

# К 100-ЛЕТИЮ CHEMICAL ABSTRACTS SERVICE

УДК (051.6):54

В. М. Ефременкова, Н. В. Круковская

## 100-летний юбилей Chemical Abstracts Service: факты и цифры

*Рассматривается история возникновения и развития реферативного журнала Chemical Abstracts от первого печатного выпуска до настоящего времени. Всемирно известная организация Chemical Abstracts Service (CAS) предлагает широкий спектр информационных ресурсов, наиболее масштабными из которых являются следующие базы данных (БД): CA/CAPlus (~27 млн рефератов), CAS Registry (~31 млн органических и неорганических соединений, ~58 млн биопоследовательностей), CASREACT (~13 млн реакций по органической химии), CHEMCATS (~14 млн записей о коммерчески доступных веществах и их поставщиках), CHEMLIST (~245 тыс. промышленно важных веществ, правила их хранения и транспортировки), MARPAT (структуры Маркуша и соответствующие им патентные документы), а также проблемно-ориентированные выпуски CA SELECTs, список журналов Chemical Abstracts Service Source Index (CASSI) и многое другое. Приведены информационно-поисковые системы, через которые осуществляется доступ к БД производства CAS: STN International и SciFinder.*

В Берлине 17-18 сентября 2007 г. прошла конференция, посвященная 100-летию юбилею выпуска первого номера реферативного журнала Chemical Abstracts (CA) [1]. Президент CAS Robert J. Massie и его коллеги рассказали о пути, пройденном службой за 100 лет. Они сформулировали концепцию развития CAS от первого печатного издания к электронным продуктам и далее к идее создания в 1983 г. совместно с Германией и Японией Международной сети научной и технической информации – STN International [2, 3]. Особое внимание было уделено одному из приоритетных направлений современной науки – биотехнологии. Большой интерес вызвал доклад лауреата Нобелевской премии 2004 г. доктора А. Ciechanover “Убиквитивная протолитическая система: от основных механизмов через болезни человека к лекарственным средствам” [4].

Созданный по инициативе Американского химического общества реферативный журнал (РЖ) Chemical Abstracts (CA) стал мировым лидером среди аналогичных изданий по химии и смежным дисциплинам: химической технологии, биохимии, материаловедению, металлургии, физике, фармацевтике и т. д.

Американское химическое общество – профессиональная организация химиков, возникшая в 1876 г. при Нью-Йоркском университете, к настоящему времени насчитывает около 160 тыс. членов в США и других странах мира, имеющих ученые степени в области химии, химической технологии и других областей науки. С самого начала своего существования общество уделяло большое внимание издательской деятельности. Уже в 1879 г.

под редакцией Herman Endemann вышел 1-й номер “Journal of the American Chemical Society” (JACS). 1893 г. был отмечен появлением в журнале кратких обзоров изданных в США книг. В связи с неудовлетворенностью американских химиков отражением результатов их исследований в европейской литературе в 1895 г. редактором А. А. Noyes в JACS был введен раздел “Обзор американских исследований по химии”, имеющий отдельную нумерацию страниц [5].

К этому времени уже существовали следующие реферативные и справочные издания по химии и смежным дисциплинам, перечень которых представлен в докладе директора National Federation of Advanced Information Systems (NFAIS) Bonnie Lawlor [3]:

1817 г. – Handbuch der Anorganischen Chemie (Gmelin). Справочник по неорганической химии;

1830 г. – Pharmaceutisches Centralblatt;

1867 г. – Catalogue of Papers (Royal Society of London). Каталог журналов Лондонского королевского общества;

1879 г. – Index Medicus. Справочник по медицине;

1881 г. – Handbuch der Organischen Chemie (Beilstein). Справочник по органической химии;

1884 г. – Index Notes (Engineering Index). Справочник прикладного характера;

1889 г. – Merck Index. Справочник химических препаратов и лекарственных средств.

1 января 1907 г., как продолжение “Обзора американских исследований по химии”, возникло новое реферативное издание – “Chemical Abstracts”

(СА). Первый номер СА содержал 11847 рефератов — 7994 реферата статей из 396 журналов, 3853 патентных документов Франции, Германии, Великобритании и США. Издание было подготовлено 155 волонтерами-референтами, а его первым редактором стал William A. Noyes [3, 5].

Перед новым изданием была поставлена нелегкая задача — охватить всю мировую научную литературу по химии, химической технологии, биохимии, физике, материаловедению, фармакологии, металлургии, и т. д.

