

А. А. Рагойша

**ТЕКСТОВЫЙ
ПОИСК
НАУЧНОЙ
ХИМИЧЕСКОЙ
ИНФОРМАЦИИ
В ИНТЕРНЕТЕ**

**Руководство
к практикуму**

**Учебное пособие
по курсу
"Информационные
технологии
в химии"
для студентов
химического факультета
специальности 1-31 05 01**

**МИНСК
БГУ
2011**

УДК 54:004.738.52

Онлайновое учебное издание

http://www.abc.chemistry.bsu.by/lit/Rahoisha_2011.pdf

Рагойша Александр Антонович

Кафедра общей химии и методики преподавания химии
Белорусского государственного университета (БГУ)

Рагойша, А. А.

Текстовый поиск научной химической информации в Интернете : руководство к практикуму : учеб. пособие по курсу "Информационные технологии в химии" для студентов хим. фак. спец. 1-31 05 01 [Электронный ресурс] — Минск: БГУ, 2011. — 64 с. — Режим доступа: http://www.abc.chemistry.bsu.by/lit/Rahoisha_2011.pdf

Методическое руководство к практическим занятиям по разделу "Текстовые научные документы в Интернете" учебного курса "Информационные технологии в химии". В пособии рассматриваются наиболее важные веб-ресурсы, а также алгоритмы целенаправленного поиска, извлечения и анализа научной химической информации.

Для студентов и магистрантов химического факультета БГУ.

УДК 54:004.738.52

© А.А.Рагойша, 2011

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ.....	5
1. УНИВЕРСАЛЬНЫЕ ПОИСКОВЫЕ СРЕДСТВА И ВСПОМОГАТЕЛЬНЫЕ ИНСТРУМЕНТЫ.....	6
1.1. Введение.....	6
1.2. Google.....	6
1.3. Яндекс.....	8
1.4. Другие универсальные поисковые средства.....	9
1.5. Тематические каталоги и метасайты.....	10
1.6. Словари, энциклопедии.....	11
1.7. Онлайн-переводчики текстов и веб-страниц.....	12
1.8. Анализ сайта — самостоятельная работа	13
2. НАУЧНЫЕ ЖУРНАЛЫ.....	15
2.1. Введение.....	15
2.2. Сайты издательств.....	15
2.3. Списки журналов.....	19
2.4. Сайты журналов — дополнение.....	22
2.5.1. Агрегаторы.....	23
2.5.2. eLIBRARY.RU.....	25
2.5.3. EBSCOhost.....	26
2.5.4. ProQuest.....	28
3. НЕЖУРНАЛЬНЫЕ РЕЦЕНЗИРУЕМЫЕ ПУБЛИКАЦИИ.....	31
3.1. Введение.....	31
3.2. Материалы конференций.....	31
3.3.1. Диссертации.....	32
3.3.2. ProQuest Dissertations & Theses.....	34
3.4. Книги.....	34
3.5. Google Books.....	35
4. РЕПОЗИТОРИИ ЭЛЕКТРОННЫХ ПУБЛИКАЦИЙ.....	37
4.1. Введение.....	37
4.2. Репозитории электронных публикаций.....	37
4.3. Персональные сайты ученых.....	39
5. ВТОРИЧНЫЕ ИСТОЧНИКИ НАУЧНОЙ ИНФОРМАЦИИ.....	40
5.1. Введение.....	40
5.2. РЖ Химия — Реферативный журнал "Химия".....	40
5.3. Тематические (специализированные) базы данных.....	41
5.4. Google Scholar.....	42
5.5. Scirus.....	43
5.6. Личная библиографическая база данных в социальной сети.....	44
6. ПАТЕНТНЫЕ БАЗЫ ДАННЫХ (сокращенная версия).....	54
6.1. Введение.....	54
6.2. US Patent and Trademark Office (USPTO).....	54
7. НОРМАТИВНЫЕ ДОКУМЕНТЫ.....	57
7.1. Сертификат безопасности материала (MSDS).....	57
7.2. ГОСТ - Государственный стандарт.....	59
7.3. СанПиН — Санитарные правила и нормы.....	61

8. СПРАВОЧНЫЕ БАЗЫ ДАННЫХ.....	62
8.1. Введение.....	62
8.2. Базы данных для работы в режиме Browse.....	62
8.3. Базы данных для работы в режиме Search.....	63
Bibliographic Record.....	65

ПРЕДИСЛОВИЕ

Текстовый поиск, которым мы займемся в данном разделе практикума, все еще остается основным методом обнаружения научной химической информации в Интернете.

Ресурсы Интернета многочисленны, и растут они постоянно. Наша первая цель — убедиться, что в этом информационном пространстве существуют стартовые точки, предназначенные для эффективного решения стандартных научных поисковых задач стандартными методами.

Ресурсы Интернета качественно неравноценны. Вторая наша цель — попытаться осознать, что *достоверные* источники *достойны* пристального внимания, а все остальные — всего лишь развлекательный материал для ленивого созерцания в минуты досуга.

Сайты, с которыми мы будем работать,— это малая доля того, что может понадобиться исследователю-химику. В будущем, решая творческие задачи и занимаясь самообразованием, каждый из нас сформирует собственный набор онлайн-источников, абсолютно необходимых ему лично.

Теперь же мы будем знакомиться с характерными — и несомненно полезными — примерами разных типов веб-ресурсов и веб-инструментов.

Чтобы работа была успешной, нам следует свободно ориентироваться в вопросах терминологии и теории, о которых идет речь в пособиях:

- А. А. Рагойша. Поиск химической информации в Интернете. Поисковые системы и тематические каталоги. Мн.: БГУ, 2003.
- А. А. Рагойша. Поиск химической информации в Интернете: научные публикации. Мн.: БГУ, 2007

Интернет — среда чрезвычайно изменчивая, поэтому материалы практикума будут постоянно обновляться. Самую последнюю версию данного учебного пособия (в *HTML*-формате) вы можете найти на сайте *ABC Chemistry* — *Азбука веб-поиска для химиков* по адресу:

<http://www.abc.chemistry.bsu.by/intro/>

1. УНИВЕРСАЛЬНЫЕ ПОИСКОВЫЕ СРЕДСТВА И ВСПОМОГАТЕЛЬНЫЕ ИНСТРУМЕНТЫ

1.1. Введение

WWW — это динамическая система: одни документы появляются, другие по прихоти авторов исчезают бесследно. Сколько всего страниц в *World Wide Web* — не знает никто; известно только, что счет идет на многие миллиарды. В Сети отсутствует единый каталог ресурсов, но его функции частично выполняют т.н. **поисковые системы**, или **поисковые службы** (*search engines, search services*). Каждая поисковая система имеет свою программу-"робот", постоянно сканирующую веб-пространство и собирающую в базу данных сведения об обнаруженных документах. Алгоритмы действий роботов подобны, но не идентичны, поэтому и сформированные базы данных в той или иной степени различаются между собой как перечнем собранных адресов, так и структурой и объемом информации, характеризующей отдельный веб-документ.

Обратившись к поисковой системе, пользователь **на бланке формулирует задание**; программа анализирует свою базу данных и сообщает результат — список веб-страниц. Список может быть очень большим, но поскольку пункты его обычно ранжированы по релевантности, пользователю достаточно ознакомиться с началом полученного перечня. Обычными элементами списка являются *URL*, название страницы (либо сайта) и гиперсвязь к ней, фрагмент документа; дополнительно могут быть приведены и другие сведения (например, объем, дата создания, аннотация).

Каждая поисковая система использует свои правила построения поискового задания. Общей тенденцией для всех является отход от жесткой заданности булевой фразы к гибкому алгоритму *free-text search*.

Роботы основных поисковых средств прочитывают веб-документы от начала до конца и вносят в свои базы данных все встреченные слова. Кроме того, они могут накапливать и иную информацию: имена файлов иллюстраций, содержимое служебной части *HTML*-документа и т.д. — этот перечень индивидуален для каждой поисковой службы.

Многие поисковые службы предлагают пользователю **тематические каталоги** веб-страниц. Такие каталоги (***Directories***) создаются безмашинным способом, поэтому в них попадает малая доля материала, имеющегося в *WWW*.

Многие поисковые средства размещены на *порталах*.

Портал — это многопрофильный сайт, предлагающий широкий спектр информационных услуг. Типовой набор включает тематический каталог Web-страниц, средство для информационного поиска в *WWW*, службу новостей, ряд специализированных разделов (например, метеорологический, финансовый и т. д.), а также узел электронной почты.

1.2. Google

В настоящее время поисковая система **Google** является лидером в области обнаружения онлайн-информации.

Главная страница **Google** — англоязычная — находится по адресу

<http://www.google.com/> и содержит простейший бланк для формулирования задания, гиперсвязи перехода к бланкам: *Images* (поиск графической информации), *Maps* (географические карты и снимки из космоса), *News* (сообщения средств массовой информации), *Videos* (поиск видеофрагментов) и др. (ссылка *more*), а также — на кнопке *Options* — гиперсвязи к усложненному бланку (*Advanced Search*) и к языковым средствам (*Language Tools*).

Поисковая система имеет подобные Главные страницы, созданные для пользователей из других стран, в том числе:

[Google-Беларусь \(http://www.google.by/\)](http://www.google.by/), где можно выбрать белорусский или русский интерфейс;

[Google-Россия \(http://www.google.ru/\)](http://www.google.ru/) с русским интерфейсом.

Google бывает навязчиво услужливым, в результате чего по адресу www.google.com на экране может появиться не английский, а русский или белорусский интерфейс. В таких случаях для перехода к английскому интерфейсу следует воспользоваться гиперсвязью *Google.com in English*, размещенной внизу страницы.

В основном способы работы со всеми перечисленными сайтами одинаковы.

Основные правила формулирования поискового задания (простейший бланк):

- Задание составляется из *слов, разделенных пробелами*.
- *Строчные и заглавные* буквы в терминах равнозначны.
- *Оператор по умолчанию: AND*.
- Термины можно объединять оператором *OR*, напечатанным **заглавными** буквами.
- Оператор *NOT* обозначается знаком - (минус), который записывают слитно с соответствующим поисковым термином.
Пример: галогенид -хлорид.
- Фраза обозначается *двойными кавычками*.
- **Шаблоны нельзя** использовать для обозначения варьируемой части слова.
- Программа в целом не работает в режиме *stemming* (все же учет словоформ поисковых терминов частично реализован).

Записав задание, следует нажать одну из кнопок:

- **Google Search** (или **Поиск в Google**) выводит на экран список записей, обнаруженных в базе данных
- *I'm Feeling Lucky* (или *Мне повезет!*) вызывает одну веб-страницу — ту, которая в списке находилась бы на первом месте (т.е. имеющую наивысшую релевантность).

Результаты поиска

Список обнаруженных записей ранжирован по степени соответствия поисковому заданию.

Запись обычно содержит название страницы, ее адрес и — очень полезно ! — тот **фрагмент**, в котором упоминается поисковый термин.

Гиперсвязь *Cached* (*Сохранено в кэше*) ведет к копии страницы, сохраненной роботом-пауком во время индексирования сайта.

Упражнения

Проведите поиск информации, используя в качестве запроса: один термин; набор из нескольких терминов; текст на русском языке; текст на белорусском языке; текст на английском языке (например, *chemistry*); текст на французском языке (например, *chimie*).

Изучите структуру страницы с результатами поиска. Каковы различия в списках, полученных на англоязычном и на русско(белорусско)язычном сайтах?

Изучите структуру усложненного поискового бланка.

Изучите языковые инструменты системы (ссылка на Главной странице).

1.3. Яндекс

На портале [Яндекс \(http://www.yandex.ru/\)](http://www.yandex.ru/) размещается одна из самых мощных поисковых систем русскоязычного Интернета.

Поисковая программа Яндекс извлекает информацию из нескольких независимых источников, в том числе, из базы данных, сформированной роботом-пауком, и из тематического каталога.

На Главной странице портала находится простейший поисковый бланк для работы в режиме *Search*.

С Главной страницы портала можно перейти к многоуровневому тематическому каталогу (по ссылкам *еще* → *Каталог*) и работать с ним в режиме *Browse*.

Основные правила формулирования поискового задания:

- В простейшем случае поисковое задание может состоять из нескольких терминов, разделенных пробелами. Оператор по умолчанию: *AND*.
- Строка символов обозначается *двойными кавычками*.
- Для обозначения оператора *NOT* перед термином следует поставить знак - (минус).
- Программа *работает в режиме stemming*; для его отключения перед термином следует поставить оператор точной формы — восклицательный знак.

