

**УТВЕРЖДАЮ**

Проректор по учебной работе

А.Л. Голстик

26.06.14

(дата утверждения)

Регистрационный № УД-113/Баз.



## **ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ХИМИИ**

**Учебная программа для специальности**

**1-31 05 01 Химия (по направлениям)**

Направления специальности:

1-31 05 01-01 научно-производственная деятельность

1-31 05 01-02 научно-педагогическая деятельность

1-31 05 01-04 охрана окружающей среды

**1-31 05 02 Химия лекарственных соединений**

**1-31 05 03 Химия высоких энергий**

**1-31 05 04 Фундаментальная химия**

**СОСТАВИТЕЛЬ:**

А. А. Рагойша, доцент кафедры общей химии и методики преподавания химии Белорусского государственного университета.

**РЕЦЕНЗЕНТЫ:**

Ю. С. Головки, доцент Межкафедрального центра — кафедры ЮНЕСКО по естественно-научному образованию, кандидат химических наук, доцент;

А. С. Тихонов, доцент кафедры химии Белорусского государственного педагогического университета имени Максима Танка, кандидат химических наук, доцент.

**РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:**

Кафедрой общей химии и методики преподавания химии Белорусского государственного университета  
(протокол № 10 от 21.04.14);

Учебно-методической комиссией химического факультета Белорусского государственного университета  
(протокол № 5 от 26.05.14).

Ответственный за редакцию: А. А. Рагойша

Ответственный за выпуск: А. А. Рагойша

## ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Информационные технологии, согласно стандарту СТУ ОП 6.3-01-02-2011, — это совокупность процессов, методов осуществления поиска, получения, передачи, сбора, обработки, накопления, хранения, распространения и (или) предоставления информации, а также пользования информацией и защиты информации.

В настоящее время новейшие идеи и достижения информационных технологий оперативно внедряются в Интернете, который становится основным информационным источником научной сферы. С учетом этой тенденции актуальной задачей является формирование у студентов отношения к Интернету как к рутинному инструменту учебной и научной деятельности.

Основные задачи данного курса заключаются в выработке навыков осмысленной работы с наиболее важными онлайн-информационными ресурсами и поисковыми инструментами, в отработке приемов оценки достоверности документа, освоении специализированных средств конструирования, визуализации химических структур и прогнозирования физико-химических параметров, обнаружения спектральных характеристик химических объектов.

В первой части курса проводится обучение студентов-химиков начальным навыкам работы с текстовыми онлайн-научными ресурсами: журналами, материалами конференций, диссертациями, книгами, полнотекстовыми, реферативными и библиографическими базами данных, патентами, нормативными документами, справочниками. Рассматриваются правила формулирования текстового запроса, алгоритмы извлечения информации и ее анализа.

Вторая часть курса знакомит обучающихся со структурными базами данных. На этом этапе вырабатываются навыки структурного поиска, анализа строения кристаллических и молекулярных химических объектов, методов извлечения ЯМР-, ИК-, масс-спектров, идентификации вещества по его спектральным данным.

В учебном курсе используются онлайн- и офлайн-информационные источники и прикладные программы; методическая поддержка обеспечивается использованием специально созданного веб-сайта.

После изучения данного курса обучаемый должен:

**знать:**

- синтаксис многокомпонентного текстового запроса;
- типовую структуру сайта издательства, научного журнала, агрегатора научных статей, патентной базы данных, справочной базы данных;
- типовую структуру онлайн-научной и научно-технологической публикации;

- методы информационного поиска на сайтах основных научных издательств (ACS, RSC, Elsevier, Springer, Wiley-Blackwell и др.), в патентных базах данных (espacenet, USPTO и др.);
- способы отражения структуры вещества в форме линейной нотации;
- область целесообразного использования CAS RN, SMILES, InChI в информационном поиске;
- приемы извлечения кристаллографической информации и структуру CIF-файлов;
- приемы работы со структурными апплетами, с программой Mercury, с пакетом программ ACD/Labs Freeware;

***уметь:***

- оценивать достоверность информационного источника и анализировать его содержание;
- осмысленно пользоваться вспомогательными инструментами онлайн-информационного поиска;
- осмысленно отбирать круг ресурсов, предположительно содержащих искомую информацию;
- вести целенаправленный поиск искомой информации и оперативно корректировать алгоритм работы;
- формулировать структурный запрос для поиска по структуре, по субструктуре, по степени подобия;
- обнаруживать ЯМР-, ИК-, масс-спектры заданных веществ;
- идентифицировать вещество по ЯМР-спектру, используя возможности базы данных SDBS.