История реферирования литературы в СА начинается с определения основных понятий нового издания: **полнота, химия, реферат** [6].

**Полнота** охвата литературы. В реферативный журнал включались все новые и новые первоисточники:

в 1908 г. начали отражаться книги;  
с 1910 г. увеличивается количество патентов за счет роста числа национальных патентных ведомств в различных странах мира: 1910 г. — Канада, 1913 г. — Австрия, Дания, Швеция, Норвегия, 1918 г. — Япония, 1921 г. — Бельгия, Нидерланды, 1922 г. — Австралия, Венгрия, Италия, СССР, 1948 г. — Индия, 1953 г. — Испания, 1956 г. — Чехословакия;

в 1940 г. стали отражаться диссертации;  
с 1912 г. по 1952 г. СА ежегодно издавал списки отражаемых журналов — Journal Titles:

1912 г. — 600  
1922 г. — более 1000  
1932 г. — более 2000  
1952 г. — более 5000  
2007 г. — более 9500, из которых 1500 являются профильными (key journal).

Список профильных журналов публикуется на сайте CAS ([www.cas.org](http://www.cas.org)) в открытом доступе. Полный список отражаемых журналов по естественным наукам с указанием издателей и ретротрансформации (прежних названий журналов) — Chemical Abstracts Service Source Index (CASSI) издается как в печатном виде, так и на компакт-дисках и является платным ресурсом.

В последние полвека добавились такие виды первоисточников, как журналы, труды конференций, технические отчеты, обзоры, электронные ресурсы, препринты. На рис. 1а представлен рост числа рефератов, опубликованных в СА с 1907 г.

В 70-е годы прошлого века в машиночитаемую форму был переведен весь массив РЖ СА и создана БД СА, а позднее — CASPlus, включающая более широкий круг первоисточников, таких как тезисы трудов ежегодных собраний Американского химического общества, биографии и др. Проведена оцифровка наиболее интересных работ с 1840 по 1906 гг., в результате чего БД СА/CASPlus пополнилась еще на 48 тыс. документов. Кроме того, БД СА/CASPlus будет дополнена документами из немецкого реферативного журнала *Chemisches Zentralblatt*, прекратившего свое издание в 1969 г.

**Химия.** Даже в 1907 г. не всегда было легко отделить химическую литературу от нехимической. По словам Е. J. Crane (редактора СА с 1915 по 1958 гг. и директора с 1956 по 1958 гг.) “невозможно установить критерии и четкие правила, позволяющие легко и четко разделить публикуемые материалы на две группы — химическую и нехимическую. Это объясняется отсутствием четкой границы между химией и не химией, а также

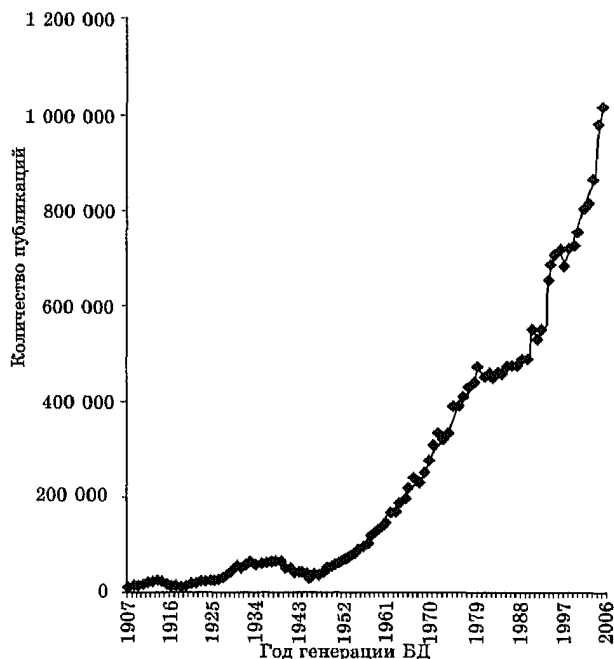


Рис. 1 Динамика потока публикаций в БД Chemical Abstracts в период с 1907 по 2006 гг.

