

БИБЛИОТЕЧНЫЕ КАТАЛОГИ И ИНФОРМАЦИОННО-ПОИСКОВЫЕ СИСТЕМЫ

УДК 004.65:[002:54] ВИНИТИ
<https://doi.org/10.33186/1027-3689-2022-10-31-51>

База структурных данных по химии ВИНИТИ РАН. Вопросы формирования, эксплуатации и создания информационных продуктов

Н. И. Чуракова¹, Ю. Е. Бессонов², Б. С. Фельдман³,
Н. В. Червинская⁴

*^{1, 2, 3, 4}Всероссийский институт научной и технической информации РАН,
Москва, Российская Федерация*

¹*nichurak@rambler.ru*

²*bessonov-ye@rambler.ru*

³*bsf@inbox.ru*

⁴*dir@viniti.ru*

Аннотация. В статье описана База структурных данных по химии ВИНИТИ РАН (База СД). Представлены краткая история создания Базы СД, её содержание, форматы данных и программное обеспечение. Дано описание современных технологий пополнения Базы СД, в основе которых лежит работа исключительно с электронными носителями информации. Представлены методы и программы адаптации данных ретрофонда Базы СД к современному программному обеспечению. Рассмотрены информационные продукты, получаемые на основе Базы СД: система структурного поиска в интерактивном режиме, автономная система структурного поиска, автономная система поиска химических реакций, электронный справочник химических соединений, электронный справочник именных реакций. Достоинством автономных систем поиска химических структур и реакций является то, что они могут работать на обычных персональных компьютерах. Представленные информационные продукты, получаемые на основе данных Базы СД, могут использоваться при проведении фундаментальных и прикладных исследований по химии, в химической промышленности, в учебном процессе, в качестве информационного обеспечения баз данных, научных библиотек, издательств. Массив данных Базы СД может использоваться в различных научных исследованиях, требующих большого объёма фактической информации.

Ключевые слова: химическая информация, база химических структур и реакций, поиск и отображение информации в базах данных по химии

Для цитирования: Чуракова Н. И., Бессонов Ю. Е., Фельдман Б. С., Червинская Н. В. База структурных данных по химии ВИНТИ РАН. Вопросы формирования, эксплуатации и создания информационных продуктов / Н. И. Чуракова, Ю. Е. Бессонов, Б. С. Фельдман, Н. В. Червинская // Научные и технические библиотеки. 2022. № 10. С. 31–51. <https://doi.org/10.33186/1027-3689-2022-10-31-51>

LIBRARY CATALOGS AND INFORMATION RETRIEVAL SYSTEMS

UDC 004.65:[002:54] ВИНТИ
<https://doi.org/10.33186/1027-3689-2022-10-31-51>

The VINITI RAS structural database in chemistry. Acquisition, operation and design of information products

Natalia I. Churakova¹, Yury E. Bessonov², Boris S. Feldman³
and Nadezhda V. Chervinskaya⁴

^{1, 2, 3, 4}RAS All-Russian Institute for Scientific and Technical Information,
Moscow, Russian Federation

¹nichurak@rambler.ru

²bessonov-ye@rambler.ru

³bsf@inbox.ru

⁴dir@viniti.ru

Abstract. The authors describe the Structural Database (SD) in chemistry supported by the All-Russian Institute of Scientific and Technical Information of the Russian Academy of Sciences (VINITI RAS). The SD history, content, data formats and soft-

ware are characterized in brief. The database development technologies based exclusively on digital media, are described. The techniques and software for adapting SD retrocollections to advanced software are discussed. The SD products are characterized, i. e. online structured search system, autonomous structured search system, autonomous system for chemical reaction search, e-dictionary of named reactions. The advantage of autonomous systems for searching chemical structures and reactions is that they can be installed on PCs. The SD-based information products can be used in fundamental and applied chemical studies, in chemical industry, education, and in information support of databases, sci-tech libraries and publishers'. The SD array may be applied in the studies based on the vast amount of factual information.

Keywords: chemical information, database of chemical structures and reactions, information search and display in chemical databases

Cite: Churakova N. I., Bessonov Yu. E., Feldman B. S., Chervinskaya N. V. The VINITI RAS structural database in chemistry. Acquisition, operation and design of information products / N. I. Churakova, Yu. E. Bessonov, B. S. Feldman, N. V. Chervinskaya // Scientific and technical libraries. 2022. No. 10. P. 31–51. <https://doi.org/10.33186/1027-3689-2022-10-31-51>

Введение

В современных условиях формирование баз структурных данных по химии является приоритетным направлением в информационном обеспечении специалистов-химиков, а также представителей смежных с химией областей науки. Подобные базы данных дают возможность оперативно получать релевантную информацию, необходимую для научных исследований.

В ВИНИТИ РАН формируется крупнейшее в России хранилище химической информации – База структурных данных (СД) по химии на основе аналитико-синтетической обработки документов, представленных в отечественных и зарубежных научных публикациях.