Контроль усвоения знаний, навыков и умений осуществляется на всех практических занятиях в форме проверочных и контрольных работ.

Курс рассчитан на 72 аудиторных часа: 14 часов лекций, 48 часов практических занятий в компьютерном классе, подключенном к сети Интернет, 10 часов контролируемой самостоятельной работы.

## УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА

Номер раздела, темы занятия	Название раздела, темы, занятия; перечень изучаемых вопросов	Количество аудиторных часов				Форма контроля знаний
		лекции	практические (семинарские) занятия	лабораторные занятия	самостоятельн ая работа	
1	2	3	4	5	6	7
	<b>1. Онлайн-текстовые научные базы данных</b>	<b>8</b>	<b>24</b>	<b>-</b>	<b>6</b>	
1	Терминология Интернета. Проблема достоверности онлайн-ресурсов.	2				
2	Синтаксис запроса в текстовых базах данных.	2				
3	Первичные и вторичные источники научной информации по химии	4				
4	Поиск и анализ химической информации на сайтах научных журналов		8			Прове- рочные задания.
5	Нежурнальные первичные информационные источники.		4			Прове- рочные задания.
6	Библиографические и реферативные базы данных.		3		1	Конт- рольные задания.
7	Патентные базы данных.		3		1	Конт- рольные задания.
8	Нормативная литература: стандарты, сертификаты безопасности веществ.		3		1	Конт- рольные задания.
9	Текстовый и числовой поиск в справочных базах данных		3		1	Конт- рольные задания.
10	Итоговая контрольная работа				2	Конт- рольные задания.

1	2	3	4	5	6	7
	<b>2. Онлайн-структурные базы данных</b>	<b>6</b>	<b>24</b>	<b>-</b>	<b>4</b>	
11	Способы условного отображения химического вещества в форме текстовой строки.	2				
12	Структурная формула и трехмерная молекулярная модель как информационные пакеты и как объекты хемоинформатики.	4				
13	Кристаллографические базы данных и визуализация кристаллических структур.		3		1	Контрольные задания.
14	Структурный поиск. Формирование запроса с помощью апплета.		4			Проверочные задания.
15	Линейные нотации. CAS RN, SMILES, InChI в информационном поиске.		4			Проверочные задания.
16	Онлайн-базы данных по органическому синтезу.		3		1	Контрольные задания.
17	Онлайн-спектральные базы данных (ЯМР, ИК, MS).		6			Проверочные задания.
18	Пакет программ ACD/Labs Freeware: молекулярный, графический редактор; моделирование трехмерных структур.		4			Проверочные задания.
19	Итоговая контрольная работа				2	Контрольные задания.

## СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

### **1. Терминология Интернета. Типы онлайн-источников научной информации. Вспомогательные инструменты. Методы оценки достоверности онлайн-ресурсов.**

Интернет — ценный, увеличивающийся в объеме и в степени доступности источник научной информации.

Терминология Интернета. Протокол HTTP, гипертекст, гиперсвязь, World Wide Web, веб-сервер, IP-адрес, домен и доменное имя, сайт, веб-страница, Главная страница сайта, URL, поисковая система, метапоисковая система, тематический каталог, метасайт, портал, браузер.

Два режима получения порций информации: перелистывание (Browse) и поиск (Search).

Общая характеристика онлайн-информационных источников. Изолированный документ; группа веб-страниц; база данных. Информация текстовая, числовая, графическая, мультимедийная; химическая структура как информационный пакет. Интерактивные элементы веб-страниц.

Проблема анонимности и авторства онлайн-документов. Проблема достоверности материалов. Первоисточник и интерпретация. Стандартные приемы формальной оценки степени достоверности онлайн-информационного источника.

Химическая информация на сайтах университетов, научных обществ, коммерческих организаций, на персональных сайтах ученых. Новости науки на веб-сайтах. Электронные каталоги библиотек.

Вспомогательные инструменты информационного поиска.

Методика распознавания смысла сокращений в химических текстах. Словари акронимов.

Онлайновые универсальные и специализированные энциклопедии.

Википедия. Принципы формирования, редактирования и рецензирования Википедии. Степень достоверности информации, публикуемой в Википедии. Сравнение объема и качества химической информации в английской и российской версиях. Википедия как стартовая точка информационного поиска.

Сайт Answers.com как стартовая точка информационного поиска.