тем, что весь реферируемый материал является новым и не легко укладывается в систему классификации, которая основывается только на том, что уже известно” [7]. Именно поэтому классификационная система СА столь четко и кратко отражает основные направления науки, техники и технологии, основываясь на широком определении химии и интересах членов американского химического общества. Если первое оглавление СА содержит 30 рубрик, последняя из которых предназначена для всех патентов, то уже в 1908 г. патенты стали помещать по их тематике в соответствующие рубрики. Классификационная схема мало менялась в течение 55 лет. Лишь в 1962 г. в связи с развитием физической химии число рубрик было увеличено до 73, в 1963 — до 74 и в 1967 г. — до 80, что отражало рост прикладной химии и химической технологии, а также развитие исследований ядерных явлений. Последнее изменение классификации было в 1974 г. в связи с ростом числа документов в области топлива и энергетики. С тех пор классификатор не менялся. Информацию о новых научных направлениях можно найти или в рубриках второго уровня, или в разделе “Index” руководства “Subject Coverage and Arrangement of Abstracts by Sections in Chemical Abstracts” [8]. Наименования кодов рубрик не всегда совпадали при изменениях классификации и для удобства пользователей был создан “CA Section Thesaurus”, позволивший установить соотношение между рубриками, которые использовались ранее. Например, для рубрики “Керамика”, история изменения кодов классификации (Classification code — CC) такова:

17 CERAMICS, 1962 ONLY/CC  
21 CERAMICS, 1963 1966/CC  
57 CERAMICS, 1967 TO PRESENT/CC.

С 1974 г. 80 рубрик были сгруппированы в пять родственных по тематике разделов: Биохимия (рубрики 1–20), Органическая химия (рубрики 21–34), Химия высокомолекулярных соединений (рубрики 34–46), Прикладная химия и химическая технология (рубрики 47–64), Физическая и аналитическая

химия (рубрики 65–80). С момента появления разделов в 1962 г. по этим пяти группам рубрик стали издаваться отдельные выпуски РЖ.

**Реферат.** В процессе развития СА большое внимание уделялось оттачиванию определения “реферат”. В статье директора СА с 1958 по 1986 гг. D. V. Baker описано три подхода к процессу реферирования: 1) сокращенный вариант статьи; 2) описание работы – краткие сведения о статье или указание на тему статьи и 3) “поисково-ориентированные” рефераты [6].

В начале своего существования в СА помещались рефераты, представляющие собой сокращенный вариант статьи и способные, в большинстве случаев, заменить первоисточник. Особенно подробными были рефераты по органической и аналитической химии. Это позволяло читателям преодолевать языковой барьер и быть в курсе последних научных достижений. Высокий уровень реферирования обеспечивался четкими инструкциями для референтов и редакторов СА, где помимо обязанностей на 4–5 страницах, давались указания в виде 35 параграфов, касающихся номенклатурных вопросов и стилистических особенностей реферирования и редактирования. Дополнения к инструкции 1910 г. включали: список 74 стандартных сокращений, классификатор и список журналов. В 1939 г. Edward J. Crane изменил формат инструкции до карманного, которая в дополнение к предыдущим сведениям содержала предметный указатель и инструкцию для реферирования патентов [9].

В дальнейшем, детальные рефераты были вытеснены индикативными и поисково-ориентированными рефератами. Поисково-ориентированный реферат также является кратким изложением статьи, но не содержит деталей эксперимента, а ограничивается описанием основного открытия, результатов и выводов. С 1970 г. в CAS в основном перешли на составление поисково-ориентированных рефератов. Краткое определение реферата, соответствующее текущему моменту, было сформулировано Американским национальным институтом стандартов: “Хорошо составленный реферат даст возможность читателям быстро и точно оценить содержание документа, определить насколько он отвечает их интересам и таким образом решить необходимо ли знакомиться с оригиналом” [10].

Е. J. Crane дал свое “определение” поисково-ориентированного реферата:

Краткость рекомендуется в речи, да,  
В юбках, я полагаю,  
В волосах, возможно,  
В дремоте на службе,  
А теперь, и в рефератах также [11].

В начале 1970-х гг. было принято решение о том: что индексировать и реферировать документ должен один и тот же сотрудник. В 1978 г. вышло 4-х томное издание “Руководство по анализу документов”. Но только первый из 4-х томов имел непосредственное отношение к составлению рефератов, остальной материал был посвящен составлению указателей. Изменения в форме реферирования, начатые в середине XX в., в настоящее время привели к тому, что в БД СА/CASplus публикуются авторефераты из англоязычных первоисточников. Основной задачей сотрудников CAS, обрабатывающих документы, стало индексирование первоисточника

“Рефераты – зачастую являются политематическими, т. е. реферат может представлять интерес для нескольких рубрик СА и может не соответствовать однозначно ни одной из рубрик”, писал E. J. Crane [7]. Предпочтение проставления кода той или иной рубрики всегда отдавалось редактором, регистрирующим химическое вещество или соединения, или химические реакции, если таковые присутствовали в реферируемой работе. Кроме того, если это было возможно, добавлялись перекрестные ссылки, что значительно расширяло концепцию однократного рубрицирования.