Пример: !фосфатного

- В общем случае, программа *не чувствительна к регистру букв*. Регистр учитывается только для терминов, перед которыми стоит оператор точной формы (восклицательный знак).

Если термин с восклицательным знаком начинается с заглавной буквы, ведется поиск только тех документов, в которых это слово тоже начинается с заглавной буквы.

Если термин с восклицательным знаком начинается со строчной буквы, ведется поиск документов, где это слово начинается либо со строчной, либо с заглавной буквы.

Пример:

по запросу **!Раков** будут извлечены веб-страницы со словосочетанием *посетите Раков*, но не со словосочетанием *ведро раков*.

Упражнения

Проведите информационный поиск, используя простейший и усложненный поисковый бланки.

Изучите строение тематического каталога.

Изучите раздел "Словари".

Ознакомьтесь с разделом "Новости" и другими информационными ресурсами портала.

1.4. Другие универсальные поисковые средства

[Yahoo! \(http://www.yahoo.com/\)](http://www.yahoo.com/) — один из наиболее известных и популярных порталов, предлагающих широкий спектр услуг.

На Главной странице портала [Yahoo!](http://www.yahoo.com/) находится бланк для формулирования поискового задания, а также множество ссылок на разнообразные информационные источники.

Ссылку на тематический каталог **Yahoo! Directory** нелегко найти на Главной странице портала, и каталог проще вызвать по адресу <http://dir.yahoo.com/>.

Русскоязычный поисковый бланк поисковой системы *Yahoo!* находится по адресу <http://ru.yahoo.com/>.

Упражнения

Проведите информационный поиск в *Yahoo!*

Ознакомьтесь со структурой тематического каталога.

[Bing \(http://www.bing.com/\)](http://www.bing.com/) — поисковая система, созданная компанией Microsoft.

Упражнение

Проведите информационный поиск в *Bing*.

Поисковые средства Беларуси

[TUT.BY \(http://www.tut.by/\)](http://www.tut.by/)

Портал, на котором расположены, в частности, собственный тематический каталог, а также бланк поисковой программы.

[Open.By \(http://poisk.open.by/catalog.html\)](http://poisk.open.by/catalog.html)

Тематический каталог и поисковая система.

[Акавіта \(http://catalog.akavita.by/\)](http://catalog.akavita.by/)

Тематический каталог и система рейтинга сайтов.

Упражнение

Изучите функциональные особенности белорусских поисковых средств.

Весьма подробный (но не оперативно обновляемый) **список национальных и специализированных** поисковых средств:

[Search Engine Colossus \(http://www.searchenginecolossus.com/\)](http://www.searchenginecolossus.com/)

Контрольные задания

1. Используя поисковую систему Google, расшифруйте аббревиатуру в тексте научной статьи (индивидуальное задание будет выдано преподавателем).

2. Как следует составить запрос, чтобы поиск проводился не по всему вебу, а только по заданному сайту?

3. Найдите достоверную информацию о том, кому была присуждена последняя по времени Нобелевская премия по химии.

4. Найдите первоисточник информации, использованной журналистом для написания научно-популярного сообщения (индивидуальное задание будет выдано преподавателем).

1.5. Тематические каталоги и метасайты

В настоящее время в Интернете накопилось такое количество информации, что задача *"найти что-нибудь по заданной теме"* почти всегда поддается решению. Сложнее *"найти что-нибудь толковое"*, а еще сложнее — *"найти только толковое"*.

Выбрать правильный алгоритм поиска, проанализировать список результатов, оценить релевантность и достоверность отобранных документов — все эти стадии требуют и затрат времени, и определенного уровня знаний.

В онлайн-работе полезно полагаться на мнение профессионалов, которые ранее уже прошли такой же поисковый путь, проделали черновую работу и составили для коллег советы, руководства, списки полезных страниц и сайтов. Хорошее подспорье для пользователя — это указатели, информирующие о веб-ресурсах, достойных внимания.

[Open Directory \(http://dmoz.org/\)](http://dmoz.org/) — это самый большой в мире универсальный тематический каталог. Авторами и рецензентами каталога являются энтузиасты — знатоки Интернета.

Упражнение

Ознакомьтесь со структурой каталога *Open Directory*.

Найдите и исследуйте структуру категории *"Chemistry"*.

Найдите и исследуйте структуру категории *"Top: World: Russian: Наука"*.

Примеры общехимических тематических каталогов

[Links for Chemists — Chemistry section of the WWW Virtual Library \(http://www.liv.ac.uk/Chemistry/Links/links.html\)](http://www.liv.ac.uk/Chemistry/Links/links.html)

Тематический каталог.

[ChemDex \(http://www.chemdex.org/\)](http://www.chemdex.org/)

Тематический каталог с поисковой программой.

[The Virtual Chemistry Center](http://www.martindalecenter.com/GradChemistry.html)

[\(http://www.martindalecenter.com/GradChemistry.html\)](http://www.martindalecenter.com/GradChemistry.html)

Тематический каталог. Содержит большое количество информации, постоянно обновляется, однако имеет далеко не самую удобную структуру.

[Chemistry: A Guide to Web Resources](http://libguides.library.albany.edu/chem_web_guide)

[\(http://libguides.library.albany.edu/chem_web_guide\)](http://libguides.library.albany.edu/chem_web_guide)

Компактный тематический каталог, содержащий ссылки только на самые важные информационные ресурсы.

Метасайт (*сборник ссылок на веб-страницы*) является удобным исходным пунктом при ознакомлении с ресурсами Интернета. Материал для него обычно подбирается профессиональным химиком, что в известной степени гарантирует качество проиндексированных документов. Поскольку метасайт — это продукт неавтоматизированного труда, то он относительно невелик по объему.

Структура метасайта может варьироваться от простейшего одностраничного списка адресов до многостраничного массива, упорядоченного по тематике.

Четкая граница между понятиями "метасайт" и "каталог" отсутствует.

Примеры специализированных метасайтов:

[Data and Property Calculation Websites](#)

http://tiger.uic.edu/~mansoori/Thermodynamic.Data.and.Property_html

Термодинамические ресурсы WWW

[Organic Chemistry Resources Worldwide](#)

<http://www.organicworldwide.net/>

Информационный центр по органической химии

[Каталог аналитических ресурсов \(http://www.anchem.ru/catalog/\)](http://www.anchem.ru/catalog/)

Русскоязычные ресурсы по аналитической химии

[The Lipid Library \(http://www.lipidlibrary.co.uk/\)](http://www.lipidlibrary.co.uk/)

Все о липидах

Упражнение

Из приведенных выше списков выберите несколько каталогов и метасайтов и посмотрите, какого рода информацию можно получить на них.

1.6. Словари, энциклопедии

Словари акронимов

(Акроним — слово, образованное из первых букв нескольких слов, например: БГУ, WWW, DDT).

[Acronyma \(http://www.acronyma.com/\)](http://www.acronyma.com/)

База данных содержит около полумиллиона акронимов из различных предметных областей.

[Akronymsuche \(http://www.chemie.de/tools/acronym.php3?language=e\)](http://www.chemie.de/tools/acronym.php3?language=e) на странице **Tools** портала **Chemie.DE** — пример специализированного словаря акронимов.

База данных содержит более 12 тыс. акронимов — названий химических веществ.

Несмотря на то, что язык интерфейса — немецкий, этот инструмент можно использовать для расшифровки англоязычных терминов.

Упражнение

Расшифруйте аббревиатуру **PVC**.

Энциклопедии

Wikipedia (http://en.wikipedia.org/wiki/Main_Page) — энциклопедия, авторами, рецензентами и редакторами которой могут быть любые пользователи Интернета.

Как правило, статьи в *Wikipedia* весьма информативны и содержат ссылки на авторитетные первоисточники. С другой стороны, многочисленность авторов означает и многочисленность мнений; случается, что некомпетентные, но очень активные авторы помещают в энциклопедию ложные сведения. Характерный пример: основателю Википедии пришлось чуть ли не десяток раз исправлять статью о себе самом — из-за таких "знатоков".

Суммируя плюсы и минусы Википедии, сформулируем наше отношение к ней таким образом: она хороша для предварительного ознакомления с предметом, но не для научной работы, а содержащиеся в Википедии сведения перепроверять следует обязательно.

Медленно, но верно растут Википедии и на иных языках:

Вікіпедыя (http://be.wikipedia.org/wiki/Main_Page) и

Вікіпедыя (<http://be-x-old.wikipedia.org/wiki/>) — белорусские;

Википедия (http://ru.wikipedia.org/wiki/Main_Page) — русская.

Упражнение

Сравните статьи **Hydrogen peroxide** и **Пероксид водорода** в *Wikipedia* и *Википедии*. Обратите внимание на различия в объеме и качестве представленной информации.

Answers.com (<http://www.answers.com/>) — универсальная справочная система, содержащая информацию из сотни энциклопедий и справочников.

Упражнение

В *Answers.com* найдите, что такое аспартам (*aspartame*). Рассмотрите структуру страницы результатов поиска.

Контрольное задание

По заданной аббревиатуре найдите химическую формулу вещества (индивидуальное задание будет выдано преподавателем).

1.7. Онлайн-переводчики текстов и веб-страниц

С Главной страницы поисковой системы *Google* по ссылке, размещенной справа вверху, можно перейти к странице инструментов, предназначенных для перевода текста с одного языка на другой:

Google.com: [Google Language Tools](http://www.google.com/language_tools?hl=en)
(http://www.google.com/language_tools?hl=en)

Google.ru: [Google Языковые инструменты](http://www.google.ru/language_tools?hl=ru)
(http://www.google.ru/language_tools?hl=ru)

Google.by: [Google Языковые инструменты](http://www.google.by/language_tools?hl=ru)
(http://www.google.by/language_tools?hl=ru)

Переводчик от Google имеется и на портале [TUT.BY \(http://slovari.tut.by/\)](http://slovari.tut.by/).
Бланк на закладке *Словари* служит для перевода отдельных слов.
Бланк на закладке *Переводчик* предназначен для перевода больших фрагментов текста.
Этот инструмент легко доступен и поэтому удобен.

Примеры онлайн-переводчиков:

[Yahoo! Babel Fish \(http://babelfish.yahoo.com/\)](http://babelfish.yahoo.com/)
[ПРОМТ \(http://translate.promt.by/\)](http://translate.promt.by/)
[SDL FreeTranslation.com \(http://translate.promt.by/\)](http://translate.promt.by/)

Упражнение 1

Изучите структуру каждого из указанных инструментов. Особое внимание обратите на то, с каких и на какие языки можно переводить на каждом из этих сайтов.

Упражнение 2

Оцените качество перевода с английского языка на русский, выполненного указанными выше переводчиками.

В качестве рабочего примера возьмите следующую фразу :)

Our cat gave birth to three kittens - two whites and one black

(Правильный перевод: "Наша кошка родила трех котят — двух белых и одного черного").

Пример позаимствован на странице "Художественный или технический: нелепости перевода" <http://www.techtranslator.com/index.php?object=nelepost>

Контрольные задания

Выполните контрольные задания, выданные преподавателем (перевод веб-страниц; информационный поиск на сайтах, язык которых незнаком пользователю).

1.8. Анализ сайта — самостоятельная работа

1. О чем говорит доменное имя?

Лишь один из следующих сайтов является официальным сайтом президента США.

Определите, какой именно.

По каким признакам можно решить такую задачу?

(Во всех адресах мы видим домен whitehouse, т. е. Белый дом.)

<http://whitehouse.com/>
<http://www.whitehouse.net/>
<http://www.whitehouse.gov/>
<http://www.whitehouse.ie/>
<http://whitehouse.lk/>

2. Очень специфический сайт

Проанализируйте материал указанного ниже сайта.

Как можно охарактеризовать степень достоверности информации, приведенной на нем?

Как бы Вы классифицировали сайты с подобным содержанием?

(Кстати, о каком веществе здесь идет речь?)

<http://www.dhmo.org/>

3. Где в Рунете должно быть самое лучшее? Конечно же на сайте AllBest.ru?

Проанализируйте, насколько содержимое сайта соответствует названию.

Сначала окиньте взглядом типы материалов, затем в разделе *"База знаний"* в рубрике *"Химия"* прочитайте один из документов.

Мне особенно приглянулась курсовая работа «рН в живых организмах», очень рекомендую :)

<http://www.allbest.ru/>

4. Периодическая система

Интерактивная периодическая система — весьма распространенный элемент химического веба.