Языковые инструменты Интернета. Словари на порталах российских поисковых систем и на специализированных лингвистических порталах.

Автоматизированный перевод фрагментов текста и веб-страниц с помощью онлайн-средств. Информационный поиск по документам, язык которых незнаком пользователю.

## **2. Текстовые базы данных.**

Классификация баз данных по типу хранящейся информации: текстовые, числовые, содержащие структурные формулы веществ, содержащие сведения о химических реакциях, спектральные, кристаллографические; архивы программ. Классификация по способу доступа: офлайновые и онлайнные. Метабазы данных. Литературные базы данных: библиографические, реферативные, полнотекстовые.

Структура текстовой базы данных: запись, поле, вспомогательные указатели (индексы). Принципиальная схема поиска и извлечения информации. Поисковый процесс как посимвольное сравнение текста задания с текстами, хранящимися в базе данных. Запрос (поисковое задание). Синтаксис запроса. Логические (булевы операторы). Оператор по умолчанию. Операторы расстояния. Использование шаблона для отображения словоформ; для отображения варьируемых фрагментов термина. Режим stemming как способ учета словоформ поисковых терминов. Регистр букв. Стоп-слова. Формулирование запроса на естественном языке. Поле как элемент поискового задания. Поисковые бланки: простейшие и усложненные. Список результатов поиска. Ранжирование списка результатов по степени соответствия поисковому заданию — релевантности.

## **3. Универсальные поисковые средства.**

Принципы работы поисковых средств (поисковой системы, тематического каталога, метапоисковой системы).

Поисковая система. Структура поисковой системы (робот, индекс, пользовательская поисковая программа). Автоматизированный способ формирования информационной базы данных поисковой системы. Особенности информационного поиска в индексе поисковой системы: поиск только по терминам, которые присутствуют в документе. Особенности работы поисковых программ, ориентированных на профессионалов; на широкий круг пользователей.

Особенности синтаксиса запроса для текстовой поисковой системы (простейший бланк). Конструирование запроса на усложненном бланке.

Список результатов поиска. Ранжирование по релевантности. Структурные части списка. Сведения о статистике поиска. Копии документов в кэше поисковой системы. Приемы уточнения поиска.

Тематический каталог. Многоуровневая древовидная структура. Безмашинный способ формирования базы данных тематического каталога.

Аннотация сайта как структурная информационная единица тематического каталога. Принципиальные отличия в использовании терминов в текстах исходного документа и аннотации.

Особенности выбора терминов для поиска в базе данных



тематического каталога и для поиска в индексе поисковой системы.

Принципы функционирования метапоисковой системы. Область целесообразного использования метапоисковой системы.

Понятие о видимом и скрытом (глубоком) Интернете. Стратегии извлечения информации в каждом из этих сегментов Интернета.

#### **4. Приемы работы с основными универсальными поисковыми средствами.**

Поисковая система Google. Краткое описание поисковой системы. Объем проиндексированной информации. Лидирующее положение в эффективности ранжирования извлекаемых документов. Многоязычность интерфейсов. Поисковая система как комплекс средств для обнаружения ресурсов разнообразной структуры и различного предназначения.

Главная страница (английский, русский, белорусский интерфейсы).

Синтаксис запроса, формулируемого на основном поисковом бланке. Правила формулирования запроса на усложненном бланке.

Стандартные приемы формулирования запроса при ведении тематического поиска, поиска по автору, по типу извлекаемых документов. Структура элементов списка документов, обнаруженных в результате поиска. Приемы формального анализа списка в целях отсеивания недостоверного материала и отбора ограниченного числа записей, максимально соответствующих запросу. Приемы оперативной коррекции запроса.

Сравнение списков, извлеченных англоязычной и русскоязычной поисковыми программами. Области целесообразного использования каждой из языковых версий.

Языковые инструменты Google.

Портал Яндекс: поисковая система, тематический каталог, вспомогательные информационные ресурсы.

Синтаксис запроса, формулируемого на основном поисковом бланке. Специфика использования регистра букв в запросе, учета словоформ поисковых терминов. Краткие сведения о функционировании операторов расстояния.

Правила формулирования запроса на усложненном бланке.

Особенности обнаружения информации в каталоге.

Область целесообразного использования комплекса поисковых средств Яндекса.

Тематический каталог и поисковая система Yahoo! Область целесообразного использования.

Основные характеристики и приемы работы с иными зарубежными универсальными поисковыми системами (Рамблер, Bing и др.).