Массивы накопленной в Базе СД информации используются при создании различных информационных продуктов, к которым относятся системы структурного поиска химической информации и электронные справочники.

База СД (краткая история)

Работа по формированию Базы данных началась в ВИНТИ РАН в 1975 г. в рамках сотрудничества с фирмой SPRESI (ГДР) на основе её программного обеспечения (Speicherung und Recherche Struktur chemischer Information), позволяющего находить и накапливать структурную химическую информацию.

В 1996 г. для формирования Базы СД была создана первая отечественная графическая программная оболочка CBASE16 (16-разрядная версия), предназначенная для ввода и обработки структурных, фактографических и библиографических данных по химии [1]. В 2010 г. в технологический режим формирования Базы СД было введено программное обеспечение CBASE32, которое является развитием программной оболочки CBASE16 и представляет собой её 32-разрядную версию (используется до настоящего времени) [2].

В становлении и развитии Базы СД можно выделить три периода, связанных с использованием различного программного обеспечения и, соответственно, различным форматом представления данных (табл. 1). Различные форматы представления данных, накопленных в прежние годы, создают значительные сложности при их использовании в условиях современного программного обеспечения.

Данные Базы СД с 1996 г. хранятся на сервере ВИНТИ РАН. Остальные данные находятся на съёмных носителях информации.

Таблица 1

Содержание Базы СД и формат представления данных в разные временные периоды

Период	Число структур	Число реакций	Формат
1975–1995 гг.	3 636 696	Неизвестно	ПБ-код, Spresi
	2 117 786	–	Т-граф
1996–2009 гг.	2 500 758	1 014 022	CBASE16, SDF/RDF

Период	Число структур	Число реакций	Формат
2010–2020 гг.	1 175 510	698 112	CBASE32, SDF/RDF
<i>Всего</i>	9 430 750	1 712 134	
<i>Всего доступно для обработки</i>	5 794 054	1 712 134	

Примечание 1

ПБ-код (произвольно-блочный код) – способ кодирования структуры органических и неорганических соединений на основе описания фрагментов химической структуры и взаимосвязей между фрагментами.

Формат Spresi – система кодировки данных о химических структурах и реакциях, которая использовалась в 1975–1995 гг. в процессе сотрудничества с фирмой SPRESI (ГДР).

T-граф – запись структуры химического соединения в виде текстовой строки.

Форматы CBASE16, CBASE32 – бинарные форматы внутреннего представления информации о химических структурах в системах CBASE16 и CBASE32.

Форматы SDF/RDF – общепринятые форматы для обмена информацией о химических структурах (SDF) и реакциях (RDF) между базами данных.

Примечание 2

Данные в форматах ПБ-код и Spresi (3 636 696 соединений) в настоящее время не задействованы в информационных процессах ввиду отсутствия необходимого программного обеспечения. Для остальных форматов записей химических структур и реакций программное обеспечение имеется. Таким образом, на данный момент доступным для использования является массив данных, содержащий информацию о 5 794 054 химических соединениях и 1 712 134 химических реакциях.

Формирование входного потока документов для пополнения Базы СД

Одним из основных элементов технологического режима создания Базы СД является формирование входного потока документов, подлежащих обработке. База СД ориентирована на обработку научно-технической литературы, связанной с важнейшим разделом химии – органической химией и особенно с органическим синтезом. Современная цивилизация требует постоянного создания новых жизненно необходимых химических соединений с заранее заданными свойствами.

База СД содержит структурную информацию о химических соединениях и реакциях, которая сопровождается разнообразной описательной текстовой информацией, представленной в виде так называемых предметных характеристик. База СД позволяет химику-исследователю быстро получать необходимые сведения о строении и свойствах химических соединений, методах их получения и применении, а также об их активности (биологической, фармакологической, токсикологической и др.).

На протяжении многих лет входной поток документов, предназначенных для пополнения Базы СД, состоял преимущественно из зарубежных ядерных журналов по химии. В настоящее время для поддержки российских исследований по органической химии и популяризации достижений отечественной науки, учитывая санкционные ограничения на доступ к зарубежным журналам, принято решение при пополнении Базы СД в первую очередь использовать российскую научную литературу по химии.

Программно-технологический комплекс CBASE32.

Новые технологии.

Современное программное обеспечение

За годы эксплуатации Базы СД сформировался программно-технологический комплекс CBASE32 с развитым математическим, лингвистическим и информационным обеспечением. В настоящее время он содержит четыре основных компонента (табл. 2).

