

## СОВРЕМЕННЫЕ СИСТЕМЫ ПОИСКА ХИМИЧЕСКОЙ ИНФОРМАЦИИ

**Сальникова Е.В., канд. хим. наук, доцент,  
Сальникова В. И., Дроконова А. В., Юдин А. А., Аманов П. Ч.  
Федеральное государственное бюджетное образовательное учреждение  
высшего образования «Оренбургский государственный университет»**

Любое исследование должно начинаться с изучения опыта предшественников, зафиксированного и накопленного в литературе. Статистика показывает [1], что половину книг фундаментальных библиотек ни разу не открывал читатель – и не потому, что они были не нужны, а потому, что читатели не знали об их существовании.

Зарождение химии как науки, отделение ее от физики и других ветвей естествознания, произошло на рубеже XVII – XVIII веков. И уже в то время поток информации, был настолько обширный, что возникла острая необходимость создания вспомогательных изданий, облегчающих исследователям ознакомление с тем, что сделано в науке ранее. В 1830 годах стали появляться первые реферативные журналы, а также фундаментальные справочники [2].

Более столетия достаточно хорошо функционировала система информационных изданий, созданная в химии, однако в конце XX века ситуация резко изменилась. В 1970 году прекратил своё существование старейший реферативный журнал «*Chemisches Zentralblatt*» [2] и во всем мире осталось только два общехимических реферативных издания – советский «Реферативный журнал Химия» и американский «*Chemical Abstracts*».

В настоящее время заметно увеличился рост научно-технической информации. Это связано с тенденцией в развитии различных областей науки, а также с возрастанием числа инженеров и ученых, создающих эту информацию. Химия приобрела огромное значение в жизни современного человека. Наука стала настолько колоссальной, что давно уже невозможно удержать в памяти содержание даже узкой её области.

Большое внимание в мире уделяется естественным наукам, и современная химия достигла высокого уровня развития: стремительное расширение применения химических процессов и продуктов в материальном производстве находит многоплановое отражение [3], с каждым годом всё сильнее растет объём проведенных исследований и сделанных открытий, тем самым затрудняется поиск нужных сведений.

Поток информации в современном мире настолько велик, что многие правила работы с литературой теперь уже потеряли свой смысл. Рост числа научных документов практически исключает возможность ознакомления с ними путем непосредственного чтения или хотя бы просмотра всех документов. Информационный поиск приходится практически проводить, используя поисковые образы документов, отражающие их содержание хотя бы в краткой форме. С ростом объёма информации, появилась потребность в изменении подхода

к её использованию, именно поэтому проблема поиска химической информации является актуальной. Профессионал-химик должен уметь решать эту проблему, причем решать ее с использованием самых различных, в том числе и современных методов [4].

Информационный поиск состоит в сравнении индексов документов, имеющих в поисковом массиве, с индексами запросов потребителей [2]: при совпадении обоих индексов документ считается релевантным – то есть отвечающим запросу. Необходимо отметить, что работа с химической информацией осложнена ее спецификой – кроме естественного языка в химии применяются специальные языки брутто- и структурных формул. С появлением глобальной сети Интернет произошла автоматизация процедуры поиска, были созданы информационно-поисковые системы, электронные базы данных.

Решением проблемы специфики химического языка являются системы правил однозначного описания состава и структуры молекул веществ. Одна из таких систем – спецификация упрощенного представления молекул в строке ввода – SMILES. Еще одним международным текстовым химическим идентификатором является код InChi, разработанный специалистами IUPAC и NIST. Из кода InChi вытекает новый хешированный формат – InChiKey, фиксированной длины, более удобный для индексации в информационно-поисковых системах. Необходимо отметить, что SMILES имеет преимущества перед InChi, например, лучшее восприятие формул человеком, более простую программную поддержку, связанную с обширной теоретической базой – теорией графов.

У большинства сайтов, которые находятся в свободном доступе, отсутствует гарантия надежности, и сообщаемые данные указаны не в полном объеме. Для того чтобы найденная информация была достоверной, необходимо обращаться к профессиональным поисковым источникам.

База данных обеспечивает хранение информации и предоставляет поименованную совокупность данных, организованных по определенным правилам, включающим общие принципы описания, хранения и манипулирования данных [5].

Первые базы данных научно-технической информации возникли как побочный продукт компьютеризации реферативных журналов [6]. Постепенно в результате появления мощных вычислительных систем, возможности практически каждого человека работать с электронной техникой и сетей передачи данных – роль таких баз данных значительно возросла. Стало возможным их широкое применение, в том числе индивидуальное использование в режиме удаленного доступа. Современные базы данных обеспечивают подробнейший быстрый поток необходимых сведений.

Рассмотрим некоторые базы данных по химии глобальной сети, действующие как в России, так и за рубежом.

STN International – Международная сеть научно-технической и деловой информации, создана в 1984 году на базе научных центров Германии, США и Японии, которые имеют многолетний опыт предоставления услуг в области информации. В контексте химической информации особыми достоинствами сети STN

могут являться: многообразие форм поиска химических соединений (по молекулярной структуре или ее фрагментам, формулам, названиям, номерам CAS), самый большой в мире репозиторий [6] полипептидных и нуклеотидных последовательностей, гибкие форматы (в том числе и бесплатные) вывода найденной информации.