Тематический каталог Open Directory. Принципы формирования, редактирования и рецензирования каталога. Многоуровневая структура; место сайтов химической тематики в каталоге; место русскоязычных сайтов. Поиск информации в каталоге методами Browse и Search.

Основные характеристики и приемы работы с универсальными поисковыми инструментами Беларуси. Поисковые ресурсы портала TUT.BY. Поисковая система и тематический каталог Open.BY. Акавита: поисковая система, каталог и инструмент рейтинга сайтов.

Метапоисковые системы. Кластерирование результатов поиска. Преимущества и недостатки метапоисковых систем. Области их использования.

Метасайты–каталоги поисковых систем.

Специализированные тематические каталоги и метасайты. Общая характеристика, структура, тематика каталогов, содержащих информацию об онлайн-химических ресурсах. Примеры каталогов.

Портал как информационный центр научного сообщества.

Ресурсы химических порталов Chemport.ru, Chemie.de, ChemSpider.

## **5. Офлайновые и онлайн-научные публикации.**

Научные публикации. Типы печатных научных изданий. Первичные и вторичные источники научной информации. Рецензируемые и нерецензируемые источники информации.

Научный журнал как архив научных знаний и как инструмент оценки качества результатов научной деятельности. Структура журнала. Типы публикаций в научном журнале (статья, краткое сообщение, письмо в редакцию, обзор). Рецензируемые и нерецензируемые части журнала.

Научная статья. Структура статьи (название, авторы, реферат или аннотация, основной текст, ссылки и список использованной литературы, вспомогательные элементы статьи).

Препринт, постпринт.

Материалы конференции. Тезисы докладов, расширенные тезисы, презентации, труды конференции. Диссертация. Автореферат диссертации. Научный отчет. Депонированная рукопись. Продолжающееся издание. Книга: монография, сборник статей, учебное пособие.

Реферативный журнал. Структура журнала. Структура отдельной записи в реферативном журнале. Универсальные и специализированные реферативные журналы.

Библиографический указатель.

Библиографическое описание научной публикации: краткое и полное. Импакт-факторы научных журналов. Индексы цитирования.

Онлайновые научные журналы. Общая характеристика онлайновых периодических и продолжающихся научных изданий. Преимущества онлайновых изданий по сравнению с печатными.

Платные и бесплатные ресурсы. Причины размещения бесплатной информации в Интернете. Инициатива Open Access. Приемы визуального определения общедоступности полного текста публикации.

Постоянно доступные ресурсы и временно доступные ресурсы. Виды открытого доступа к отдельным частям тома научного журнала или его выпуска. Период эмбарго и движущаяся граница доступности.

Форматы онлайновых публикаций. Организация информационного массива онлайнового журнала как комплекса PDF- и HTML-документов.

Информационные ресурсы издательства в открытом и скрытом Интернете.

Структура сайта издательства и сайта журнала. Структурные элементы Главной страницы издательства, Главной страницы журнала. Работа в режиме Browse. Стандартная цепочка гиперсвязей от Главной страницы сайта издательства к тексту статьи. Структура оглавления выпуска журнала; списка статей, подготовленных к публикации; страницы реферата; статьи в формате PDF; статьи в формате HTML. Иллюстрации встроенные и внешние. Ссылки в статье. Список использованной литературы и гиперсвязи, ведущие к процитированным документам или к записям в реферативных базах данных. Приложения к статьям, публикуемые только в электронном формате.

Особенности электронных журналов, не имеющих печатных аналогов. Библиографическое описание онлайн-научной публикации в таком журнале.

Особенности сохранения статьи в HTML-формате, содержащей внешние иллюстрации, на компьютере пользователя.

Методы открытого онлайн-рецензирования научных публикаций.

Текстовые поисковые программы на сайтах издательств. Инструменты для поиска по открытому и для поиска по закрытому Интернету. Навигационные элементы на страницах сайта. Структура поисковых бланков Quick, Basic, Advanced Search. Области применения каждого из типов бланков. Особенности работы с бланком Quick Search.

Общие правила формулирования текстового запроса для поисковой программы издательства. Методика поиска информации по сведениям об авторе публикации. Особенности поиска по фамилии с диакритическими знаками. Транслитерация кириллических фамилий. Методика проведения

тематического поиска. Отбор поисковых терминов для запроса. Тематический поиск по полям записей в базе данных. Возможности поиска по брутто-формулам химических веществ.