**Основные компоненты программного комплекса CBASE32
и их роль при формировании Базы СД**

№ п/п	Компонент программного комплекса CBASE32	Перечень вводимой информации
1	Программное обеспечение для ввода информации о соединениях	Молекулярная формула, систематическое и/или тривиальное название, химическая структура, ключ InChIKey*; предметные характеристики: физико-химические свойства, получение, реакционная способность, биологическая активность, применение
2	Программное обеспечение для ввода информации о реакциях	Участники реакции (реагенты, растворители, катализаторы, прочие участники), условия реакции (температура, давление, время), выход, предметные характеристики, описание реакций по стадиям
3	Электронный справочник химических соединений (Глоссарий)	Стандартизация ввода химических структур и их названий
4	Электронный справочник именных реакций	Автоматизированный ввод имени реакции (при наличии его в статье)

* InChIKey – международный химический идентификатор IUPAC (International Union of Pure and Applied Chemistry), использующийся для поиска информации о химических соединениях в базах СД по химии.

Основные функции программного комплекса CBASE32:

обеспечение ввода и научного редактирования текстовой и графической информации о химических соединениях и реакциях;

отображение содержимого Базы СД с помощью многооконного интерфейса;

импорт и экспорт данных с использованием обменных форматов.

В 2017 г. начался переход к безбумажной технологии формирования Базы СД, то есть к технологии, в которой все обрабатываемые данные являются электронными на всех стадиях процесса обработки информации – от получения первичных данных до вывода (генерации) результирующих данных [3].

Для автоматизации обработки электронных статей было создано соответствующее программное обеспечение, включающее ряд программных модулей, функции которых заключаются в формировании заданий на загрузку электронных статей, просмотре всех статей, распределённых в соответствующий номер подготовки базы СД в формате PDF, распределении заданий сотрудникам для обработки, активизации, просмотра и редактирования, для архивации результатов обработки.

Внедрение безбумажной технологии привело не только к экономии материальных средств и временного ресурса специалистов-химиков, но и к повышению качества вводимой в Базу СД информации.

Технологии пополнения Базы СД:

Отбор статей из журналов, предназначенных для извлечения информации о химических соединениях и реакциях органического синтеза.

Формирование интегрированного электронного ресурса статей для Базы СД в формате PDF на базе статей, отобранных в п. 1.

Проведение компьютерной разметки документов. Разметка статьи в электронном виде (выделение существенных ключевых соединений и реакций) является необходимым этапом работы специалиста-химика при обработке документов. Для выполнения разметки использовалась программа Adobe Reader 11, предназначенная для работы с текстовыми документами, включающими также различные графические компоненты (рисунки, диаграммы и пр.), в формате PDF.

Создание библиотек (единиц постатейного хранения структурных данных) с использованием профессиональной программной среды, которая позволяет осуществлять ввод и научное редактирование структурных и текстовых данных с параллельным просмотром и разметкой электронной статьи. По завершении обработки библиотека наполняется данными о химических структурах и химических реакциях, а электронная статья получает разметку в виде набора комментариев о ключевых параметрах статьи.

Автоматическое соединение библиографических данных статьи и учётных данных ВИНТИ с библиотеками структурных данных и текстами статей в формате PDF.

Автоматическая проверка корректности структурных данных для химических соединений средствами CBASE32, не допускающими ввод некорректных структурных данных.

Проверка корректности следующих введённых данных: предметных характеристик химических соединений; предметных характеристик химических реакций; систематических (номенклатурных) названий химических соединений.

В настоящее время технологический процесс пополнения Базы СД включает четыре основных этапа:

1. Формирование интегрированного электронного ресурса документов.
2. Извлечение структурной химической информации из документов и её ввод в Базу СД.
3. Научное редактирование информации, введённой в Базу СД.
4. Автоматическая проверка корректности введённой в Базу СД информации.

Система структурного поиска в интерактивном режиме

В 2013 г. была разработана программная система, реализующая рациональный поиск структурной химической информации с учётом стереохимических особенностей. Для кодировки химических структур применена рекомендованная IUPAC технология InChIKey [4].

Предложен новый алгоритм сжатия InChIKey, в результате применения которого база данных из 100 млн структур требует всего 1,5 Гб ОЗУ. Для InChIKey определена метрика, на основе которой при поиске по точной структуре предложено использовать алгоритм бисекций [5]. В 2014–2015 гг. система была усовершенствована добавлением функции поиска структур по фрагментам [6].

Соответствующее веб-приложение [7] апробировано в ВИНТИ РАН. Объём тестовой интерактивной базы составляет приблизительно 200 тыс. структур.

В 2016 г. интерфейс системы был модернизирован: улучшен дизайн окна отображения результатов, добавлены новые информационные поля. Система функционирует в постоянном режиме на отдельном компьютере (сервере), подключённом к сети Интернет.

Система структурного поиска в автономном режиме

В 2015 г. была разработана первая версия программы ChemBD.exe [8, 9], позволяющая автономно находить информацию о различных характеристиках химических соединений. ChemBD.exe выполняет поиск в локальной базе данных (пользовательской БД), получаемой из Базы СД с помощью программы **Конвертор АСП**. Особенностью пользовательской БД является иерархическая организация её структуры, позволяющая эффективно выполнять поиск.

