, именам цитируемых и цитирующих авторов, ключевым словам и т. п.

Поисковые массивы WoS были дополнены описаниями примерно 850 тыс. статей из 262 журналов, выходящих в свет в первой половине XX века. Все записи снабжены ссылками на полные электронные первоисточники. WoS была открыта для использования в январе 2005 г.

В WoS используются пять основных БД, созданных еще в ISI: Science Citation Index Expanded, Social Science Citation Expanded, Arts&Humanities

Citation Index, Index Chemicus, Current Chemical Reactions.

Дальнейшим развитием WoS стала онлайн-овая информационная система Web of Knowledge (WoK), которая охватывает все отрасли науки и используется не только для поиска информации, но и для получения ответов на вопросы о конкретных научных исследованиях и исследователях. Для этого WoK была дополнена междисциплинарными и специализированными информационными ресурсами, а также средствами для анализа, оценки и управления библиографической информацией.

К междисциплинарным информационным ресурсам, кроме ресурсов WoS, относятся следующие БД: Current Contents Connect (CCC), Science Citation Index (SCI), Derwent Innovations Index (1963–).

К специализированным информационным ресурсам WoK относятся следующие БД: BIOSIS Previews (1969–), CAB Abstracts (1973–), FSTA = Food Science and Technology Abstracts (1969–), INSPEC (1969–), MEDLINE (1950–), Zoological Record (1979–).

Аналитические и оценочные средства включают следующие инструменты: Essential Science Indicators (1990–), Journal Citation Reports, Journal Use Reports, The Analyze Tools, EndNot Web, EndNot, ProCite, Reference Manager.

Напомним, что *Current Contents Connect (CCC)* — это междисциплинарная БД, содержащая оглавления 7,6 тыс. рецензируемых научных журналов, 2 тыс. книг, трудов научных конференций, а также несколько тысяч веб-сайтов, отобранных редакторами ISI. В этой БД выделено 7 дисциплинарных серий: Биологические науки — содержит оглавление 1370 журналов по 14 дисциплинам; Физика, химия и науки о Земле — содержит оглавления 1050 журналов по 10 дисциплинам; Машиностроение, вычислительная техника и технология — содержит оглавления 1100 журналов по 18 дисциплинам; Искусства и гуманитарные науки — содержит оглавления 1110 журналов по 9 дисциплинам; Общественные и поведенческие науки — содержит оглавления 1620 журналов по 15 дисциплинам; Сельское хозяйство, биология и науки об окружающей среде — содержит оглавления 1040 журналов по 12 дисциплинам; Клиническая медицина — содержит оглавления 1120 журналов по 25 дисциплинам.

Кроме того, в *CCC* приводятся оглавления двух тематических коллекций журналов — по бизнесу (содержит оглавления 235 журналов) и по электронике и телекоммуникациям (содержит оглавления 200 журналов).

Другим главным источником научной информации в WoS являются базы цитирования, которые выпускаются в трех сериях — *SCI Expanded*, *SSCI* и *A&HCI*:

- *SCI Expanded* отражает цитирование в 5800 журналах. Ежегодно в эту БД вводится 18 700 новых записей с ретроспективой до 1945 г. С 1992 г. в эту БД вводятся также авторские резюме цитируемых публикаций. Ежегодно в *SCI Expanded* вводится более 1,1 млн записей и более 23 млн ссылок по 230 дисциплинам. Обновление БД происходит еженедельно;

- *SSCI* полностью отражает цитирование в 1730 журналах, а выборочно — еще в 5800 журналах.

Еженедельно в эту БД вводится 2700 новых записей с ретроспективой до 1956 г. Авторские резюме вводятся с 1992 г.;

• *A&HCI* отражает цитирование в 1140 журналах и выборочно еще в 7000 журналах. Еженедельно в эту БД вводится 2100 новых записей с ретроспективой до 1975 г. БД охватывает 18 научных дисциплин. Авторские резюме цитирующих публикаций вводятся с 2000 г.