Онлайновые адреса статей. Причины непостоянства и неопределенности URL статьи. Динамические адреса онлайн-объектов. Принципы работы CrossRef — агентства, регистрирующего онлайн-научные объекты. Цифровой идентификатор объекта DOI. Структура кода DOI. Использование DOI в онлайн- и печатной литературе в целях однозначного описания местонахождения публикации в Интернете.

## **6. Основные издательства химической научной периодики.**

Портал ScienceDirect — информационный центр издательства Elsevier. Портал SpringerLink — информационный центр издательства Springer. Портал Wiley Online Library — информационный центр издательства Wiley-Blackwell. Сайты издательств научных обществ American Chemical Society и Royal Society of Chemistry.

Общая характеристика типа, объема, доступности информации, размещенной на сайтах. Распознавание платных и бесплатных порций информации. Работа в режиме Browse. Обнаружение цепочки гиперсвязей, ведущих с Главной страницы издательства к материалам конкретной научной публикации. Степени доступности оглавлений, рефератов, полных текстов статей, препринтов. Структура онлайн-статьи. Гиперсвязи в списке процитированной литературы; базы данных, к которым ведут гиперсвязи. Возможность получения списка статей, цитирующих данную публикацию.

Работа в режиме Search. Типы поисковых бланков, их достоинства, недостатки, области применения. Синтаксис текстового запроса.

Дополнительные сервисы, предоставляемые зарегистрированным пользователям. Возможность сохранения запросов и автоматического уведомления о новых поступлениях в базу данных.

Краткая характеристика сайтов других издательств, публикующих химическую научную периодику (Bentham Science, Taylor & Francis Group, МАИК Наука /Interperiodica, IOP Publishing, НАН Беларуси и др.).

Инструменты обнаружения онлайн-журналов. Методика работы с метасайтами, содержащими списки журналов и их адреса. Методика работы с базой данных Genamics JournalSeek.

Списки бесплатных полнотекстовых научных химических журналов на сайте ABC-Chemistry: постоянно и временно доступные журналы.

Полнотекстовые журналы, доступные из сети БГУ, из внутренних сетей библиотек Беларуси.

Онлайновые инструменты, предназначенные для нахождения полного названия журнала по его сокращенному названию.

## **7. Научные публикации в неперидических изданиях.**

Внежурнальные электронные публикации. Типы публикаций, объединяемых термином e-prints.

Материалы научных конференций. Типичная структура сайта конференции на разных стадиях подготовки и проведения мероприятия, а также после его окончания. Программа конференции, сведения об авторах и докладчиках. Онлайновые тезисы докладов, презентации, тексты докладов; их местонахождение на сайте. Приемы обнаружения сведений о конференции с помощью метасайтов, а также с помощью универсальных поисковых систем. Приемы обнаружения научных работ, представленных на конференции.

Онлайновые формы проведения научной конференции.

Университетские, национальные и международные репозитории научных публикаций. Поиск репозитория по его тематике, местонахождению, наполнению; полнотекстовый поиск в документах, хранящихся в репозиториях; получение перечня организаций — владельцев репозитория.

Специализированные базы данных с материалами диссертаций. Электронная библиотека "Авторефераты" на сайте ВАК Беларуси. Электронная библиотека авторефератов диссертаций Украины, Электронная библиотека диссертаций Российской государственной библиотеки.

Книги. Финансовые и юридические проблемы, ограничивающие распространение книжной продукции через Интернет. Копирайт. Форматы представления книжной продукции в Интернете. Электронные библиотеки. Архивы книг по химии и смежным дисциплинам; архивы книг универсальной тематики. Методика поиска книг с помощью универсальных поисковых систем. Специализированные поисковые системы Рунета, примеры и принципы использования.

Поисковая система Google Book Search. Структура, объем и тип информации в базе данных. Методика извлечения информации. Список результатов поиска, индикация объема доступности материалов; проблема сохранения извлеченной информации на локальном компьютере.

## **8. Агрегаторы. Библиографические и реферативные базы данных.**

Агрегаторы. Общенациональные и региональные агрегаторы

журналов (на примере J-STAGE и SCIELO). Тематический агрегатор PubMed Central. Коммерческие агрегаторы (EBSCOhost, IngentaConnect и др.).

Агрегатор EBSCOhost. Общая характеристика объема, доступности информации, размещенной в базах данных агрегатора.

База данных Academic Search Complete. Полнотекстовые, реферативные и библиографические ресурсы. Структура Главной страницы; навигационные элементы; поисковые бланки. Синтаксис текстового запроса. Кластерированный список результатов поиска. Вспомогательные указатели.