Пользовательская БД (ПБД) может быть тематической – ориентированной на конкретного заказчика. Программа ChemBD.exe вместе с ПБД называется автономной системой поиска (АСП). Схема формирования ПБД приведена на рис. 1.

На этой схеме показано, что **Конвертор АСП** читает файлы CSV с библиографией и добавляет библиографические данные к структурным данным Базы СД в файлах RDB, которые затем преобразуются в ПБД.

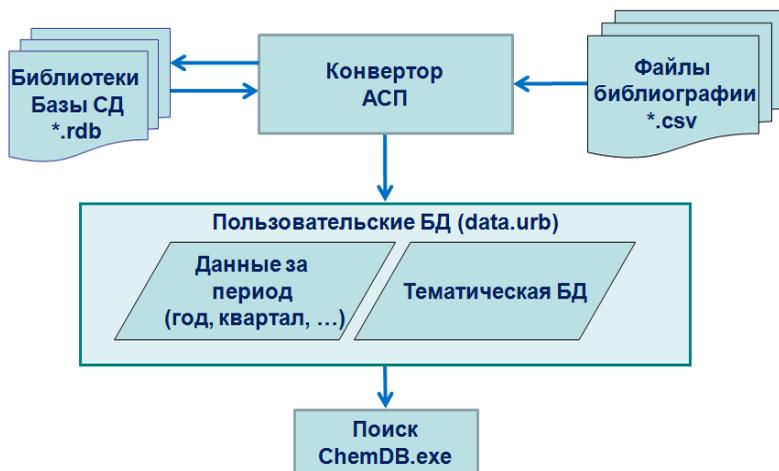


Рис. 1. Схема формирования пользовательских баз данных

Первоначальная версия программы **Конвертор АСП** позволяла представлять химическую реакцию лишь в виде её графического изображения. Однако данные, содержащиеся в Базе СД, дают возможность отражать в пользовательских БД расширенную информацию о химических реакциях, которая максимально полно описывает экспериментальные детали данной реакции и отражает реальную последовательность проведения отдельных стадий реакции, приводящих к получению целевого продукта.

В настоящее время разработана новая версия АСП, в которой существенно изменён интерфейс в части отображения подробной информации о химических реакциях (рис. 2, 3). С использованием актуальной версии АСП были созданы ПБД на основе массивов данных Базы СД, накопленных в периоды с 2010 по 2020 г.

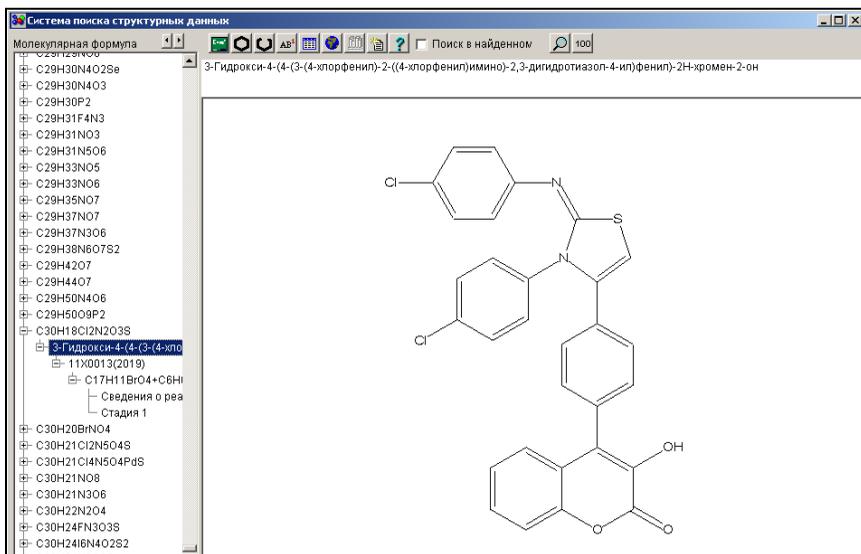


Рис. 2. Отображение данных о химической структуре

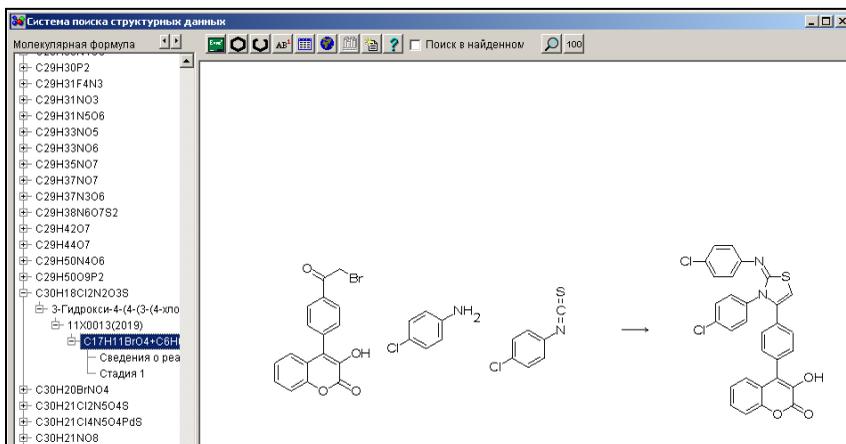


Рис. 3. Отображение данных о химической реакции

В совокупности эти базы данных содержат информацию о 1 175 510 структурах и 698 112 реакциях из 55 693 документов Базы СД. Данные по каждой ПБД приведены в табл. 3.