Ниже приведен адрес сайта, на Главной странице которого расположен справочник такого рода (кстати, внешне весьма привлекательный и удобный).

Засеките время, которое Вам понадобится, чтобы найти в таблице типичную методическую ошибку, особенно вредную в учебной литературе (сайт-то предназначен для школьников).

Есть там еще и смешные химглупости, связанные с автоматизацией содержимого. Найдете?

<http://midakva.ucoz.ru/>

5. Где можно обнаружить самые интересные новые сайты? Вероятно, в числе номинантов конкурсов?

Пройдемся по весьма солидному [Конкурсу Золотой сайт](http://www.goldensite.ru/) (<http://www.goldensite.ru/>), а именно по его разделу [Образовательные программы, курсы](http://www.goldensite.ru/list.aspx?id=587) (<http://www.goldensite.ru/list.aspx?id=587>).

Зайдите на сайты номинантов-химиков.

Кому бы из них Вы присудили приз?

Свою оценку любого из сайтов-номинантов изложите в письменной форме.

2. НАУЧНЫЕ ЖУРНАЛЫ

2.1. Введение

В основе современной системы хранения, распространения и оценки качества научной работы лежит **научный журнал** — рецензируемое периодическое издание, в котором результаты научных исследований представлены в форме **научных статей**.

Планируя поиск научной информации в Сети, следует учитывать **три важных обстоятельства**:

а) Не стоит надеяться, что из WWW можно бесплатно извлечь все статьи, которые были опубликованы в печатных научных журналах. Будь это так — кто бы тогда тратил деньги на подписку?!

б) Но, с другой стороны, не стоит пренебрегать ресурсами WWW — они огромны, здесь есть такие статьи, журналы, книги, которые не попадали (а может, и не попадут) в наши самые фундаментальные библиотеки.

в) Степень достоверности онлайн-материалов весьма разнообразна, поэтому не каждому источнику информации стоит доверять безоговорочно. Предпочтение следует отдавать официальным сайтам издательств, научных обществ, университетов.

В настоящее время практически вся современная **печатная научная периодика** в той или иной степени представлена в WWW. Минимум информации, на который может рассчитывать читатель, — это *оглавления журналов (ToC — Tables of Contents)*. Нормой становится размещение в Сети *рефератов* статей. К сожалению, невелика доля тех журналов, которые открывают свободный доступ к самим *статьям*. (Вышесказанное касается только бесплатной информации — для подписчиков в WWW есть (почти) всё.)

Целостный комплект научного журнала может находиться на сайте издательства, на своем собственном сайте, на сайте агрегатора.

2.2. Сайты издательств

Наиболее вероятным местом, где содержится самая полная информация о журнале, является сайт издательства или научного общества, публикующего это издание.

Ниже приведены ссылки на сайты нескольких особенно важных для нас издательств.

Самостоятельная работа

Ознакомьтесь со структурой **каждого** сайта.

Найдите гиперсвязи, ведущие к Главным страницам журналов.

С Главной страницы одного из журналов перейдите к списку томов, выпусков, к оглавлению выпуска, к реферату одной из статей и, если это возможно, к полному тексту статьи.

Изучите структуру научной статьи. Каковы особенности *HTML*-версий статей на сайтах разных издательств?

Какую служебную информацию издательства размещают на сайтах журналов?

Какие вспомогательные инструменты доступны на сайтах издательств?

Найдите поисковые бланки, предназначенные для проведения информационного поиска по опубликованному в журнале (журналах) материалу. Проведите поиск (любой, например, по термину *black*) и ознакомьтесь со структурой списка результатов поиска.

Обратите внимание на функциональные возможности простейших и усложненных поисковых бланков.

[Национальная Академия Наук Беларуси](http://www.nasb.gov.by/rus/publications/index.php)
(<http://www.nasb.gov.by/rus/publications/index.php>)

Издания академии. Доступны оглавления журналов, рефераты статей с 1996 г.

[МАИК "Наука/Интерпериодика"](http://www.maikonline.com/maik/showCatalogs.do?type=alphabet&lang=ru)
(<http://www.maikonline.com/maik/showCatalogs.do?type=alphabet&lang=ru>)

На сайте Международной академической издательской компании размещена информация о статьях, опубликованных в последние годы в российских академических журналах.

Бесплатны оглавления, рефераты, списки процитированной литературы.

[American Chemical Society \(ACS\)](http://pubs.acs.org/action/showPublications?display=journals) (<http://pubs.acs.org/action/showPublications?display=journals>)

Бесплатны оглавления, рефераты, приложения к статьям (*Supporting info*), фрагменты иной информации. Для зарегистрированных пользователей доступны рукописи, принятые к публикации (*Just Accepted Manuscripts*).

Полностью открыты *Sample Issues* — по одному выпуску (как правило, № 1 текущего года) каждого журнала издательства.

Гипертекстовая структура статей в журналах American Chemical Society

Гипертекстовую структуру статей мы рассмотрим на примере публикации в журнале [ACS Applied Materials & Interfaces](http://pubs.acs.org/journal/aamick) (<http://pubs.acs.org/journal/aamick>).

Ссылка, ведущая к *Sample Issue*, находится в разделе [Browse the Journal](#) навигационного меню Главной страницы журнала.



В оглавлениях журнала рядом с библиографическим описанием статьи находятся навигационные ссылки на информационные блоки:

Abstract — реферат;

Supporting Info (*необязательный элемент*) — бесплатный дополнительный материал, доступный на сайте и отсутствующий в печатном журнале;

Full Text HTML — статья в формате HTML;

Hi-Res PDF — статья в формате PDF (без встроенных гиперсвязей);

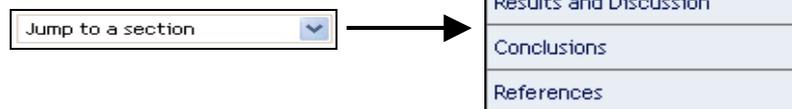
PDF w/ Links — статья в формате PDF (с гиперсвязями).



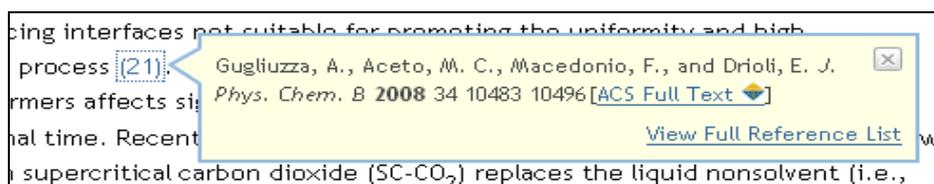
Пометка **Sponsored Access** указывает, что данная статья **бесплатна**.

HTML-версия статьи, как и ее печатный вариант, состоит из стандартных *разделов (sections)*.

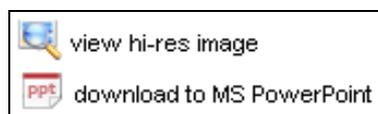
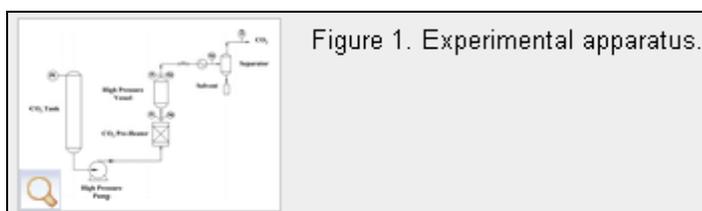
Для быстрого перехода к соответствующему разделу предназначены меню [Jump to a section](#), размещенные на границах разделов.



Ссылки на литературные источники обозначены в тексте статьи стандартным образом — числами. При щелчке по числу-ссылке появляется всплывающее окошко с библиографическим описанием первоисточника:



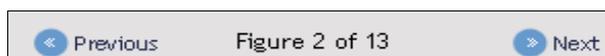
В текст статьи встроены миниатюрные копии *иллюстраций*. При щелчке по миниатюре открывается новое окно, содержащее среднemasштабную иллюстрацию и элементы управления.



Управляющие кнопки окна иллюстрации:

[view hi-res image](#) — вызов изображения максимального разрешения;

[download to MS PowerPoint](#) — сохранение иллюстрации с подписью на компьютере пользователя (в формате PPT).



Все иллюстрации данной статьи можно перелистывать, используя кнопки [Previous](#) и [Next](#).

Список использованной литературы (References) содержит гиперсвязи:

1. Gugliuzza, A., and Drioli, E. J. *Membr. Sci.* **2007** 300 51 >62 [[CrossRef](#)], [[ChemPort](#)]

- [ACS Full Text](#) — к статье, если она была опубликована в одном из журналов ACS;
- [CrossRef](#) — к статье на сайте соответствующего издательства, если это издательство является членом ассоциации *CrossRef*.

(Следует быть готовым к тому, что полный текст статьи скорее всего окажется недоступным для неподписчика и читатель сможет ознакомиться только с рефератом).

- [CAS](#) — к реферату в реферативной базе данных *ChemPort*.
- [PubMed](#) — к реферату в реферативной базе данных *PubMed*.

Рефераты одной и той же статьи, полученные с сайта издательства и из реферативной базы данных, могут различаться, если создателей реферативной базы данных не удовлетворяет качество или объем информации, содержащейся в авторском варианте реферата.

Продолжим изучение сайтов издательств.

Royal Society of Chemistry (RSC) (<http://pubs.rsc.org/en/Journals?key=Title&value=Current>)

Для зарегистрированных пользователей бесплатны полные тексты статей, опубликованных в 1996–2007 гг.

Полностью открыты оглавления, рефераты, приложения к статьям (*Electronic supplementary information*), а также частично некоторая иная информация.

ScienceDirect (<http://www.sciencedirect.com/science/journals>) — информационный центр издательства **Elsevier Science**.

Бесплатны оглавления, рефераты статей, а для отдельных журналов — и полные тексты статей.

Пояснение.

Elsevier Science размещает информацию о своих журналах на двух сайтах: www.sciencedirect.com и www.elsevier.com. Объем и тип имеющейся там информации различны. В Интернете можно встретить ссылки на журналы издательства, ведущие к обоим сайтам.

ScienceDirect (www.sciencedirect.com) — это архив публикаций; именно здесь стоит искать статьи.

Сайт www.elsevier.com предоставляет общие сведения о журнале, содержание последних номеров. Если вы все же попали сюда — ищите ссылку, перенаправляющую на *ScienceDirect* (например, *Access Full-Text* или *Online Journal*).

SpringerLink (<http://www.springerlink.com/journals/>) — информационный центр издательства **Springer**.

Бесплатны оглавления, рефераты, первые страницы статей.

Через сеть БГУ мы имеем возможность читать и полные тексты статей из многих журналов.

Wiley Online Library (<http://onlinelibrary.wiley.com/>) — информационный центр издательства **John Wiley & Sons**.

Бесплатны оглавления, рефераты статей.

Контроль

Обсуждение структуры информационных источников и устные ответы на вопросы преподавателя.

Индивидуальное контрольное задание 1

Найдите импакт-фактор журнала, указанного преподавателем, за 2010 г.

(Информацию, обнаруженную на сайте журнала, будем считать достоверной).

Индивидуальное контрольное задание 2

Определите число работ, опубликованных данным автором (данными авторами) в журналах каждого из указанных издательств (индивидуальное задание получите у преподавателя).

2.3. Списки журналов

Адреса многих журналов можно узнать, если обратиться к специализированным сайтам, содержащим списки представленных в WWW периодических изданий.

Ниже приведены примеры таких источников.

Самостоятельная работа

Ознакомьтесь со структурой и функциональными возможностями каждого метасайта (базы данных).

Исследуйте, каким образом эти инструменты можно использовать для поиска сайта заданного журнала.

Выполните проверочные задания 1–7.

На сайтах журналов найдите страницы с перечнями томов и выпусков; определите уровни доступности опубликованного материала.

Веб-страницы с результатами выполнения каждого из проверочных заданий необходимо будет предъявить преподавателю.

[Electronic Journals Library - University Library of Regensburg](http://rzblx1.uni-regensburg.de/ezeit/fl.phtml?bibid=UBS&colors=7&lang=en¬ation=V&frames=&toc=&ssg=)
(<http://rzblx1.uni-regensburg.de/ezeit/fl.phtml?bibid=UBS&colors=7&lang=en¬ation=V&frames=&toc=&ssg=>)

Списки журналов упорядочены по тематике и по алфавиту. Языки сайта: немецкий и английский.