Общая характеристика баз данных Medline и Eric.

Directory of Open Access Journals (DOAJ) как база данных: библиографическая, реферативная и полнотекстовая. Поиск информации в DOAJ.

Реферативный журнал "Химия". Общая характеристика онлайн-базы данных. Работа в режимах Browse и Search. Поиск по каждому из полей, по их комбинации и по всей записи в целом.

Научная электронная библиотека eLibrary. Объем информации в библиотеке; степени ее доступности. Методика информационного поиска в базе данных.

Сведения о литературе по радиационной химии, радиохимии, экологии, материаловедению в базе данных INIS.

Принципы использования информационных ресурсов коммерческого агрегатора как общедоступной реферативной или библиографической базы данных.

Специализированная поисковая система Google Scholar. Типы первоисточников, индексируемых поисковой системой. Отличие специализированной поисковой системы Google Scholar от универсальной Google: преимущества и недостатки. Синтаксис запроса. Структура списка результатов поиска. Возможность использования Google Scholar для обнаружения библиографической, реферативной и полнотекстовой информации.

Общее понятие о содержании и структуре ресурсов, хранящихся в базах данных ISI Web of Knowledge и Scopus.

Методы обнаружения онлайн-версий научных статей по известным библиографическим данным.

Поиск статьи по коду DOI.

Общие сведения о программах, предназначенных для создания библиографических баз данных индивидуального пользования.

Онлайновые инструменты хранения библиографических данных. Методика работы с сервисом CiteULike.

## **9. Патентные базы данных.**

Патентная литература. Патент, патентная заявка, авторское свидетельство. Патент как юридический документ и как источник научной информации. Структура патентного документа: страница библиографического описания, формула изобретения, описание изобретения. Понятие о патентном семействе.

Национальные, международные и региональные патентные бюро. Общие сведения о процедуре регистрации изобретения.

Патентная классификация.

Основные онлайн-патентные базы данных.

Характеристика объема и типа информации, содержащейся в базах данных информационного центра Европейского патентного бюро Espacenet. Структура Главной страницы сайта, назначение ссылок, находящихся на Главной странице. Усложненный поисковый бланк: структура, методика формулирования запроса. Синтаксис запроса. Особенности выбора языка запроса для поиска в каждой из баз данных информационного центра.

Структура списка результатов поиска. Методика коррекции поискового процесса. Структура страницы патента. Получение факсимильной копии патента.

Серверы национальных патентных бюро на портале Espacenet: особенности информационного наполнения.

Патентная база данных Национального центра интеллектуальной собственности Республики Беларусь.

Базы данных USPTO. Методика формулирования запроса и коррекции поискового процесса.

Базы данных Роспатента. Краткая характеристика объема и типа информации. Методика информационного поиска.

Базы данных патентов Японии. Методика поиска информации в базе данных PAJ. Машинный перевод текста с японского языка на английский.

Патентные базы данных КНР, стран Юго-Восточной Азии. Сайт Евразийской патентной организации.

Специализированная поисковая система Google Patents. Область применения поисковой системы и сравнение возможностей Google Patents с другими аналогичными инструментами патентного поиска.

## **10. Научная коммуникация и онлайн-коллективная научная работа.**

Научный форум как источник информации, как оперативно работающий справочный инструмент. Форумы на портале Chemport.ru.

Этика онлайн-общения — netiquette.

Научные и учебные ресурсы категории Web 2.0.

Подкасты с материалами научной химической тематики.

Блог как источник научной информации. Степень достоверности данных, публикуемых в блоге. Примеры блогов научных сотрудников.

«Открытая наука». Открытый лабораторный журнал (Open Notebook). Выполнение научного проекта неформальным Интернет-коллективом (Crowdsourcing).

Научные социальные сети.

Web 3.0. Проблема машинного распознавания смыслового содержания научного документа; проблема распознавания химических объектов в научных документах. Метаданные в теле научного документа. Химическая онтология. Прототипы семантического веба.

## **11. Способы условного отображения химического вещества в форме текстовой строки.**

Регистрационные номера химических веществ, используемые в больших базах данных.

CAS Registry Number: алгоритм нумерации химических объектов, формат записи. Проблема соотнесения кода и вещества. Онлайн- и офлайн-источники сведений о CAS RN. Возможность использования CAS RN в информационном поиске.

Отображение состава и структуры химического вещества в форме линейной записи (линейной нотации).