Таблица 3

Содержание ПБД за 2010–2020 гг.

Год	Число документов	Число структур	Число реакций
2010	7 797	145 084	84 775
2011	9 050	154 776	87 324
2012	7 801	149 877	92 102
2013	6 955	142 662	95 769
2014	6 242	133 203	89 527
2015	5 584	120 248	57 083
2016	3 950	107 225	53 770
2017	3 619	101 604	61 820
2018	3 695	96 134	61 086
2019	746	20 550	12 378
2020	254	4 147	2 478
<i>Всего</i>	55 693	1 175 510	698 112

С использованием актуальной версии АСП на основе аналитико-синтетической обработки статей, опубликованных в «Журнале органической химии» («ЖОрХ») за 2019 и 2020 гг., были сформированы годовые ПБД химических соединений.

В табл. 4 приведены характеристики ПБД, сформированных для «ЖОрХ» в 2019 и 2020 гг.

Таблица 4

Характеристики ПБД для «ЖОрХ» в 2019 и 2020 гг.

Год публикации, номер тома, номер выпуска	Количество обработанных статей	Количество химических соединений	Количество химических реакций
2019, т. 55, № 01 – 12	256	3 720	2 243
2020, т. 56, № 01 – 12	247	3 888	2 456

Данный информационный продукт является первым примером возможного сотрудничества специалистов в области обработки структурной химической информации и отечественными издательствами профильных журналов.

Следует отметить, что в ВИНТИ РАН отработана технология формирования электронных пономерных/годовых ПБД химических соединений, содержащихся в статьях, опубликованных в русскоязычных журналах химического профиля. В дальнейшем их можно размещать в согласованных форматах на сайтах соответствующих журналов или научно-технических библиотек, что может расширить круг авторов и читателей этих журналов.

Система поиска химических реакций

Чтобы получить данные о химических реакциях в программе АСП, сначала необходимо выполнить запрос по химическому соединению – участнику реакции. В 2021 г. была разработана система поиска химической информации, которая в отличие от АСП позволяет эффективно выполнять запросы, касающиеся непосредственно химических реакций [10]. Быстрый поиск обеспечивается иерархической организацией данных о реакциях (рис. 4).

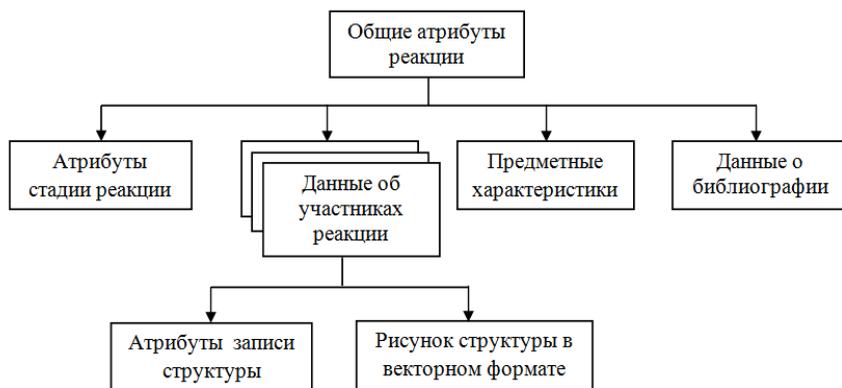


Рис. 4. Иерархическая организация данных о химической реакции

Особенность разработанной системы поиска реакций – её функционирование в автономном режиме на локальных компьютерах.

При этом данные о химических реакциях хранятся в локальных БД ограниченного объёма (в пользовательских базах данных химических реакций – ПБДХР), получаемых из Базы СД. Объём локальных БД соответствует календарному периоду формирования массива Базы СД: месяц, квартал, полугодие, год. В настоящее время созданы ПБДХР, содержащие данные о химических структурах и реакциях в статьях, опубликованных в «ЖОрХ» в 2019 г. (2 243 реакций) и в 2020 г. (2 456 реакций).

Электронные справочники

Электронный справочник химических соединений (Глоссарий) – структурированная и пополняемая база химических соединений [11]. Он позволяет стандартизировать представление структурных и фактографических данных о химических соединениях и упрощать процедуру ввода информации о соединениях и реакциях в Базу СД, а также осуществлять поиск соединения по молекулярной формуле, структуре и названию. Если соединение найдено в Глоссарии, то непосредственно из него оно вводится в Базу СД со всеми атрибутами и при этом получает номер, под которым оно зарегистрировано в Глоссарии. Пополнение и редактирование Глоссария осуществляются средствами CBASE32. В настоящее время Глоссарий содержит более 3 тыс. соединений.