Проверочное задание 1

Найдите онлайн-ресурсы журнала *Journal of Industrial and Engineering Chemistry*.

[Chemistry Journals - University of Cambridge](http://www-jmg.ch.cam.ac.uk/data/c2k/cj/)
(<http://www-jmg.ch.cam.ac.uk/data/c2k/cj/>)

Сайт содержит несколько вариантов списков журналов: алфавитный (*alphabetical*), упорядоченный по издательствам (*by Publisher*), упорядоченные по тематике.

Проверочное задание 2

Найдите онлайн-ресурсы журнала *Journal of Heterocyclic Chemistry*.

[Genamics JournalSeek](http://journalseek.net/) (<http://journalseek.net/>) — база данных, содержащая адреса и описания периодических изданий разной тематики. Здесь, в частности, можно найти URL по сокращенному названию журнала.

Проверочные задания 3, 4

Найдите ресурсы журнала *Polymer Journal* (это полное название).

Найдите ресурсы журнала *Polymer Polymer Compos* (это сокращенное название).

ABC Chemistry: Списки бесплатных полнотекстовых научных химических журналов

[Постоянно доступные журналы](http://www.abc.chemistry.bsu.by/current/a.htm)
(<http://www.abc.chemistry.bsu.by/current/a.htm>)

[Временно доступные журналы](#)

<http://www.abc.chemistry.bsu.by/current/trialjournal.html>

[Полнотекстовые журналы, доступные только из сети БГУ](#)

<http://www.abc.chemistry.bsu.by/current/bdu.htm>

Проверочные задания 5, 6

Найдите места хранения полных текстов химических публикаций *Вестника БГУ*.

Найдите на официальном сайте БГУ путь от Главной страницы к химическим публикациям *Вестника БГУ*.

[Directory of Open Access Journals - DOAJ](http://www.doaj.org/) (<http://www.doaj.org/>)

Каталог бесплатных журналов. Многодисциплинарный, наднациональный.

В каталог включены только те журналы, которые **полностью соответствуют принципам *Open Access*** (по этой причине здесь нет журналов коммерческих издательств, предоставляющих избирательный доступ к своим материалам).

Имеется поисковая система, которая позволяет проводить информационный поиск по статьям, опубликованным в части журналов, включенных в каталог.

Проверочное задание 7

Сколько статей опубликовано в журнале *International Journal of Inorganic Chemistry* ?

Дополнительные инструменты

Самостоятельная работа

Ознакомьтесь с функциональными возможностями справочников.

Выполните проверочные задания 8–9.

Веб-страницы с результатами выполнения проверочных заданий предъявите преподавателю.

Для **обнаружения полного названия журнала по его сокращенному названию**, кроме базы данных *Genamics JournalSeek*, можно

- воспользоваться справочником:

[Science and Engineering Journal Abbreviations](http://www.library.ubc.ca/scieng/coden.html) (<http://www.library.ubc.ca/scieng/coden.html>)

- или обратиться к поиску в ресурсах CAS:

[CAS Source Index \(CASSI\) Search Tool](http://cassi.cas.org/search.jsp) (<http://cassi.cas.org/search.jsp>)

Проверочное задание 8

Определите полное название журнала, если его сокращенное название — *Pap. Celul.*

Для **обнаружения сокращенного названия журнала по его полному названию**, кроме баз данных *Genamics JournalSeek* и *CAS Source Index*, можно воспользоваться справочником:

[crossref.org :: Title List](http://www.crossref.org/titleList/) (<http://www.crossref.org/titleList/>)

Порядок работы. В поисковом бланке переключатели поставить в положение *Genre: journals, Search: title*; в графу вписать по меньшей мере одно слово из названия журнала.

В списке результатов поиска кнопка "*зеленый треугольник*" открывает таблицу с вариантами записи названия соответствующего журнала.

Проверочное задание 9

Определите сокращенное название журнала, если его полное название — *Advanced Composite Materials*.

Индивидуальный контроль 1 (задание получите у преподавателя)

Вам выдано полное название журнала.

а) Найдите сайт журнала по его известному полному названию.

б) Определите степень доступности информации, содержащейся в последнем выпуске журнала текущего года.

(Возможные варианты: библиографическое описание, реферат, список литературы, полный текст статьи, дополнительный электронный материал).

в) Выведите на экран доступные ресурсы одной из статей последнего выпуска и предъявите преподавателю.

Индивидуальный контроль 2 (задание получите у преподавателя)

Вам выдано *краткое* библиографическое описание статьи.

Необходимо найти ее *полное* библиографическое описание — обязательно в **первоисточнике**.

(*Полное библиографическое описание* должно содержать в себе список всех авторов, название статьи, полное название журнала, год, номер тома, номер выпуска, начальную и конечную страницу публикации).

Выведите на экран веб-страницу с полным библиографическим описанием. Сформируйте полное библиографическое описание данной статьи в формате *TXT* или *DOC*.

Предъявите веб-страницу и библиографическое описание преподавателю.

Индивидуальный контроль 3

Предположим, что возникла необходимость найти статьи о коллоидах.

(Справка: *коллоид* по-английски — *colloid*).

Поиск будем вести в базе данных *DOAJ*.

Сколько документов по коллоидной тематике можно извлечь из базы данных *DOAJ*, если поиск вести по полю "Название", а в качестве поискового термина использовать имя существительное?

2.4. Сайты журналов — дополнение

Большая часть онлайн-журналов располагается на сайтах издательств и агрегаторов. Это удобно для пользователя, поскольку таким образом он получает возможность вести информационный поиск по большому информационному массиву. (К сожалению, одновременный поиск по всем информационным массивам нам недоступен, но будем надеяться, что эта проблема в будущем исчезнет).

За пределами крупных издательств и агрегаторов существует много мест, где тоже публикуются научные журналы.

Ознакомимся с характерными примерами.

Некоторые научные сообщества самостоятельно занимаются выпуском научной продукции по своей тематике. Это может быть один журнал или несколько журналов, а располагаться они могут либо на сайте общества, либо на собственном сайте.

Пример небольшого тематически однородного информационного массива первичной литературы — *Цифровая библиотека* международного *Электрохимического общества* (*The Electrochemical Society*):

[ECS Digital Library \(http://ecsdli.org/\)](http://ecsdli.org/)

Здесь размещено несколько (*сколько?*) журналов общества.

Определите *уровень доступности материала в последних выпусках этих журналов*.

Отделение химического образования Американского химического общества издает авторитетный, по сути международный, журнал **Journal of Chemical Education**.

У этого журнала есть два сайта:

[Новый сайт \(http://pubs.acs.org/journal/jceda8\)](http://pubs.acs.org/journal/jceda8) является частью сайта издательства ACS. Здесь размещены опубликованные статьи.

[На старом сайте \(http://jchemed.chem.wisc.edu/\)](http://jchemed.chem.wisc.edu/) хранятся иные материалы, имеющие отношение к химическому образованию (*какие?*).

Многие университеты издают журналы для оперативной публикации результатов научных исследований своих сотрудников.

Нередко у такого журнала название начинается с "Вестник...", "Известия...", "Труды...", "Ученые записки...", а сайт журнала является составной частью университетского сайта (и может быть спрятан в разделах "Наука", "Публикации", "Библиотека", "Структура" и др. — в этом отношении фантазия у сайтостроителей богатая).

Упражнение

Определите, издает ли **[Воронежский государственный университет \(http://www.vsu.ru/\)](http://www.vsu.ru/)** научные журналы химической тематики.

В своем большинстве онлайн-версии журналов являются электронными версиями печатных изданий. Тем не менее, в вебе есть и электронные журналы, не имеющие печатных аналогов.

Пример — расположенный на собственном сайте журнал Европейской полимерной ассоциации (EPF):

[e-Polymers \(http://www.e-polymers.org/\)](http://www.e-polymers.org/)

Упражнение

Определите, как можно получить список статей, опубликованных в журнале *e-Polymers* в 2009 году.

Определите, сколько выпусков журнала было опубликовано в 2009 году.

По каким параметрам библиографическое описание статьи, опубликованной в *e-Polymers*, будет отличаться от стандартного библиографического описания?

2.5.1. Агрегаторы

Агрегатор — это информационный центр, в котором собраны комплекты журналов разных издательств.

Рассмотрим примеры агрегаторов нескольких типов.

а) Коммерческий агрегатор

Агрегаторы такого типа создаются фирмами-посредниками, занимающимися продажей научной продукции.

Одни информационные провайдеры работают только с подписчиками, как правило, с крупными клиентами — организациями. Пример — *EBSCOhost*, с которым БГУ имеет соответствующее соглашение. С ресурсами *EBSCOhost* мы позже ознакомимся детально.

Другие провайдеры в большей степени ориентируются на розничную торговлю. Для привлечения потенциальных клиентов они предоставляют бесплатный доступ к библиографической и реферативной информации, а деньги зарабатывают на продаже полных текстов статей. Пример — *ingentaconnect*, о котором речь ниже.

[ingentaconnect \(http://www.ingentaconnect.com/content\)](http://www.ingentaconnect.com/content)

Многодисциплинарный агрегатор англоязычных журналов коммерческих издательств и научных обществ.

Некоторые малые издательства не размещают свою периодику на собственных сайтах, а распространяют ее через *ingentaconnect*. Материалы иных издательств, наоборот, лишь частично представлены в *ingentaconnect*, а в большем объеме — на собственных сайтах издательств.

Агрегатор ценен большим объемом и разнообразием библиографической информации, поэтому мы еще не раз будем возвращаться к нему.

Упражнения и проверочные задания

а) Определите, сколько записей имеется в базе данных *ingentaconnect*.

б) Проанализируйте, насколько разнообразен набор издательств, ресурсы которых включены в базу данных. Есть ли здесь ресурсы ведущих издательств химической литературы (таких как *Elsevier*, *American Chemical Society* и т.п.)?

в) Проверьте, есть ли в базе данных статьи о графене (*graphene*). Основываясь на полученных результатах, выскажите предположение, для решения какого типа проблем ресурсы *ingentaconnect* стоило бы применять в научной работе.

г) Проверьте, какова здесь степень доступности материала.

б) Библиотека (архив) научной периодики страны или региона

В целях сохранения научной информации, в целях общемировой рекламы достижений страны, в целях помощи своим гражданам некоторые государства интенсивно формируют электронные архивы отечественной научной периодики. Как правило, коммерческая выгода в таких случаях вторична; престиж страны — первичен. (В российской *eLIBRARY.RU*, с которой позже будем работать детально, пытаются совместить оба фактора). *Примеры:*

[J-STAGE \(http://www.jstage.jst.go.jp/browse/_journalist\)](http://www.jstage.jst.go.jp/browse/_journalist)

Многодисциплинарный. Национальный.

Содержит практически все современные **научные журналы Японии** (публикации на английском и японском языках).

К очень многим имеющимся здесь журналам предоставляется бесплатный доступ.

Контрольный вопрос 1

Установите недостающее: в *WWW* имеются публикации журнала *Journal of Chemical Engineering of Japan* за (**какие?**) годы.

Контрольный вопрос 2

Установите имеющееся в базе данных агрегатора *J-STAGE* число публикаций 2008 года, в которых содержатся результаты исследования коллоидных систем (англ.: *colloid, colloids, colloidal, ...*).

SciELO (<http://www.scielo.org/php/index.php?lang=en>)

Многодисциплинарный. Наднациональный.

Содержит **научные журналы стран Латинской Америки** (публикации на испанском, португальском и английском языках).

К журналам предоставляется бесплатный доступ.

Контрольный вопрос 3

Для какой категории читателей предназначен журнал *Revista Virtual de Quimica*?

в) Тематический архив научных изданий

Национальная медицинская библиотека США при государственной финансовой и юридической поддержке формирует уникальный архив:

PubMed Central (<http://www.ncbi.nlm.nih.gov/pmc/journals/>)

Тематически **специализированный**: журналы биомедицинской направленности (тематика статей включает биохимию, органическую химию, да и не только органическую). Архив содержит американские и международные издания на английском языке.

Бесплатный полный доступ к статьям предоставляется либо сразу, либо через небольшой период эмбарго после опубликования.

В разделе *Special Collections* накапливаются избранные статьи из платных журналов крупных международных издательств (*ACS, Elsevier, Wiley-Blackwell* и др.) — те научные работы, которые публикуются по схеме *Open Access*.