В системе WoK большое развитие получили также средства анализа и оценки информации. Важнейшими из таких средств являются *Отчеты о цитировании журналов*, *Существенные показатели развития науки* и *Отчеты об использовании журналов*. Далее даны их краткие характеристики.

Journal Citation Reports (Отчеты о цитировании журналов) выходят ежегодно с 1975 г. В них приводятся сведения, позволяющие судить о “научном весе” журналов, которые используются для подготовки информационных продуктов ISI. В этих отчетах приводятся следующие данные о каждом журнале: библиографическое описание журнала — полное и сокращенное название, его рубрикация, страна издания, язык публикации и т. п.; импакт-фактор журнала; индекс оперативности; индекс полужизни цитирования статей из этого журнала.

Импакт-фактор — это средняя цитируемость статей данного журнала, опубликованных за последние два года. *Индекс оперативности* — это средняя цитируемость статей в журнале в год их опубликования. *Индекс полужизни цитирования* — это число лет, на которое приходится 50% ссылок на статьи данного журнала. По этим показателям ежегодно вычисляются и публикуются характеристики процитированных журналов. По импакт-фактору оценивается “научный вес” журнала. В БД JCR-Science Edition за 2000 г. наибольший импакт-фактор имел журнал *Annual Review of Biochemistry* — 50,340. На шестом месте находился журнал *Nature* (25,814), а на седьмом — журнал *Science* (23,852). Около трети журналов имели импакт-фактор меньше 0,5. В JCR за 2005 г., в котором было процитировано 6086 журналов, импакт-фактор находился в пределах от 49,749 до 14,325.

Essential Science Indicators (ESI) — *Существенные показатели развития науки* охватывают 10 млн статей из более 11 тыс. журналов по 22 отраслям науки. Эта БД обновляется каждые два месяца и позволяет получать информацию о

важнейших исследователях страны, цитировании их публикаций, месте работы и т. д.; т. е. информацию, которая отражает развитие той или иной отрасли науки.

Journal Use Reports (JUR) — *Отчеты об использовании журналов* начали выходить с 2006 г., в них указываются особенности использования каждого журнала, количественные показатели его цитирования, связи с другими журналами и другие показатели, помогающие библиотекарям оценивать необходимость наличия этого журнала в фондах библиотеки.

Для доставки полнотекстовых копий документов потребителям в WoK создана служба *Document Solution*, в которой используются все возможные для этого средства — курьеры, факс, почта и т. п. Доставка копий осуществляется в форматах PDF и TIFF. Сроки доставки копий — 24 часа, в среднем они не превышают 2,5 дней. Все доставленные копии имеют правовую защиту.

В настоящее время через WoK доступны следующие источники научной информации: 23 тыс. научных журналов, 110 тыс. трудов конференций, 23 млн патентов, 9 тыс. веб-сайтов, 2 млн химических соединений, 40 млн цитированных первоисточников и 700 млн библиографических ссылок. Эти источники в ретроспективе охватывают более 100 лет. Пользователями WoK являются более 20 млн ученых и специалистов из 20 стран. Число ежедневных посещений WoK составляет 150 тысяч!

В 2008 г. Thomson Corp. приобрела одно из крупнейших в мире телеграфных агентств — Reuter, за которое она заплатила 17,34 млрд долларов. Новое объединение, получившее название Thomson-Reuter, стало по словам западных средств массовой информации — “информационным монстром” (только в агентстве Reuter работают 50 тыс. человек из 93 стран мира): объединенная фирма стала контролировать до трети мирового рынка научной, медицинской, юридической и финансовой информации!

Мы полагаем, что такого слияния не произошло бы или оно произошло бы позже, если не было бы полувека целеустремленного труда Юджина Гарфилда, его идей и веры в огромное значение научной информации для развития современного информационного общества.

Материал поступил в редакцию 19.02.09.