ИПС SciFinder – одна из наиболее востребованных среди мировых научных информационных ресурсов. Это единая платформа которая представляет собой шесть различных баз данных: библиографические (Chemical Abstracts Plus, MEDLINE), структурно-химические (CAS Registry, CASREACT) и т.д. В 1999 году появился SciFinder Scholar – тот же самый по наполнению ресурс, предназначенный для университетов. БД охватывают более 66 млн. химических веществ, 1,1 млн. экспериментальных спектров, 41,7 млн. одно- и многостадийных реакций, 35 млн. ссылок из более чем 60 мировых патентных ведомств.

ИПС «Термические константы веществ» – электронная версия справочника Термические константы веществ: Вып. 1-10 / Отв. ред. В. П. Глушко.- М.: ВИНТИ, 1965-1982.; Thermal Constants of Substances: Vol. 1-8 / Ed. V.S. Yungman. – NY: Wiley, 1999. Подготовка этого издания проводилась группой из более чем восьмидесяти экспертов в области химической термодинамики. Работа продолжалась около двадцати лет и привела к публикации в 1965-1982 годах справочного здания, включающего 10 томов, содержащих сведения о более чем 25000 веществ, образованных практически всеми химическими элементами. Библиография включает больше 51500 ссылок [7].

LIFBASE – программа для расчета спектроскопических параметров (частоты и коэффициенты поглощения и эмиссии для колебательно-вращательных переходов, вероятности переходов в зависимости от вида колебательно-вращательных волновых функций и момента электронного перехода и т.д.) двухатомных молекул. Плюсом данного программного обеспечения является то, что оно распространяется исключительно на свободной основе.

Базы данных Национального института стандартов и технологий США (NIST): ASD – БД атомных спектров, Химическая Кинетика, Computational Chemistry Comparison and Benchmark Database – коллекция теоретических и экспериментальных термодинамических данных. Особо хотелось бы отметить БД Chemistry WebBook, содержащую данные о термохимии более 10 тысяч реакций, ИК спектров более 9 тыс. веществ, более 13 тыс. масс-спектров. Помимо аналогового изображения спектров практически на каждое вещество имеется оцифрованная версия, которую можно скачать. На сайте также имеется онлайн-версия редактора ИК-спектров.

Базы данных Международного союза теоретической и прикладной химии (IUPAC) включают кинетические и фотохимические данные, таблицы по гетерогенным реакциям, термодинамические таблицы.

БД FACT – Facility for the Analysis of Chemical Thermodynamics – объединенная термодинамическая база данных. Применяется во многих иных областях химической термодинамики: пирометаллургии, гидрометаллургии, геохимии, геологии, электрохимии.

Объединенная спектральная система баз данных органических веществ (SDDBS, Япония): содержит огромное количество ЯМР, масс-, ЭПР, ИК-, СКР спектров. Доступ к базе данных бесплатен, однако накладывается ограничение на количество скачанных спектров: не более 50 за день.

Огромное количество полезной информации в области химии, предоставляет сайт Химического факультета Московского государственного университета. Множество углубленной литературы, наглядных лекций, представленных в виде презентаций, научных статей и справочников находится в свободном доступе. Также на сайте МГУ имеется информация о нескольких базах данных, таких как:

1. база данных Всероссийского института научной и технической информации;
2. реферативный журнал «Chemical Abstracts»;
3. INSPEC (Physics Abstracts);
4. Inorganic Crystal Structure Database (структуры неорганических кристаллов), и т.д.

Рассмотренные базы данных и информационно-поисковые системы не являются исчерпывающими, существует еще огромное количество источников полезной информации в глобальной вычислительной сети. Однако рассмотренный перечень будет весьма полезен студентам-химикам на протяжении всего периода обучения и дальнейшей профессиональной деятельности, так как большинство БД содержат не только справочную информацию, но и имеют ссылки на иные базы данных (в том числе и междисциплинарные), литературные источники, информацию о современных исследованиях, наукометрические данные, фундаментальные справочники (Гмелина, Меллора, Фрезениуса, Бейльштейна и др.). Оцифрованная версия вышеуказанных упрощает процесс поиска в несколько раз.

#### Список литературы

1. Рагойша, А. А. Поиск химической информации в Интернете: научные публикации: учеб. пособие для студентов хим. фак. спец / А. А. Рагойша. – Минск: БГУ, 2007. – 71 с.
2. Потапов, В. М. Химическая информация: Что, где и как искать химику в литературе / В. М. Потапов, Кочеткова, Э. К. – Москва: Химия, 1979. – 304 с.
3. Лукина, Ю. А. К вопросу о взаимодействии теории и практики в развитии химии / Ю. А. Лукина, И. Д. Третьяков // Вестник ИрГТУ. 2011. №4 (51). Режим доступа: <http://cyberleninka.ru/article/n/k-voprosu-o-vzaimodeystvii-teorii-i-praktiki-v-razviii-himii> (дата обращения: 23.12.2017).
4. Высоцкий, В. И. Поиск химической информации в Интернет: Метод. указания для студентов Ин-та химии и приклад. экологии / В. И. Высоцкий. – Владивосток: Изд-во Дальневост. ун-та, 2000. - 19 с.
5. Голенищев, Э. П. Информационное обеспечение систем управления / Э. П. Голенищев, И. В. Клименко. – Ростов н/Д: Феникс, 2003. – 352 с.

6. Зибарева, И. В. Поиск химической информации в базах данных сети STN International / И. В. Зибарева. – Новосибирск: НИИ НГУ, 2015. – 81 с.

7. База данных ТКВ. [Электронный ресурс]. Режим доступа: <http://www.chem.msu.ru/cgi-bin/tkv.pl> (дата обращения: 23.12.2017).