Упражнение

Проблема токсичности/безопасности наночастиц стала особенно актуальной в настоящее время (*кстати, почему?*).

Проверьте, в какой степени ресурсы *PubMed Central* пригодны для сбора первичной информации по указанной проблеме.

Термины: *наночастица = nanoparticle, токсичный = toxic, токсичность = toxicity*.

О синтаксисе запроса: шаблон * разрешен; оператор следует записывать в явной форме.

О поисковых бланках: далеко не лучшие в вебе; к более-менее удобному перейдите по ссылкам *Search* → *Limits*.

Контрольное задание

В дискуссии о токсичности/безопасности наночастиц профессор А., ссылаясь на литературные данные, упомянул, что даже такой химически неактивный металл как золото (*gold*) может быть опасным для организма. Его оппоненты, собиравшиеся сорвать большой куш на наномоде (как выяснилось, на продаже нанозолотосодержащей настойки), потребовали доказательств. Проф. А. оказался в некотором затруднении: он помнил только, что статья была опубликована в 2010 году и что суть проблемы звучала в названии статьи.

Найдите в архиве *PubMed Central* статью, подтверждающую или опровергающую заявления профессора А.

2.5.2. eLIBRARY.RU

eLIBRARY.RU (<http://elibrary.ru/defaultx.asp>) — междисциплинарный информационный центр Российской Федерации. Содержит англо- и русскоязычные научные издания.

- В полном объеме ресурсы *eLIBRARY.RU* открыты только подписчикам.
- Библиографическая и реферативная информация доступна любому пользователю.
- Зарегистрированные пользователи (регистрация бесплатна) могут, кроме того, получать полные тексты статей из нескольких сотен российских журналов.

Информационный центр имеет очень удобный интерфейс, интуитивно понятный и дружелюбный по отношению к пользователю. С другой стороны, поскольку центр находится в стадии развития и отладки, при работе с ним приходится сталкиваться с заметным числом недоработок.

Индивидуальная работа 1: изучение структуры сайта

Как можно извлечь список журналов заданной тематики заданного издательства?

Извлеките один из выпусков журнала, материалы одной из статей.

Какие гиперсвязи присутствуют в типичной записи?

По каким признакам можно судить об объеме материала, доступного бесплатно?

Как можно извлечь список работ заданного автора?

Определите, каким образом можно узнать показатели цитирования заданного автора.

Существенное примечание. Количественные показатели цитирования, которые сообщаются в *eLIBRARY.RU*, характеризуют состояние дел только в этой части российского информационного пространства. Если бы мы провели аналогичный анализ в международных базах данных *ISI Web of Science* или *Scopus*, мы бы убедились в несравненно большей цитируемости работ этого же автора.

Если судить по инструкции, имеющейся на сайте, **правила формулирования запроса** в *eLIBRARY.RU* в основном стандартны.

Индивидуальная работа 2: тематический поиск и исследование функциональных возможностей поисковой программы

Найдите на сайте правила формулирования текста тематического запроса.

Проверьте, насколько адекватно поисковая программа реагирует на тексты запросов, сформированные по этим правилам.

(Следует проверить, в частности, работу логических операторов и оператора расстояния, поиск по словоформам, использование шаблона).

Для проведения эксперимента проводите пробные поиски по теме *токсичность наночастиц* (используйте русскоязычные термины).

Выберите такую область поиска, которая позволит оперативно оценивать содержание и качество извлекаемого списка результатов поиска. Сначала проведите поиск по одному термину, затем по двум; проанализируйте списки результатов поиска на соответствие поисковому заданию.

Какие "правила оформления поисковых запросов" исполняются верно?

Есть ли правила, которые исполняются неверно?

Ответ оформите в письменном виде и сдайте преподавателю.

2.5.3. EBSCOhost

EBSCOhost (<http://search.ebscohost.com/>) — многодисциплинарный агрегатор. Содержит в основном англоязычные научные, технические и общественно-политические издания.

EBSCOhost работает только с подписчиками, но **через сеть БГУ нам открыт полный доступ** ко всем материалам агрегатора.

EBSCOhost состоит из нескольких баз данных, в том числе:

- **Academic Search Complete** — научная и научно-популярная литература разнообразной тематики.
Academic Search Complete содержит ок. 12 тыс. изданий, из них ок. 8 тыс. — с полными текстами статей, остальные представлены библиографическими описаниями и рефератами.
- **Medline** — реферативная база данных, содержащая литературу, прямым или косвенным образом связанную с науками о живом (*Life Sciences*).
- **ERIC** — база данных, в которой накапливаются образовательные ресурсы, как реферативные, так и полнотекстовые.

Плюсы:

(1) Агрегатор *EBSCOhost* ценен тем, что здесь можно проводить одновременный информационный поиск по ресурсам, собранным от разных издательств.

(2) *EBSCOhost* — единственное для нас место, где можно найти полные тексты статей, опубликованных в малотиражных или узкоспециализированных журналах.

Следует однако учитывать и ограничения:

(1) база данных *EBSCOhost* не охватывает все издательства, а в пределах издательства не охватывает все его ресурсы;

(2) *EBSCOhost* не предоставляет доступ к полным текстам статей, публикуемых ведущими коммерческими издательствами.

Правила формулирования запроса в EBSCOhost в основном стандартны:

- разрешены булевы операторы *AND, OR, NOT*;
- порядок выполнения операций назначается круглыми скобками;

- фраза обозначается двойными кавычками;
- шаблон * заменяет произвольное число символов (используется только в конце слова);
- шаблон ? заменяет один символ (используется в середине и в конце слова);
- разрешены операторы расстояния *near* и *with*, которые отображаются буквами *n* или *w* с соответствующим числом. Примеры:
 - по запросу **andrey n3 rogach** извлекаются документы, содержащие фрагменты **Andrey L. Rogach, Andrey Rogach, Rogach Andrey** (любой порядок слов);
 - по запросу **andrey w3 rogach** извлекаются документы, содержащие фрагменты **Andrey L. Rogach, Andrey Rogach** (заданный порядок слов);
- запрос для поиска по полям (*название документа, авторы, предметная область* и т.д.) стоит формировать на бланке расширенного поиска.

Самостоятельная работа 1

Информационный поиск в базе данных *Academic Search Complete*.

Изучите структуру бланка Расширенного поиска. Проведите пробный тематический поиск, поиск по автору. Изучите возможности коррекции списка результатов поиска (по тематике, по доступности первоисточника, по типу публикации).

Как можно извлечь список журналов, имеющих в базе данных?

Какие вспомогательные указатели имеются в распоряжении пользователя? Определите, как следует переносить сведения из вспомогательного указателя в поисковый бланк.

Контрольные задания 1

Выполните информационный поиск (индивидуальное задание получите у преподавателя).

Самостоятельная работа 2

Проведите информационный поиск в базе данных *Medline*.

Проведите информационный поиск в базе данных *ERIC*.

Проведите информационный поиск одновременно во всех базах данных агрегатора.

Контрольное задание 2

Охарактеризуйте базу данных *Eric* как источник библиографической, реферативной и полнотекстовой информации по теме "Демонстрационный эксперимент на уроках химии".

2.5.4. ProQuest

ProQuest (<http://proquest.umi.com/login>) — это междисциплинарный информационный центр, состоящий из множества текстовых баз данных. Тематика их широка: от коммерции до науки. Типы первоисточников разнообразны: газеты, бизнес-отчеты, диссертации, научные журналы и др.

Библиотека БГУ подписана на ресурсы *ProQuest*, и это дает нам возможность работы с ними — но **только из сети БГУ**.

Информация из естественнонаучных журналов сконцентрирована главным образом в базах данных *ProQuest Science Journals*, *ProQuest Education Journals*, *ProQuest Biology Journals*, *Research Library*.

Кроме того, для химиков весьма полезен диссертационный архив *Dissertations & Theses* — с ним будем работать позже.

Ссылка [ProQuest \(http://proquest.umi.com/login\)](http://proquest.umi.com/login) ведет к усложненному поисковому бланку (авторизация в сети БГУ протекает автоматически).

Обратим внимание, что на странице, содержащей поисковый бланк, имеются следующие полезные элементы:

- Меню *Interface language* (справа вверху);
- Закладка *Topics (Темы)*;
- Закладка *Publications* (Перечень названий журналов, газет и т. п.);
- Кнопка *More Search Options (Расширить выбор)* (под бланком).

More Search Options (Расширить выбор) выводит на экран вторую — нижнюю — половину поискового бланка.

Нижняя половина бланка содержит существенное дополнение — ссылки *Look up... (Найти...)*, ведущие к алфавитным указателям.

Пользователь может выбирать термины из алфавитного указателя и вносить их в поисковый бланк.

Некоторые пояснения терминологии, встречающейся на бланках:

- *Citation* означает сумму полей *Название документа + Авторы + Место публикации + Тематическая область (Subject)*.
- В русском интерфейсе *Краткий обзор*, как ни странно, означает *Реферат, Abstract*.
- В русском интерфейсе *Название документа* означает "название статьи", а *Название публикации* означает "название журнала (или газеты, или иного издания)".
- Поле *Author* содержит имена и фамилии авторов документа, а поле *Person* содержит имена и фамилии тех людей, о которых рассказывается в статье.
- Поиск по полям *Citation and abstract* означает поиск по библиографическому описанию и реферату.
- Поиск по полям *Citation and document text* означает поиск по библиографическому описанию, полному тексту документа и реферату.
- При поиске по полю *Document text* не ведется поиск по реферату.

Основные правила формулирования запроса:

- Разрешены логические операторы *AND, OR, AND NOT*.
- Точная фраза обозначается кавычками.
- Если запрос состоит из двух слов без оператора, такой текст по умолчанию считается точной фразой.
- Если запрос состоит из трех слов без операторов, ведется поиск документов, в которых эти слова должны находиться по соседству.
- Поиск по словоформам терминов (*stemming*) не ведется.
- Шаблон * (*звездочка*) может заменять собой произвольное число символов в конце слова.
- Шаблон ? (*вопросительный знак*) может заменять собой один символ в середине или в конце слова.

- Разрешены операторы расстояния:
W/n указывает, что в документе термины должны быть на расстоянии не более чем в *n* словомест, причем в любом порядке,
PRE/n указывает, что в документе термины должны быть на расстоянии не более чем в *n* словомест и именно в таком же порядке,
W/PARA указывает, что в документе термины должны быть в одном абзаце на расстоянии не более чем в 1000 символов.

Упражнение

Исследуйте поисковые бланки, работу поисковой программы, структуру списка результатов поиска, структуру отдельной записи.

Сконструируйте задачи с операторами расстояния и исследуйте результаты поиска.

Проверочное задание

Исходное состояние: поиск по всем базам данных.

Сколько изданий содержат в своих названиях слова, начинающиеся с фрагмента *chem* ?

Сколько документов опубликовано в этих изданиях?

Из числа этих документов: сколько опубликовано в научных журналах?

Из последнего списка: число доступных полнотекстовых документов = ...

В том числе, опубликовано в июне 2011 года: ...

Из них: находятся в базе данных *ProQuest Science Journals* ...

Из последнего списка: рассказывающих об ученом по фамилии *Shelley* ...

Контрольное задание

Получите у преподавателя текст заметки, опубликованной в научно-популярном издании.

а) В базе данных *ProQuest* найдите научную статью, на основании которой журналист написал выданную вам заметку.

б) В реферате статьи найдите подтверждение или опровержение численных данных, сообщенных в заметке.

в) Поскольку полный текст этой статьи отсутствует в базе данных *ProQuest*, найдите сайт соответствующего журнала и на нем — эту статью.

г) На сайте журнала проанализируйте структуру страницы реферата и определите:

- который из авторов ведет переписку с заинтересованными читателями?
- где работает этот автор?
- каков его электронный адрес?

д) Извлеките статью в формате PDF и укажите, в каких местах текста размещается перечисленная выше контактная информация.

3. НЕЖУРНАЛЬНЫЕ РЕЦЕНЗИРУЕМЫЕ ПУБЛИКАЦИИ

3.1. Введение

WWW открывает для читателя огромный информационный пласт рецензируемых научных ресурсов, которые ранее были труднодоступны — **материалов конференций, научных отчетов, диссертаций.**

Каждая уважающая себя конференция организует сайт и размещает на нем свои материалы задолго до начала работы. Такой сайт может функционировать длительное время и по завершении конференции. Здесь можно найти не только программу научной встречи, но и тезисы докладов, а нередко в последнее время и презентации.