Указывая, что в СА должна быть “информация, которую хочет получить читатель”, и отмечая особенности как хорошего реферата, так и хорошего реферативного журнала, в 1958 г. E. J. Crane писал. “полнота охвата, оперативность, точность, ясность и удобство пользования являются основными особенностями хорошего реферативного журнала. Это необходимо, если реферативный журнал с уверенностью гарантирует, что существующая важная информация не будет пропущена” [7]. В 70-е г. прошлого века в связи с развитием отдельных направлений химии, физики, биологии, вычислительных наук, возникновением приоритетных направлений в науке и технологии изменился подход к формированию рубрик

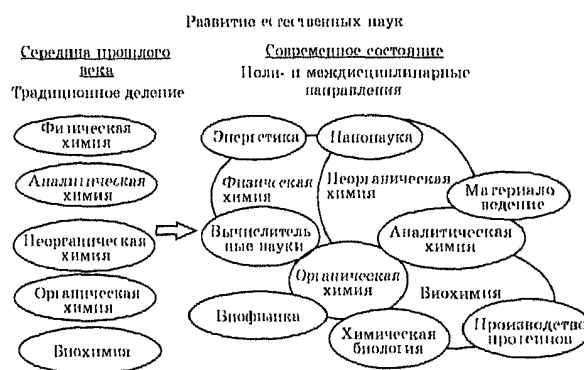


Рис. 2 Отражение тенденций развития естественных наук, начиная с середины прошлого века (Peter Stang <http://www.cas.org/about/cas100/annivhistory.html>)

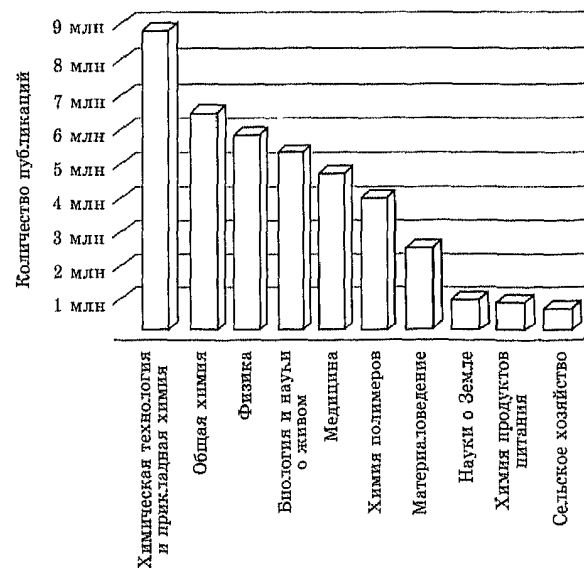


Рис. 3 Распределение потока публикаций по отраслям знания в БД Chemical Abstracts (2007 CAS Catalog, [www.cas.org](http://www.cas.org))

РЖ. Был сформулирован новый подход к организации реферативной службы, основу которого составила концепция отражения публикаций, "ориентированных на определенную проблему, отрасль знания" (Рис. 2) [7]. Большую роль в успешном решении задачи отражении приоритетных направлений сыграли предметно-систематический указатель General Subject Index и Index Guide, содержащий тезаурус. Анализ тематического содержания БД СА показал, что она в настоящее время является политематической, мультидисциплинарной научно-технической БД (Рис. 3) [12, 13].

**Указатели.** Как любой другой реферативный журнал, СА начал издавать традиционные виды указателей: предметный, авторский и патентный. Стали выпускаться пономерные, полугодовые и кумулятивные указатели (индексы) [14]. История развития кумулятивных указателей кратко может быть представлена в таблице.

Формульный указатель способен дать однозначный ответ на вопрос: в каком первоисточнике можно найти информацию об интересующем читателя соединении и описано ли это вещество в научной литературе? Появление указателя химических соединений связано с принятой в начале 1970-х гг. номенклатурой CAS. Эти два вида указателей, объединенные в электронном виде и дополненные регистрационными номерами CAS стали основой для создания БД Registry и информационного ресурса CAS Registry Handbook, издаваемого в печатном варианте и на микрофишах.

Информацию о циклических соединениях содержит печатный выпуск RING SYSTEMS HANDBOOK. Расшифровка тривиальных названий с помощью номенклатуры CAS, предметный тезаурус к General Subject и номенклатурные правила CAS помещены в Index Guide - путеводителе к указателям, издаваемом с 1968 г.