Электронный справочник именных реакций – структурированная и пополняемая база именных реакций, содержащая более 300 именных реакций в русскоязычном и англоязычном вариантах, в которой возможен поиск по записям названий реакций на русском и английском языках [12, 13]. Справочник был сформирован в результате обработки данных из интернета и литературных источников.

Работа с ретрофондом

Ретрофонд Базы СД представляет собой данные, накопленные с 1975 по 1995 г. (2 117 786 структур) и с 1996 по 2009 г. (2 500 758 структур, 1 014 022 реакции). Химическая информация очень быстро накапливается и не устаревает. Поэтому данные из ретрофонда могут представлять интерес для специалистов по химии в наше время. Массив данных ретрофонда также может использоваться в различных научных исследованиях, требующих большого объёма фактической

информации. Примером может служить задача по определению взаимосвязи между структурой и свойствами химических соединений, решаемая с помощью нейросетей, в которых в качестве обучающей выборки используется база СД по химии.

Таким образом, задача адаптации данных ретрофонда к современным программным средствам обработки химической информации является актуальной.

Данные 1996–2009 г. хранятся в Базе СД в формате CBASE16. На схеме (рис. 5) показаны два пути их преобразования для работы с современным программным обеспечением. С одной стороны, данные преобразуются в формат CBASE32 с помощью специальной программы **Конвертор CBase16**. С другой стороны, они выгружаются системой CBase16 в файлы старых обменных форматов (SDF и RDF), которые затем преобразуются в современные форматы SDF и RDF с помощью разработанной программы **Конвертор SDF/RDF**. На схеме показано, как адаптированные данные могут предоставляться пользователям либо непосредственно в виде файлов SDF и RDF, либо в виде пользовательских баз данных АСП и Системы поиска данных о химических реакциях.

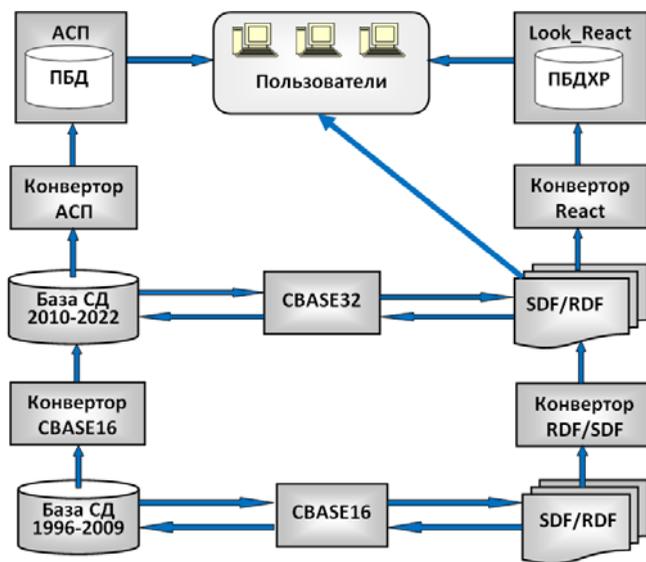


Рис. 5. Схема адаптации ретроданных к современному ПО

Данные 1975–1995 гг. хранятся в виде Т-графов, в которых закодированы сведения об атомах и химических связях. Существенным недостатком этих данных является отсутствие информации об изображениях структурных формул. Это объясняется тем, что ввод информации о химических структурах и реакциях стал выполняться с помощью графического редактора только в 1996 г. Автоматическое построение изображений химических структур (визуализация) по данным о химических связях между атомами – это сложная научная проблема.

В 2015 г. в ВИНТИ РАН были разработаны алгоритм и программа автоматической визуализации структур химических соединений [14]. Применение указанной программы к ретро Данным 1975–1995 гг. показало, что доля полученных качественных изображений структур составляет 99%. Результаты разработки опубликованы в статье [Там же].

Представленные выше информационные продукты могут использоваться:

- при проведении фундаментальных и прикладных исследований по химии;

- в химической промышленности;

- в учебном процессе;

- в качестве информационного обеспечения баз данных, научных библиотек, издательств.

Заключение

Представленная в статье База СД является единственной в России крупнейшей базой структурных данных по химии, которая содержит уникальный для нашей страны массив сведений по органическому синтезу, доступный для современного анализа и использования. Её значение сегодня выросло в связи с современной политической обстановкой. Российская научная общественность начинает осознавать, что необходимо создавать и поддерживать свои национальные ресурсы (журналы, различные базы данных и др.), которые содержали бы данные о достижениях как отечественной, так и зарубежной науки и которые не были бы подвержены никаким зарубежным санкциям.