Размещение научных отчетов — явление в Сети пока что скорее редкое, чем обыденное.

Лучше дело обстоит с диссертациями. В последнее время стремительно растет число университетских онлайн-архивов — **репозиториев**, которые заполняются, в том числе, и текстами бакалаврских, магистерских и докторских диссертаций.

Ситуация с книгами неоднозначна. С одной стороны, книга как коммерческий продукт не должна была бы присутствовать в открытом доступе. С другой стороны, правовой нигилизм сайтовладельцев (подкрепленный правовым нигилизмом читателей) настолько силен, что онлайн-архивы оказались наполненными книгами, размещенными там безо всякого согласия авторов.

3.2. Материалы конференций

Сайты конференций зарождаются в Сети, развиваются, затем с разной скоростью умирают. И остаются от них только мертвые ссылки, ведущие в никуда, в многочисленных каталогах.

Но случается и другое. Если оргкомитет жив, то материалы архивируются и долго могут быть доступны заинтересованной аудитории.

Пример такого информационного источника — сайт [INFORUM](http://www.inforum.cz/archiv/inforum2010/en/) (<http://www.inforum.cz/archiv/inforum2010/en/>), где размещены материалы серии конференций, проходивших в течение многих лет.

Самостоятельная работа

Сайт *INFORUM* почти идеален с точки зрения своего содержания и структуры. Исследуйте его строение.

Найдите типичные элементы хорошего сайта научной конференции: общую информацию для участников, программу последнего заседания, материалы последнего и предыдущих заседаний.

Какого элемента сайту явно недостает?

В Сети отсутствует единое хранилище бесплатных материалов прошедших конференций. В последние годы кое-что начало накапливаться в онлайн-репозиториях (*о них — позже*).

Примеры метасайтов, информирующих о *предстоящих* конференциях:
[План проведения \(http://belisa.org.by/ru/actions/\)](http://belisa.org.by/ru/actions/) научных и научно-технических мероприятий в Беларуси.

[Наука-форум \(http://www.science-forum.ru/index.php?name=conf\)](http://www.science-forum.ru/index.php?name=conf) — Поиск конференций и выставок (страны СНГ, другие страны).

[Конференции. Conferences. Konferenzen \(http://e-conference.ru/modules.php?name=ConfSearch\)](http://e-conference.ru/modules.php?name=ConfSearch) —

Место проведения: Россия и другие страны.

[Конференции.RU \(http://www.konferencii.ru/\)](http://www.konferencii.ru/) —

Место проведения: СНГ.

3.3.1. Диссертации

Наиболее вероятным местом обнаружения текста диссертации является электронный архив (*репозиторий*) того университета, где проходила процедура защиты.

Пример. В [Диссертационном архиве \(http://arrow.dit.ie/\)](http://arrow.dit.ie/) Дублинского технологического института (Ирландия) хранятся полные тексты магистерских и докторских диссертаций, защищенных в этом институте.

Ознакомьтесь со структурой архива и найдите ссылки, ведущие к спискам диссертаций по химии и к полным текстам диссертационных работ.

Во время занятий полные тексты извлекать не надо — файлы слишком большие.

После занятий найдите свободное время и просмотрите какую-нибудь диссертацию; зрелище познавательное и любопытное, особенно в разделе "*Acknowledgement*".

Некоторые страны имеют общенациональные архивы диссертационных работ и/или онлайн-центры, в которых накапливаются сведения из университетских репозиториях.

Пример: [DissOnline \(http://www.dissonline.de/eng/index.htm\)](http://www.dissonline.de/eng/index.htm) — центр электронных диссертаций Германии.

Упражнение

Ознакомьтесь со структурой *DissOnline*, найдите поисковые бланки (задача не из простых!) и определите функциональные возможности каждого из них.

Контрольное задание

Выпускник химического факультета БГУ, бывший научный сотрудник НИИ ФХП Дмитрий Талапин защитил докторскую диссертацию в Германии.

В каком году проходила защита, в каком университете, как звучала тема диссертации? Имеется ли в каком-нибудь архиве бесплатная электронная версия диссертации? Если есть — каким образом ее можно получить? Если нет — какие сведения из диссертации доступны без оплаты?

Еще примеры национальных информационных центров:

[Theses Canada Porta, Канада \(http://www.collectionscanada.gc.ca/thesescanada/index-e.html\)](http://www.collectionscanada.gc.ca/thesescanada/index-e.html)

[DiVA, Швеция \(http://www.diva-portal.org/smash/search.jsf\)](http://www.diva-portal.org/smash/search.jsf)

[NARCIS, Нидерланды \(http://www.narcis.info/index/tab/narcis\).](http://www.narcis.info/index/tab/narcis)

[ADT, Австралия \(http://adt.caul.edu.au/\).](http://adt.caul.edu.au/)

Как можно убедиться на этих примерах, далеко не каждый информационный центр имеет структуру, удобную для пользователя.

Для обнаружения нужных диссертационных работ, рассеянных по многочисленным архивам, прибегают к вспомогательным инструментам.

Около 2 млн. описаний диссертаций и ссылок на соответствующие архивы имеется в базе данных международной организации **Networked Digital Library of Theses and Dissertations (NDLTD)** — но здесь содержатся сведения лишь о части онлайн-ресурсов такого типа.

[На странице поискового бланка \(http://thumper.vtls.com:6090/\)](http://thumper.vtls.com:6090/) можно рассортировать содержимое базы данных *NDLTD* по следующим критериям: язык, страна, формат документа, предметная область,— а также сформулировать запрос.

Вопросы

Имеются ли в базе данных *NDLTD* сведения о русскоязычных диссертациях?

Имеются ли в базе данных *NDLTD* сведения о диссертациях по химии?

В странах СНГ диссертант перед защитой должен публиковать **автореферат** — брошюру с кратким изложением своей диссертационной работы.

Автореферат диссертации тоже является ценным источником научной информации.

[Электронная библиотека АВТОРЕФЕРАТЫ](http://referat.vak.org.by/index.php)

[\(http://referat.vak.org.by/index.php\)](http://referat.vak.org.by/index.php)

ВАК Беларуси публикует авторефераты диссертаций, представленных к защите. (Архив авторефератов защищенных диссертаций отсутствует).

Упражнение

Определить, имеются ли авторефераты по химии в этой базе данных.

[Электронная библиотека авторефератов диссертаций Украины \(http://www.nbu.gov.ua/eb/ard.html\)](http://www.nbu.gov.ua/eb/ard.html)

Упражнение

Определить, имеются ли авторефераты по химии в этой базе данных; если да — то каково их примерное количество?

Разрешен ли свободный доступ к полным текстам авторефератов?

[Электронная библиотека диссертаций \(http://diss.rsl.ru/\)](http://diss.rsl.ru/) Российской государственной библиотеки содержит не только авторефераты, но и сами диссертации.

Онлайн-библиографический поиск доступен всем пользователям.

Полные тексты документов доступны в Зале электронных информационных ресурсов БГУ (БГУ является подписчиком этой библиотеки).

Упражнение

Определить, имеются ли в базе данных диссертационные работы, в названиях которых упоминается графен. Если да, то сколько из них было представлено на соискание ученой степени кандидата или доктора химических наук?

3.3.2. ProQuest Dissertations & Theses

В числе ресурсов *ProQuest* имеется диссертационный архив *Dissertations & Theses*, содержащий докторские и магистерские диссертации США, Канады, а также иных стран за период с 1997 года по настоящее время.

Более миллиона документов здесь — полнотекстовые.

Информационный центр *ProQuest* открыт для нас **только через сеть БГУ**.

Напоминаем основные правила формулирования запроса:

- Разрешены логические операторы *AND*, *OR*, *AND NOT*.
- Точная фраза обозначается кавычками.
- Если запрос состоит из двух слов без оператора, такой текст по умолчанию считается точной фразой.
- Если запрос состоит из трех слов без операторов, ведется поиск документов, в которых эти слова должны находиться по соседству.
- Поиск по словоформам терминов (*stemming*) не ведется.
- Шаблон * (*звездочка*) может заменять собой произвольное число символов в конце слова.
- Шаблон ? (*вопросительный знак*) может заменять собой один символ в середине или в конце слова.
- Разрешены операторы расстояния *W/n*, *PRE/n*, *W/PARA*.

Упражнение

Исследуйте поисковые бланки, работу поисковой программы, структуру списка результатов поиска, структуру отдельной записи.

Проверочное задание

Сколько диссертаций тематики "Химия" содержится в базе данных?

Сколько магистерских и докторских диссертаций, в названиях которых упоминается проблема токсичности наночастиц, имеется в базе данных? (наночастица — *nanoparticle*, токсичный — *toxic*, токсичность — *toxicity*).

Извлеките список диссертаций, в названиях которых термины *токсич(-ный, -ость и т.п.)* и *наночастиц(-а, ы)* находятся рядом.

Контрольное задание

Из базы данных *ProQuest Dissertations & Theses* извлеките список докторских диссертаций, выполненных под руководством Нобелевских лауреатов 2010 года по химии. (Справка: научный руководитель — *Advisor*).

3.4. Книги

В соответствии с законами, защищающими авторские права, издательства обычно запрещают перевод печатных изданий в электронный формат. Тем не менее, книги можно встретить в *WWW*. Такое случается, во-первых, если права

принадлежат автору и автор желает открыть свою работу для широкой публики. Во-вторых, издательства в целях рекламы иногда публикуют в Интернете избранные главы только что выпущенных книг. В-третьих, энтузиасты публикуют ту литературу, для которой истек срок копирайта. В-четвертых, есть энтузиасты, для которых законы не писаны. В-пятых, в Сети все больше появляется чисто электронных изданий — но это, в основном, учебники.

Рассмотрите строение книги "Основы ЯМР" ([Joseph P. Hornak. The Basics of NMR \(http://www.cis.rit.edu/htbooks/nmr/nmr-main.htm\)](http://www.cis.rit.edu/htbooks/nmr/nmr-main.htm)). Это типичное онлайн-издание, насыщенное гиперсвязями, статическими и динамическими иллюстрациями.

Тем, кто испытывает затруднения с английским языком, рекомендую обратиться к расположенной на сайте химфака МГУ [Электронной библиотеке по химии \(http://www.chem.msu.ru/rus/elibrary/\)](http://www.chem.msu.ru/rus/elibrary/). Выберите одно из имеющихся здесь учебных пособий и ознакомьтесь с его структурой.

Электронные версии печатных книг присутствуют в Сети, как правило, в форматах PDF и DJVU. Вот некоторые полезные **архивы и метасайты** научной и, в частности, химической литературы:

[Большая Научная Библиотека](http://sci-lib.com/subject.php?subject=3&pp=1)

[\(http://sci-lib.com/subject.php?subject=3&pp=1\)](http://sci-lib.com/subject.php?subject=3&pp=1)

[Библиотека Ихтика \(http://ihtik.lib.ru/dreamhost_chem_8janv2007.html\)](http://ihtik.lib.ru/dreamhost_chem_8janv2007.html)

[CHEMPORT.ru \(http://www.chemport.ru/?cid=29\)](http://www.chemport.ru/?cid=29)

[NeHudLit.Ru \(http://nehudlit.ru/books/cat352.html\)](http://nehudlit.ru/books/cat352.html)

Адреса и описания книжных архивов самой разнообразной тематики собраны на метасайте

[Free Books \(http://www.e-book.com.au/freebooks.htm\)](http://www.e-book.com.au/freebooks.htm) и его клоне

[Open Access e-Books \(http://www.digitallibrary.edu.pk/OAEBooks.html\)](http://www.digitallibrary.edu.pk/OAEBooks.html).

Для поиска электронных книг предназначены **специализированные поисковые системы Рунета**:

[eBdb \(http://www.ebdb.ru/Default.aspx?\)](http://www.ebdb.ru/Default.aspx?)

[Поиск электронных книг \(http://www.poiskknig.ru/index.html\)](http://www.poiskknig.ru/index.html)

При поиске книг с помощью универсальной поисковой системы **Google** стоит учитывать формат, в котором чаще всего может присутствовать книга в вебе. Нередко к успеху приводит конструкция запроса подобного типа:

автор filetype:pdf

название filetype:djvu

Контрольное задание

Выполните информационный поиск полнотекстовой монографии по заданию преподавателя.

3.5. Google Books

Книжный архив, формируемый компанией *Google*:

[Google Books \(http://books.google.com/books?hl=en\)](http://books.google.com/books?hl=en) —

английский интерфейс,

[Google Книги \(http://books.google.ru/books?hl=ru\)](http://books.google.ru/books?hl=ru) — русский интерфейс.