**Применение электронно-вычислительных машин.** Экспоненциальный рост научно-технической литературы в 50-х гг. XX в. потребовал выработки новых стратегических решений по технологии обработки стремительно возрастающего

массива документов. В 1955 г. создается новое подразделение R&D и журнал СА выходит из состава издательского подразделения Американского химического общества. Так в 1956 г. возникла организация Chemical Abstracts Service (CAS). С этого времени стали осваиваться новые технологии обработки данных с применением вычислительных машин. При содействии Национального фонда научных исследований CAS включилась в 5-летнюю программу по созданию оперативной информационной системы. В 1965 г. была создана автоматизированная система CAS Registry System. В 1967 г. - началось освоение компьютерного ввода рефератов и библиографии и диалога с машиной по адекватному представлению данных; с 1970 г. - переход на полностью автоматизированную подготовку РЖ, БД, Указателей. Результаты внедрения новых технологий и подходов к отражению документов впечатляющие: с 1907 по 1967 гг. количество записей в СА составляло 2,5 млн, с 1968 по 2007 гг. - 23,5 млн.

В это же время CAS осваивает новые фотонаборные технологии печати, микроформы и, наконец, магнитные ленты, позволяющие хранить введенную информацию и многократно использовать ее для пономерных, годовых и кумулятивных указателей. Тогда же были предприняты попытки с помощью ЭВМ создавать рефераты и индексировать первоисточник, но, как оказалось, компьютер не способен заменить человека в процессе интеллектуального труда.

С переводом РЖ СА в машиночитаемую форму было выпущено новое информационное издание, представляющее узкотематическую подборку по актуальным проблемам современной науки - СА SELECT. Начав с 6 наименований выпусков в 1976 г., к январю 1978 г. было предложено уже 110, а в 2007 г. около 240 выпусков. Тематика СА SELECT определяется потребностями специалистов. Так, СА SELECT по одной тематике может выпускаться в трех вариантах: только по статьям

Динамика создания указателей к РЖ СА

Коллективный индекс	Годы издания/указатели				
	основной предметный	названия химических веществ и соединений	авторский	формульный	патентный
1	1907 1916		1907 1916		
2	1917 1926		1917 1926		1907 1936
3	1927 1936		1927 1936	1920 1946	
1	1937 1946		1937 1946		1937 1946
5	1947 1956		1947 1956	1947 1956	1947 1956
6	1957 1961		1957 1961	1957 1961	1957 1961
7	1962 1966		1962 1966	1962 1966	1962 1966
8	1967 1971		1967-1971	1967 1971	1967 1971
9	1972 1976	1972 1976	1972 1976	1972 1976	1972 1976
10	1977 1981	1977 1981	1977 1981	1977 1981	1977 1981
11	1982 1986	1982 1986	1982 1986	1982 1986	1982 1986
12	1987 1991	1987 1991	1987 1991	1987 1991	1987 1991
13	1992 1996	1992 1996	1992 1996	1992 1996	1992 1996
11	1997 2001	1997 2001	1997 2001	1997 2001	1997 2001

из журналов, только по патентным документам или по всем видам документов. Кроме того, с течением времени тематика выпусков может меняться в зависимости от развития отдельных научных направлений и интереса специалистов. Например, ранее выпускавшийся CA SELECT "Fullerenes" в настоящее время прекратил свое существование в связи с изменившимися приоритетами работ в США в области фуллеренов (переход от исследовательских работ к промышленному производству). В настоящее время CA SELECT's выпускаются как в печатном виде, так и в электронном (в Интернете).

**Теледоступ.** Новейшая история развития информационных ресурсов CAS связана с организацией доступа в режиме реального времени. В начале 80-х г. прошлого века руководство CAS сумело оценить преимущества доступа в режиме онлайн, а также убедить в перспективности этого направления и других производителей научно-технической информации. В 1983 г. было подписано соглашение с FIZ Karlsruhe (Германия) о создании международной сети научно-технической информации – The Scientific and Technical Information Network (STN International). В 1984 г. CAS организовала онлайн-новый доступ со своего сервера к БД как собственного производства, так и других производителей научно-технической информации. В 1987 г. стал возможен доступ через сервер FIZ Karlsruhe, а в 1988 г. — через сервер The Japanese Information Center of Science and Technology (JICST, Япония). Интересно то, что STN International была создана и начала функционировать до появления Всемирной паутины — Интернета, предвосхищая возможность научных коммуникаций специалистов разных стран. В 2007 г. в STN International представлено около 220 БД, генераторами которых являются Elsevier B. V., DECIHEMA e. V., Thomson Scientific, Cambridge Scientific Abstracts, Gale Group, Institution of Electrical Engineers и т. д. Необходимо отметить одно из очень важных достоинств созданной сети — все БД представлены в едином стандарте. Для удобства пользователей БД объединены в кластеры или по тематическому признаку: CHEMISTRY, PHYSICS, ENGINEERING, PETROLEUM, FUEL, ENVIRONMENT и т. д. или по типу информации ALLBIB, FULL TEXT, STRUCTURES и т. д.