Назрела необходимость создания российской национальной базы данных по химии. Требуется широчайшая кооперация российских научных организаций химического профиля по скорейшему наполне-

нию этой базы отечественными и зарубежными научными разработками, её популяризация для всех российских субъектов, занимающихся органическим синтезом. Необходимо создание отечественных технологических процессов, уменьшающих и нивелирующих зависимость России от импортных веществ и технологий в химическом синтезе.

Список источников

1. **Воронежева Н. И., Чуракова Н. И., Нечаева К. С. [и др.]** Индексирование и ввод химических реакций с помощью программы графической обработки данных CBASE. Временная инструкция 21–97. Москва, 1997. 89 с.
2. **Королёва Л. М., Фёдоровская М. А., Чуракова Н. И. [и др.]** Индексирование и ввод сведений о химических соединениях при подготовке базы структурных данных по химии с использованием программного комплекса CBASE32. Инструкция ВИНИТИ РАН 81-2010. Москва : ВИНИТИ РАН, 2010. 103 с.
3. **Фельдман Б. С., Бессонов Ю. Е., Кирьянова Н. С. [и др.]** Разработка и внедрение нового технологического процесса формирования Базы структурных данных по химии ВИНИТИ РАН с использованием первоисточников НТЛ в электронном виде. Тезисы доклада / Международная конференция «Информация в современном мире» (Москва, 25–26 октября 2017 г.). С. 338–344.
4. **Официальный сайт** группы InChI Trust. URL: <http://www.inchi-trust.org/downloads/> (дата обращения: 09.09.2022).
5. **Нефедов О. М., Трепалин С. В., Королева Л. М., Бессонов Ю. Е.** Быстрый поиск точных химических структур в больших базах данных с использованием InChI Key кодировки структур // Научно-техническая информация. Сер. 2. 2013. № 12. С. 27–33.
6. **Нефедов О. М., Трепалин С. В., Королёва Л. М. [и др.]** База структурных данных по химии ВИНИТИ РАН: проблемы поиска по фрагменту структуры // Научно-техническая информация. Сер. 2. 2014. № 12. С. 19–29.
7. **Система** поиска химических структур в интерактивном режиме. URL: <http://chem.viniti.ru/> (дата обращения: 9.09.2022).
8. **Trepalin S. V., Bessonov Yu. E., Fel'dman B. S. [et al.]** The Structural Chemical Database of the All-Russian Institute for Scientific and Technical Information, Russian Academy of Sciences. An Autonomous System for Structural Searches // Automatic Documentation and Mathematical Linguistics. 2018. Vol. 52. № 6. P. 297–305.
9. **Трепалин С. В.** Свидетельство о государственной регистрации программы для ЭВМ № 2017613588 ChemDB. Правообладатель: Федеральное государственное бюджетное учреждение науки Всероссийский институт научной и технической информации Российской академии наук (ВИНИТИ РАН) (RU). Заявка № 201760837. Дата поступления

01 февраля 2017 г. Дата государственной регистрации в Реестре программ для ЭВМ
22 марта 2017 г.

10. **Бессонов Ю. Е., Фельдман Б. С., Чуракова Н. И. [и др.]** Поиск и отображение информации о химических реакциях в базе структурных данных по химии ВИНТИ РАН // Научно-техническая информация. Сер. 2. 2022. № 3. С. 10–22.
11. **Voronezhcheva N. I., Trepalin S. V., Churakova N. I. et al.** Glossary as an Element of Data Input Standardization in the Cbase32 Program Complex // Automatic Documentation and Mathematical Linguistics. 2007. Vol. 41. № 3. P. 124–129.
12. **Вацууро К. В., Мищенко Г. Л.** Именные реакции в органической химии. Москва : Химия, 1976. 528 с.
13. **Ли Дж. Дж.** Именные реакции. Механизмы органических реакций. Москва : БИНОМ. Лаборатория знаний, 2006. 456 с.
14. **Nefedov O. M., Koroleva L. M., Trepalin S. V., Bessonov Yu. E. et. al.** The Development of an Integrated System for Structural Chemical Information. Automatic Documentation and Mathematical Linguistics. 2015. Vol. 49. № 6. P. 213–220.

References

1. **Voronezhcheva N. I., Churakova N. I., Nechaeva K. S. [и др.]** Индексирование и ввод химических реакций с помощью программы графического обработки данных CBASE. Временная инструкция 21–97. Москва, 1997. 89 с.
2. **Korolyova L. M., Fyodorovskaia M. A., Churakova N. I. [и др.]** Индексирование и ввод сведений о химических соединениях при подготовке базы структурных данных по химии с исползованием программного комплекса CBASE32. Инструкция ВИНТИ РАН 81-2010. Москва : ВИНТИ РАН, 2010. 103 с.
3. **Fel'dman B. S., Bessonov Iu. E., Kir'ianova N. S. [и др.]** Разработка и внедрение нового технологического процесса формирования Базы структурных данных по химии ВИНТИ РАН с исползованием первоисточников NTL в электронном виде. Тезисы доклада / Международная конференция «Информация в современном мире» (Москва, 25–26 октября 2017 г.). С. 338–344.
4. **Ofitcial'ny'i' sai't gruppy`** InChI Trust. URL: <http://www.inchi-trust.org/downloads/> (data obrashcheniia: 09.09.2022).
5. **Nefedov O. M., Trepalin S. V., Koroleva L. M., Bessonov Iu. E.** Быстрые поиски точных химических структур в больших базах данных с исползованием InChI Key кодировки структур // Научно-техническая информация. Сер. 2. 2013. № 12. С. 27–33.
6. **Nefedov O. M., Trepalin S. V., Korolyova L. M. [и др.]** База структурных данных по химии ВИНТИ РАН: проблемы поиска по фрагменту структуры // Научно-техническая информация. Сер. 2. 2014. № 12. С. 19–29.