Компания *Google* приступила к переводу печатных фондов библиотек в электронный формат. Доступность к полным текстам книг, накапливающихся в базе данных, варьируется в широких пределах; есть здесь книги и платные, и бесплатные.

Важно понимать, что *Google Books (Google Книги)* - это автономный архив. Поисковая программа *Google Books (Google Книги)* ищет книги не по всему вебу, а только по своей базе данных.

- *Google Book Search* проводит информационный поиск по всему тексту книги.
- *Google Book Search* проводит информационный поиск в основном по тем же правилам, что и поисковая система *Google*.

В списке результатов поиска о степени доступности материала можно судить по следующим пометкам:

- *No preview* — *Без предварительного просмотра* — читатель может ознакомиться только с библиографическим описанием книги;
- *Snippet view* — *Просмотр фрагмента* — на экран можно вывести несколько строк, содержащих поисковые термины;
- *Preview* — *Предварительный (ограниченный) просмотр* — можно прочитать ограниченное число страниц (иногда — значительное); сохранение книги на компьютере пользователя запрещено;
- *Full view* — *Полный просмотр* — книга полностью бесплатна для чтения.

Значительную часть книг категории "Полный просмотр" можно не только читать на экране, но и сохранять на своем компьютере. В таких случаях на веб-странице имеется ссылка со словом PDF.

Примечание.

Если в списке результатов поиска помечено, что полный текст данной книги недоступен, стоит обратить внимание на пункт "*Другие издания*" ("*Other editions*"). Бывает, что более раннее издание этой же книги доступно читателю полностью.

Упражнение 1

Проверьте, имеются ли в базе данных книги химической тематики. Есть ли в их числе книги, полные тексты которых доступны бесплатно? Есть ли в архиве книги по химии на русском языке?

Упражнение 2

Проверьте, можно ли сохранять на своем компьютере информацию, обнаруженную в базе данных *Google Books*.

4. РЕПОЗИТОРИИ ЭЛЕКТРОННЫХ ПУБЛИКАЦИЙ

4.1. Введение

В последние годы в Сети заметно увеличивается количество онлайн-архивов (**репозитории**), которые медленно, но верно наполняются научными и учебными материалами разнообразных типов: текстовыми, графическими, видео.

Репозитории бывают университетские, ведомственные; национальные и межнациональные; тематически специализированные и многодисциплинарные; хранящие материалы нескольких типов или одного из них (*пример*: электронный университетский архив диссертационных работ).

В репозитории каждый хранящийся объект совершенно автономен, никак не связан с другими объектами.

С нашей точки зрения, все более ценными источниками информации становятся *университетские репозитории*. Теперь считается хорошим тоном, если университет предоставляет своим сотрудникам возможность (или поощряет, или даже требует) размещать электронные версии статей, отчетов, докладов в архиве, открытом для всего Интернета. И университет, и его сотрудники таким простым способом рекламируют свою деятельность, а мы получаем доступ к уникальным ресурсам, причем не разбросанным, как ранее, по многочисленным сайтам ученых, а относительно сконцентрированным.

Есть один нюанс, из-за которого мы выделили репозитории в особую группу онлайн-научных источников.

Научная надежность материалов, содержащихся в таких архивах, строго говоря, неопределенная.

В репозитории можно встретить и никем не отрецензированный препринт, и материал, прошедший стадию рецензирования, и факсимильную копию статьи, опубликованной в авторитетном международном журнале.

Препринт и постпринт одной и той же статьи могут иметь различия, иногда существенные.

Оценка достоверности документа — обязательная стадия работы с материалами, полученными из репозитория.

В английской терминологии электронные публикации, структурно не входящие в журнал, часто объединяют общим названием **e-prints** — это препринты, постпринты, тезисы докладов, презентации, научные и технические отчеты, диссертации, информационные бюллетени с выставок и т. д. В репозиториях основная часть хранящегося материала обычно относится к категории *e-prints*.

4.2. Репозитории электронных публикаций

Самостоятельная работа

Ознакомьтесь со структурой и функциями перечисленных ниже сайтов.

Пример **университетского** репозитория: [University of Twente Publications](http://doc.utwente.nl/) (<http://doc.utwente.nl/>) — архив университета Твенте, Нидерланды.

Проверочное задание

Из базы данных репозитория извлеките список тех статей, в названиях которых присутствует слово *phosphate*, причем отберите статьи, доступные всем пользователям, не только сотрудникам университета.

Спрогнозируйте вероятность получения полных текстов этих же статей с сайта соответствующего журнала и затем проверьте свой прогноз.

Содержимое репозитория находится частично в области "открытого Интернета" (меньшая доля), частично в области "скрытого Интернета" (большая доля).

Google и другие универсальные и специализированные поисковые системы пока что не могут дать исчерпывающей информации о содержимом всех онлайн-репозиториях.

Возможны следующие *пути целенаправленного информационного поиска* в репозиториях:

а) исходя из неких соображений определить для себя круг университетов, на серверах которых может содержаться интересующая нас информация, и анализировать архивы этих университетов;

б) найти межведомственные специализированные репозитории, хранящие информацию по интересующей нас тематике;

в) воспользоваться специальными средствами полнотекстового поиска в серии репозиториях.

Инструменты для обнаружения заданного репозитория

[Directory of Open Access Repositories — OpenDOAR](http://www.open_doar.org/index.html) **(http://www.open_doar.org/index.html)**

Содержит информацию почти о 2 тыс. репозиториях, расположенных на всех континентах и содержащих документы, составленные на многих языках.

На сайте *OpenDOAR* имеются следующие инструменты:

[Search or Browse for Repositories \(http://www.open_doar.org/find.php\)](http://www.open_doar.org/find.php) — поиск репозитория по его тематике, местонахождению, наполнению;

[OpenDOAR — Countries and Organisations](http://www.open_doar.org/countrylist.php)
(http://www.open_doar.org/countrylist.php) — перечень организаций-владельцев репозиториях.

[Registry of Open Access Repositories — ROAR \(http://roar.eprints.org/\)](http://roar.eprints.org/)

Список, адреса и краткие описания более тысячи репозиториях.

Список можно сортировать по географическому признаку, по типу архива и по некоторым другим параметрам.

Инструмент для информационного поиска в группе репозиториях:

[Search Repository Contents \(http://www.open_doar.org/search.php\)](http://www.open_doar.org/search.php) — полнотекстовый поиск в документах, хранящихся в тех репозиториях, которые зарегистрированы в **OpenDOAR**.

Проверочное задание.

Определите, имеются ли в репозиториях такие полнотекстовые документы, где обсуждается проблема наличия воды в лунных минералах.

[Электронная библиотека БГУ \(http://elib.bsu.by/\)](http://elib.bsu.by/)

Репозиторий БГУ находится в стадии отладки и только-только начал пополняться материалами.

Проверочное задание

Извлеките списки документов химической тематики, имеющихся в Электронной библиотеке БГУ.

4.3. Персональные сайты ученых

Если на сайте журнала статья доступна только за плату, стоит попытаться найти ее бесплатный препринт или репринт на персональном сайте автора или в репозитории той организации, где работает автор.

Возможности репозиториев мы рассмотрели, теперь обратим наше внимание на сайты ученых.

Рассмотрим несколько путей поиска списка публикаций заданного автора.

1) Место работы автора известно.

Пример.

Выпускник химического факультета БГУ *Олег Выдров* в настоящее время является сотрудником знаменитого Массачусетского технологического института (*Massachusetts Institute of Technology, MIT*).

Задача 1

Дано: фамилия автора (*Выдров*) и место его работы.

Найти:

- а) сайт Массачусетского технологического института, а затем
- б) на сайте института — персональный сайт О. Выдрова и, наконец,
- в) список публикаций О. Выдрова.

Есть ли на сайте автора какая-либо порция информации, которую мы не смогли бы получить на сайтах соответствующих журналов?

2) Место работы автора неизвестно.

Пример.

Выпускник химического факультета БГУ Андрей Рогач, в настоящее время работающий за границей, находится на 51-м месте в списке Top 100 Materials Scientists прошедшего десятилетия.

Задача 2

Дано: фамилия и один инициал автора (*A. Rogach*).

Найти: список публикаций А. Рогача.

Использовать *Google*, поиск провести в одну-две стадии. Какие термины целесообразно включать в запрос для решения данной задачи?

Контрольное задание

Дипломник А. изучает процессы, проходящие в растворах на поверхности диоксида кремния. Случайно он узнал, что по его теме существует монография Ю. Холина, профессора Харьковского университета. Найдите эту книгу.

5. ВТОРИЧНЫЕ ИСТОЧНИКИ НАУЧНОЙ ИНФОРМАЦИИ

5.1. Введение

Научные статьи, материалы конференций, диссертации — это *первичные* информационные источники, потому что именно здесь публикуются результаты научных исследований. Сведения о первичных информационных источниках суммируются (а иногда и обобщаются) во *вторичных* источниках.

В сфере печатной литературы особенно важными вторичными информационными источниками являются *реферативные журналы*. Традиционный реферативный журнал — это сборник библиографических описаний и рефератов рецензированных публикаций.

В области химии издаются универсальные реферативные журналы (*Chemical Abstracts*, *РЖ Химия*) и специализированные (например, *Analytical Abstracts*).

В сфере онлайн-литературы вторичные информационные источники тоже есть, их много, причем они характеризуются разнообразием типов и функциональных возможностей. Некоторые из них только отдаленно напоминают своих предков — бумажные реферативные журналы.

Индекс любой поисковой системы — это тоже вторичный информационный источник. Проблема состоит в том, что в этом индексе свалено в кучу огромное количество научных и антинаучных публикаций. *Google* легко выдаст длинный список веб-страниц, на которых упоминается некая статья, но вот выбрать из этого списка нужный пункт — отдельная задача, и не всегда простая. (А если учесть, что нужная статья может находиться не в открытом, а в скрытом Интернете, то поисковая система ее и не обнаружит...)

Специальные базы данных и поисковые программы, предназначенные для ученого мира, не утратили своего значения даже в эпоху *Google*. Более того, этот поисковый гигант, понимая проблему, разработал и совершенствует собственный инструмент поиска научных публикаций — *Google Scholar*.

К сожалению, онлайн-доступ к непревзойденной химической реферативной базе данных — *Chemical Abstracts* — только платный.

Бесплатные вторичные информационные источники имеют один общий досадный недостаток: они не универсальны, каждый из них охватывает только определенную часть онлайн-овых (иногда и офлайн-овых) ресурсов.

К примеру, на сайте *ScienceDirect* размещены рефераты всех публикаций издательства *Elsevier* — но только этого издательства.

OpenDOAR имеет инструмент поиска электронных документов — но только в тех репозиториях, которые зарегистрированы в *OpenDOAR*.

В этом разделе практикума мы рассмотрим возможности нескольких вторичных информационных ресурсов, базы данных которых характеризуются заметной шириной охвата.

5.2. РЖ Химия — Реферативный журнал "Химия"

Реферативный журнал "Химия" — один из пакета *Реферативных журналов*, размещенных на сайте зональной научной библиотеки Воронежского государственного университета.

- Онлайн-версия *Реферативного журнала "Химия"* является обогащенной копией бумажной версии.
 - Журнал реферирует только рецензируемые первоисточники.
 - В онлайн-версии приводятся описания процитированных источников, но отсутствуют гиперссылки к статьям.
- База данных содержит рефераты, опубликованные начиная с 2004 г.

[Поисковый бланк реферативного журнала](http://www.lib.vsu.ru/cgi-bin/zgate?Init+rj.xml,simple_rj.xsl+rus) находится по адресу:
http://www.lib.vsu.ru/cgi-bin/zgate?Init+rj.xml,simple_rj.xsl+rus

Особенности формулирования запроса:

- Поиск по словоформам поисковых терминов *не проводится*.
- Разрешен шаблон * для замены произвольного числа символов в конце слова.
- При поиске в поле *Заглавие* анализируются разные поля записи, не только название статьи.

Упражнение

Ознакомьтесь со структурой онлайн-версии *Реферативного журнала "Химия"*.

Проведите пробный поиск информации, проанализируйте список результатов поиска.

Есть ли в онлайн-версии такие структурные элементы, которые отсутствуют в печатной версии?

Контрольное задание

Предположим, возникла необходимость быстро ознакомиться с литературой, в которой обсуждается метод *электрохимической импедансной спектроскопии*.