Номенклатура ИЮПАК. Офлайн- и онлайн-средства генерирования систематических названий веществ. Проблемы, возникающие при использовании систематического названия химического вещества в качестве компонента поискового задания. Особенности ведения информационного поиска по названию вещества в индексе универсальной поисковой системы и в специализированной справочной базе данных.

Онлайн-формульный указатель. Порядок размещения молекулярных формул химических веществ по системе Хилла.



Линейные нотации SMILES. Причины, обусловившие потребность в подобной форме отображения состава и структуры вещества. Основные правила формулирования кода SMILES (кодирование линейных, разветвленных, циклических, ароматических молекулярных структур, ионных соединений, стереоизомеров, схем химических реакций). Офлайновые и онлайнные средства генерирования кодов SMILES. Достоинства и недостатки SMILES в сфере кодирования химических структур. Онлайнные базы данных, содержащие SMILES. Использование SMILES в информационном поиске.

Международный химический идентификатор InChI. Причины, обусловившие потребность в разработке InChI. Общее представление о формате кода InChI (слоевая структура, набор слоев в текущей версии спецификации, подслои главного слоя). Понятие об InChIKey. Офлайновые и онлайнные средства генерирования кодов InChI. Использование InChI и InChIKey в информационном поиске.

## **12. Двумерное и трехмерное отображение химической структуры.**

Двумерные и трехмерные модели структуры химического вещества. Параметры трехмерной модели: координаты атомов, углы между связями, торсионный (двугранный) угол. Способы визуализации трехмерной структуры; проволоочная, стержневая, шаростержневая, объемная (СРК) модели. Визуализация трехмерной структуры биомолекул. Многогранники как элементы визуализации трехмерной структуры неорганических соединений.

Молекулярная поверхность. Поверхности Ван-дер-Ваальса, Конноли, Ли-Ричардса. Поверхности: доступная растворителю и недоступная. Визуальное отображение электростатического потенциала.

Понятие о структуре Маркуша. Принципы использования структур Маркуша в патентных базах данных.

Расчетные и экспериментально определенные параметры — две категории сведений о структуре вещества.

Молекулярные редакторы и прикладные программы генерирования и визуализации двумерных и трехмерных молекулярных структур и расчета молекулярных параметров.

Пакет программ ACDLabs Freeware: молекулярный редактор, графический редактор, генератор трехмерных моделей. Приемы конструирования и редактирования графических формул и иных химических объектов. Изготовление иллюстративного материала; конвертирование двумерных структур в трехмерные; создание файлов с информацией о структуре вещества; формирование заданий для ведения структурного поиска

в базах данных. Генерирование названий веществ, линейных нотаций; расчет молекулярных параметров.

Апплеты и плагины, предназначенные для конструирования графических формул; для визуализации трехмерной структуры.

Стандартные форматы файлов, предназначенных для хранения данных о молекулярной структуре.

Формат MDL Molfile. Таблица соединений как составная часть файла. Параметры молекулярной структуры, которые могут храниться в MOL-файле. Прикладные программы, способные генерировать MOL-файлы. Область использования файлов в формате MOL в качестве запроса в информационном поиске.

Стандарт Международного союза кристаллографии, регламентирующий порядок хранения кристаллографической информации о низкомолекулярных веществах. Файловый формат CIF: текстовая структура, блоки `_loop`. Кристаллографическая информация в онлайн-научных журналах. Онлайн-архивы CIF-файлов. Прикладные программы визуализации и анализа информации, хранящейся в CIF-файлах.

Приемы работы с программой Mercury: визуализация кристаллической структуры вещества; определение структурных параметров.

Структура, субструктура. Подобие двумерных и трехмерных структур. Структурный код, "отпечаток пальцев". Коэффициент Танимото. Принципы использования коэффициента Танимото при информационном поиске по подобию.

Понятие о молекулярных и физико-химических дескрипторах, двумерном и трехмерном фармакофоре, молекулярном докинге, Data Mining, QSAR и QSPR. Принципы предсказания биологической активности вещества, обнаружения вещества с заданными свойствами в структурной базе данных методами хемоинформатики. "Правило пяти" Липинского. Место хемоинформатики в системе наук.

Онлайн-базы данных, в которых реализован поиск по структуре, субструктуре и по подобию структур. Способы формирования запроса и коррекции информационного поиска.

Принципы информационного поиска в онлайн-базах данных, содержащих сведения о химических реакциях.