7. **Система** поиска химических структур в интерактивном режиме. URL: <http://chem.viniti.ru/> (data obrashcheniia: 9.09.2022).
8. **Trepalin S. V., Bessonov Yu. E., Fel'dman B. S. [et al.]** The Structural Chemical Database of the All-Russian Institute for Scientific and Technical Information, Russian Academy of Sciences. An Autonomous System for Structural Searches // Automatic Documentation and Mathematical Linguistics. 2018. Vol. 52. № 6. P. 297–305.
9. **Trepalin S. V.** Svidetel'stvo o gosudarstvennoi` registratsii programmy` dlia E`VM № 2017613588 ChemDB. Pravoobladatel`: Federal`noe gosudarstvennoe biudzhethoe uchrezhdenie nauki Vserossii`skii` institut nauchnoi` i tekhnicheskoi` informatsii Rossii`skoi` akademii nauk (VINITI RAN) (RU). Zaiavka № 201760837. Data postupleniia 01 fevralia 2017 g. Data gosudarstvennoi` registratsii v Reestre programm dlia E`VM 22 marta 2017 g.
10. **Bessonov Iu. E., Fel'dman B. S., Churakova N. I. [i dr.]** Poisk i otobrazhenie informatsii o himicheskikh reaktsiiakh v baze strukturny`kh danny`kh po himii VINITI RAN // Nauchno-tekhnicheskaiia informatsiia. Ser. 2. 2022. № 3. S. 10–22.
11. **Voronezhva N. I., Trepalin S. V., Churakova N. I. et al.** Glossary as an Element of Data Input Standardization in the Cbase32 Program Complex // Automatic Documentation and Mathematical Linguistics. 2007. Vol. 41. № 3. P. 124–129.
12. **Vatcuro K. V., Mishchenko G. L.** Imenny`e reaktsii v organicheskoi` himii. Moskva : Himiia, 1976. 528 s.
13. **Lee Dzh. Dzh.** Imenny`e reaktsii. Mehanizmy` organicheskikh reaktsii`. Moskva : BINOM. Laboratoriia znanii`, 2006. 456 s.
14. **Nefedov O. M., Koroleva L. M., Trepalin S. V., Bessonov Yu. E. et. al.** The Development of an Integrated System for Structural Chemical Information. Automatic Documentation and Mathematical Linguistics. 2015. Vol. 49. № 6. P. 213–220.

Информация об авторах / Information about the authors

Чуракова Наталия Исааковна – канд. хим. наук, заведующая отделом исследований и обработки структурной химической информации Всероссийского института научной и технической информации РАН, Москва, Российская Федерация
nichurak@rambler.ru

Natalia I. Churakova – Cand. Sc. (Chemistry), Head, Department for Structural Chemical Information Studies and Processing, RAS All-Russian Institute for Scientific and Technical Information, Moscow, Russian Federation
nichurak@rambler.ru

Бессонов Юрий Ефимович – канд. техн. наук, ведущий научный сотрудник отдела исследований и обработки структурной химической информации Всероссийского института научной и технической информации РАН, Москва, Российская Федерация
bessonov-ye@rambler.ru

Фельдман Борис Семёнович – старший научный сотрудник отдела исследований и обработки структурной химической информации Всероссийского института научной и технической информации РАН, Москва, Российская Федерация
bsf@inbox.ru

Червинская Надежда Викторовна – врио директора Всероссийского института научной и технической информации РАН, Москва, Российская Федерация
dir@viniti.ru

Yury E. Bessonov – Cand. Sc. (Engineering), Leading Researcher, Department for Structural Chemical Information Studies and Processing, RAS All-Russian Institute for Scientific and Technical Information, Moscow, Russian Federation
bessonov-ye@rambler.ru

Boris S. Feldman – Senior Researcher, Department for Structural Chemical Information Studies and Processing, RAS All-Russian Institute for Scientific and Technical Information, Moscow, Russian Federation
bsf@inbox.ru

Nadezhda V. Chervinskaya – Acting Director, RAS All-Russian Institute for Scientific and Technical Information, Moscow, Russian Federation
dir@viniti.ru