В настоящее время нет ни одной международной сети, в которой было бы представлено большее количество БД по научно-технической информации. Помимо информационных ресурсов CAS в STN International размещены сопоставимые по объему БД, содержащие порядка нескольких миллионов записей. Это БД научного цитирования (SCISEARCH), крупнейшие патентные БД INPADOC, WPIDS (Derwent World Patent Index), БД INSPEC, COMPENDEX, GEOREF, EMBASE, BEILSTEIN, GMELIN и многие другие, размещенные на одной платформе с единым интерфейсом, что позволяет проводить одновременный поиск в нескольких БД и исключать дубликатные документы. Специальное программное обеспечение STN AnaVist, созданное в 2005 г., предлагает дополнительные возможности для анализа результатов поиска, например, для получения статистических данных в виде таблиц и графиков. Используя преимущества STN AnaVist можно применять наукометрические методы анализа при обработке больших объемов информации.

Самыми активными пользователями ресурсов STN International являются патентные ведомства. Согласно данным CAS, количество химических веществ, описанных в патентных документах, в 1976 г. составляло 14%, а в 2006 г. выросло до 63% от общего количества соединений (проиндексированных из всех видов документов). Как правило, патентные ведомства стран, в которых активно идет процесс патентования, имеют контракты на неограниченный доступ к БД STN International. Сумма таких контрактов составляет сотни тысяч долларов США в год. Для пользователей с более скромными средствами предлагается оплата доступа с учетом фактически проведенных поисков, в соответствии с прайс-листом, опубликованным на сайтах CAS, FIZ Karlsruhe и JAICI (организации, которая стала партнером США и Германии в Японии с 2006 г.). Помимо высокой стоимости ресурсов STN International, другим недостатком, с точки зрения пользователей, является сложный интерфейс, требующий обучения информационных специалистов. Поэтому Robert J. Massie, возглавивший CAS в 1992 г., большое значение придает совершенствованию информационного сервиса. В 1995 г. был создан новый информационный ресурс CAS с дружественным интерфейсом SciFinder [15]. В 1999 г. для учебных заведений разработан вариант SciFinder Scholar, основное отличие которого, заключается в иной форме оплаты за подписку, при этом в SciFinder Scholar возможно неограниченное число поисков, что жизненно необходимо для образовательного процесса. Содержащие SciFinder/SciFinder Scholar составляют 6 БД, 5 из которых создаются в CAS: CAPlus, CAS Registry, CASREACT, CHEMLIST, CHEMCATS и 1 БД производства Национальной медицинской библиотеки США MEDLINE. Следует отметить, что SciFinder/SciFinder Scholar в отличие от STN International является эксклюзивным продуктом CAS, доступ к которому невозможен через FIZ Karlsruhe и JAICI.

Для развития поискового сервиса в 1996 г. была создана STN Easy – информационно-поисковая система с более простым поисковым языком, чем STN International, но содержащая меньшее количество БД (около 80). В 1999 г. стал возможен доступ к БД STN International в Интернете (STN on the Web). В 1997 г. появилось программное обеспечение CHEMPORT, позволяющее перейти от результатов поиска в STN International или в SciFinder/SciFinder Scholar, или в других электронных информационных ресурсах производства CAS к полным текстам документов, размещенных в Интернете на сайтах издателей. CHEMPORT установлен на компакт-дисках БД CA. Последний вариант (CA на CD) разработан специально для библиотек.

За прошедшее столетие произошли значительные изменения в мире науки и производства. Помимо США и Европы большой вклад в научно-технический прогресс вносится странами Азии. Японией, Китаем, Индией и Малайзией. Изменилась форма представления научно-технической информации, революционную роль сыграли в этом интернет-технологии, возникли новые поисковые машины (Google и другие), но при этом реферативные службы продолжают играть ведущую роль в