### **13. Справочные базы данных.**

Справочные базы данных, содержащие сведения о физических и химических свойствах химических веществ, смесей, материалов. Базы данных, содержащие информацию о структуре вещества.

Приемы формулирования запросов разных типов: текстового, числового, структурного — для целенаправленного поиска справочной информации.

Сайты ведущих компаний, производящих химические товары. Каталоги химических реактивов. Методика поиска сведений о реактиве и о фирме-изготовителе. Критерии определения оптимальной цены реактива. Структурный поиск. Метабазы данных о реактивах.

ГОСТ, ТУ, СанПиН как источники химической информации. Базы данных, содержащие сведения о национальных и межгосударственных стандартах и технологических спецификациях.

Сертификаты безопасности материала (MSDS). Характеристика достоверности информации, имеющейся в сертификате. Онлайн-базы данных, содержащие MSDS; метасайты.

Базы данных NIST. Chemistry WebBook как источник достоверных сведений в области химии. Справочник по термодинамическим и термодинамическим свойствам веществ и ионов NIST Standard Reference Database. Главная страница базы данных; назначение гиперссылок; поисковые бланки. Методика извлечения информации о веществе. Текстовый поиск. Поиск по структуре.

Термодинамические базы данных на сайте химфака МГУ.

Графическое отображение ЯМР-, ИК-, масс-спектров. Файловый стандарт обмена спектральными данными JCAMP-DX. Апплет как инструмент отображения интерактивного спектра.

Онлайн-репозитории спектральных данных. Информационный поиск в спектральных базах данных и анализ результатов.

Спектры веществ в NIST Chemistry WebBook.

Комплекс баз данных RIO-DB Национального института современной индустриальной науки и технологии (Япония) как источник достоверных сведений о строении и свойствах химических веществ и материалов. Спектроскопические ресурсы RIO-DB. Методика извлечения спектральной информации (ЯМР-, масс-, ИК-спектров) из базы данных SDBS.

ChemSpider как онлайн-центр структурной информации. Принципы формирования и редактирования базы данных; границы достоверности сведений, содержащихся в ChemSpider. Методика извлечения спектральной информации.

Онлайн-источники информации о методиках органического синтеза. Приемы извлечения и анализа сведений из баз данных ChemSynthesis, ChemSpider SyntheticPages, Organic Syntheses.

Репозитории химических структур. Принципы формирования базы данных и коррекции ее содержимого. Границы достоверности извлекаемой информации.

Метасайты, содержащие адреса, описания и рекомендации по использованию бесплатных справочных баз данных.

## ОСНОВНАЯ ЛИТЕРАТУРА

1. А. А. Рагойша. Поиск химической информации в Интернете. Поисковые системы и тематические каталоги: Учеб. пособие для студентов хим. фак. – Мн.: БГУ, 2003.
2. А. А. Рагойша. Поиск химической информации в Интернете: научные публикации : учеб. пособие для студентов хим. фак. спец. 1-31 05 01. – Мн.: БГУ, 2007.
3. А. А. Рагойша. Текстовый поиск научной химической информации в Интернете [Электронный ресурс] : практикум по курсу "Информационные технологии в химии" для студентов спец. 1-31 05 01 Химия (по направлениям). — Минск: БГУ, 2012. — <http://elib.bsu.by/handle/123456789/14599>.
4. А. А. Рагойша. Поиск информации о структуре химического вещества в онлайн-базах данных : практикум по курсу "Информационные технологии в химии" для студентов специальности 1-31 05 01 Химия (по направлениям). – Минск : БГУ, 2013. – <http://elib.bsu.by/handle/123456789/44282>.
5. В. М. Потапов, Э. К. Кочетова. Химическая информация. Где и как искать химику нужные сведения. – М.: Химия, 1988.
6. А. А. Рагойша. Азбука веб-поиска для химиков [Электронный ресурс] – Минск, БГУ, 1999-2012. — <http://www.abc.chemistry.bsu.by>.

## ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

7. Chemoinformatics: A Textbook. Edited by Johann Gasteiger and Thomas Engel. – Wiley-VCH, 2003.
8. В. А. Bunin, В. Siesel, G. A. Morales, J. Bajorath. Chemoinformatics: Theory, Practice, & Products. – Springer, 2007.
9. Andrew R. Leach, Valerie J. Gillet. An Introduction to Chemoinformatics. – Springer, 2007.
10. Е. А. Устинова. Формулы изобретения на химические объекты. – М.: ИНФРА-М, 1997.
11. Руководства пользователя к прикладным программам.