Извлеките оптимальный список публикаций.

5.3. Тематические (специализированные) базы данных

Пример:

[INIS \(http://www.iaea.org/INIS\)](http://www.iaea.org/INIS) — литературная база данных МАГАТЭ.

Тематика *INIS*: все аспекты, связанные с ядерной энергетикой, в том числе, радиационная химия, радиохимия, экология, материаловедение.

Типы документов: статьи, книги, научные отчеты, диссертации, материалы конференций и др.

База данных содержит большое количество полнотекстовых документов.

Самостоятельная работа.

Рассмотрите структуру усложненного поискового бланка, проведите информационный поиск, изучите строение стандартной записи, определите уровень доступности извлекаемого материала.

Контрольное задание

В базе данных *INIS* подберите небольшой список литературы для изучения следующей проблемы: *концентрация радона в атмосфере жилого дома*.

Может ли концентрация радона различаться на разных этажах дома? Если да, то насколько сильно? Если да, то почему?
Полнотекстовые ресурсы приветствуются.

5.4. Google Scholar

[Google Scholar \(http://scholar.google.com/\)](http://scholar.google.com/) — представитель нового поколения вторичных источников научной информации.

Google Scholar — это поисковая система, которая специализируется на индексации одного типа материалов — *научных публикаций* (статей, книг, препринтов и т. д.).

Как и универсальная поисковая система *Google*, специализированная *Google Scholar* сообщает пользователю название, фрагмент текста и гиперсвязь к документу.

Робот *Google Scholar* посещает только сайты, имеющие отношение к науке, и собирает в свой индекс информацию о местонахождении и содержании научных работ. В базу данных попадают сведения и о бесплатных полнотекстовых статьях, и о тех, у которых доступны только рефераты, и о тех, для которых известны только библиографические описания.

Google Scholar содержит сведения как о тех статьях, которые опубликованы в журналах, так и о тех, которые хранятся в репозиториях или находятся на сайтах научных коллективов или отдельных ученых.

- *Google Scholar* содержит сведения не только об онлайн-овых, но и о печатных статьях.

Робот *Google Scholar* индексирует онлайн-овые научные публикации. Если в онлайн-овом документе в списке литературы обнаруживается ссылка на офлайн-овый документ, библиографическое описание такого печатного документа тоже попадает в базу данных *Google Scholar*.

В списке результатов поиска офлайн-овые статьи имеют пометку [Citation].

- Список результатов поиска содержит гиперсвязи, ведущие к веб-страницам, где имеются сведения о статье (как минимум — библиографическое описание).

Конечно же, в списке отсутствуют гиперсвязи к офлайн-овым источникам.

- В списке результатов поиска может быть несколько ссылок на материалы, относящиеся к одной и той же статье (например, ссылки на сайт издательства, на сайт агрегатора, на реферативную базу данных, на персональный сайт автора).

- В списке результатов поиска ссылки на бесплатные полные тексты публикаций имеют пометки [PDF].

В списке может быть несколько ссылок на несколько полнотекстовых версий одной и той же статьи (например, на окончательную версию на сайте издательства и на препринт на сайте автора).

- Поисковая программа *Google Scholar* работает по тем же правилам, что и поисковая программа *Google*.

- Индексы *Google* и *Google Scholar* — это разные базы данных.

Google Scholar выполняет не только информационные, но и наукометрические функции.

Из списка результатов поиска по гиперсвязи *Cited by* можно получить

сведения о том, сколько и какие именно документы ссылаются на конкретную публикацию в пределах базы данных *Google Scholar*.

Число в *Cited by* отражает степень авторитетности и известности публикации.

Замечание о темных сторонах Google Scholar:

База данных *Google Scholar* заполняется не человеком, а роботом. Анализируя веб-страницы, робот не всегда адекватно обнаруживает элементы библиографического описания, и поэтому иногда получаются любопытные результаты.

Например, в базе данных можно найти статьи таких авторов как *Subscribe*, *Abstract*, *Belarus*. (Это утверждение проверьте экспериментально).

Проверочные задания

Проведите информационный поиск в *Google Scholar* по теме "Использование эпоксидных материалов в разного рода покрытиях" (возможные термины: *epoxy*, *coating*).

Ознакомьтесь со структурой полученного списка результатов, исследуйте назначение гиперссылок.

По каким параметрам *Google Scholar* подобна обычным библиографическим (или реферативным) базам данных и по каким параметрам отличается от них?

Проведите такой же тематический поиск в индексе *Google*. Есть ли существенные различия в результатах поиска в *Google* и в *Google Scholar*?

Определите, можно ли использовать *Google Scholar* для поиска статей, опубликованных на русском языке.

Контрольное задание

В *Google Scholar* извлеките список опубликованных в 2011 году PDF-документов, в которых может содержаться информация о наличии воды в лунных минералах.

(По какой причине в задании сказано "может содержаться информация", а не "содержится информация"?)

Проведите такой же поиск в *Google*.

Есть ли существенные различия в результатах поиска PDF-документов в *Google* и в *Google Scholar*?

5.5. Scirus

[Scirus \(http://www.scirus.com/srsapp/\)](http://www.scirus.com/srsapp/) — научная поисковая система.

Робот поисковой системы *Scirus* сканирует Интернет, находит сайты, имеющие отношение к науке, и индексирует их.

Список результатов поиска в *Scirus* практически не засорен ссылками на рекламные, коммерческие, общественно-политические, аморальные и тому подобные сайты.

С другой стороны, не следует полагать, что вся имеющаяся в Интернете научная информация зарегистрирована в базе данных *Scirus*.

- *Scirus* содержит сведения о статьях, патентах и веб-страницах.
- *Scirus* извлекает информацию из журналов ряда издательств (*Elsevier* и

- др.), репозиториев, сайтов научных коллективов и персональных страниц ученых.
- *Scirus* содержит сведения о рецензируемых и нерецензируемых первоисточниках.
 - В базе данных содержатся сведения о 450 млн. документов.

Основные правила формулирования запроса:

- Задание может быть сформулировано с использованием операторов. Разрешены операторы *AND*, *OR*, *ANDNOT*; двойные кавычки; шаблон * в конце термина.
- Для формулирования запроса, содержащего несколько поисковых параметров, целесообразно использовать бланк *Advanced Search*.
- Возможен поиск по отдельным полям записей, в т. ч., по полям *Название*, *Автор*.
- В запросе могут быть указаны тип документа, дата создания, тип и место хранения, некоторые иные параметры.
- Список результатов поиска можно разбивать на кластеры.

Упражнение

Исследуйте структуру усложненного поискового бланка. В разделе "*Content sources*" обратите внимание на перечень сайтов, информация из которых попадает в базу данных.

Проведите информационный поиск по любой теме и проанализируйте структуру списка обнаруженных документов.

Контрольное задание

Выполните контрольное задание, выданное преподавателем.

5.6. Личная библиографическая база данных в социальной сети

В процессе научной работы каждый ученый накапливает собственный массив сведений о публикациях коллег — современников и предшественников. Массив постоянно растет, и при достижении некоей критической величины ориентироваться в нем становится трудно. Выход — в систематизации собранной информации. Раньше для этого использовали бумажные библиографические карточки, теперь — библиографические базы данных.

Для создания на персональном компьютере собственной библиографической базы данных существуют прикладные программы — "менеджеры ссылок" (*Reference Manager*, *EndNote*, *ProCite* и др.). Большей частью такие программы либо платные, либо требуют немалых усилий при формировании базы данных, либо не очень дружелюбны по отношению к пользователю. Изучение программного обеспечения этого типа отложим до иных времен.

Теперь мы будем знакомиться с *онлайновым* менеджером ссылок, ориентированным именно на пользователя — ученого. Один из таких инструментов функционирует и развивается при финансовой поддержке компании *Springer*, и называется он:

[CiteULike](http://www.citeulike.org) — "Цитируйте, что Вам Нравится"
(www.citeulike.org)

- *CiteULike* предоставляет любому пользователю Интернета место на своем сервере для размещения личной или коллективной библиографической базы данных, а также программное обеспечение для работы с этой информацией.
- *CiteULike* помогает пользователю извлекать в автоматическом режиме библиографическую информацию из онлайн-источников для его базы данных.

Права пользователей CiteULike:

- Зарегистрированный пользователь имеет право формировать свою собственную библиографическую базу данных, которая хранится на сервере *CiteULike*.
- Группа зарегистрированных пользователей имеет возможность формировать свою коллективную библиографическую базу данных, которая хранится на сервере *CiteULike*.
- Составитель базы данных может частично или полностью открыть свои ресурсы для некоторых коллег или даже для всех пользователей Интернета.
- Незарегистрированные посетители сайта имеют возможность знакомиться с той частью информации, которая открыта для всех.

Общая схема составления базы данных

Рассмотрим типичную ситуацию.

Путешествующий по вебу зарегистрированный пользователь встречает интересную статью.

Для того, чтобы сохранить ее описание в своей базе данных, он должен выполнить следующие действия:

1. Перейти на страницу, на которой имеется библиографическое описание статьи — обычно это веб-страница с рефератом или с полным текстом.
2. Нажать на кнопку [Post to CiteULike] на панели закладок браузера (*где эта кнопка, как ее устанавливать — разберемся позже*).
На этом этапе *CiteULike* приступает к интеллектуальному анализу веб-страницы — находит на ней фамилии авторов, название статьи, место публикации, реферат — и всю информацию заносит в нужные поля базы данных. Анализ и копирование информации протекают в автоматическом режиме.
3. На экран выводится страница с результатом, который при необходимости можно подкорректировать вручную.
4. Пользователь на бланке указывает режим доступа к сформированной записи (т.е. должна ли она быть видна всем или только ему лично) и нажимает соответствующую кнопку. Операция выполнена.

Некоторые *ограничения*:

- *CiteULike* настаивает, чтобы составители включали в свои базы данных ссылки только на прорецензированные публикации.
- *CiteULike* умеет корректно извлекать библиографическую информацию с сайтов ведущих издательств; в остальных случаях записи приходится подправлять вручную.

CiteULike — больше, чем обычный архив. Это:

- сайт категории Веб 2.0: здесь онлайн-ресурс создается коллективными усилиями;
- инструмент оценки качества: база данных формируется из тех публикаций, которые, по мнению пользователей, являются особенно важными или интересными;
- научная социальная сеть, в которой есть группы по интересам.

Самостоятельная работа 1

Общее знакомство с *CiteULike*. Работа в качестве незарегистрированного пользователя.

Изучите структуру открытой базы данных (*Everyone's Library*), типовую структуру записи, принципы ведения информационного поиска.

Некоторые этапы работы незарегистрированного пользователя в режиме *Browse* показаны на иллюстрациях:

www.citeulike.org

http://www.citeulike.org/

CiteULike: Everyone's library

[Browse](#) | [FAQ](#) | [News](#) | [Log in](#)

Search citeulike

citeulike | sponsored by Springer

citeulike is a free service for managing and discovering scholarly references

3,224,085 articles - 7,703 added today.

- Easily store references you find online
- Discover new articles and resources
- Automated article recommendations **NEW**
- Share references with your peers
- Find out who's reading what you're reading
- Store and search your PDFs

[Join now](#)

Список самых последних добавлений в Общую библиотеку.
 Всем пользователям разрешен доступ к этой информации

Everyone's library **Export**

Кнопка, которая формирует файл со списком литературы, отображенной на странице

Some recent papers posted to CiteULike - all mixed together.
[Hide details](#)

The network is the database: data management for highly distributed systems
 In SIGMOD '01: Proceedings of the 2001 ACM SIGMOD international conference on Management of data, pp. 103-112. Edited by Julio C. Navas, Michael Wynblatt. New York, NY: ACM Press, 2001. 103 pages. ISBN 0-918986-40-9. [Posted to partitioning distributed-databases databases by ejones on 2009-11-13 18:34:41](#)

Precedence-based speech segregation in a virtual auditory environment
 The Journal of the Acoustical Society of America, Vol. 118, No. 5. (2005), pp. 3241-3251. by Douglas S. Brunqart, Brian D. Simpson, Richard L. Freyman. [Posted to segregation psych precedenceeffect intelligibility by asterix77 on 2009-11-13 18:32:23 as ★★](#)

Identification of Codon-Specific Serine to Asparagine Mistranslation in Recombinant High-Resolution Mass Spectrometry
 Analytical Chemistry, Vol. 81, No. 22. (15 November 2009), pp. 9282-9290. by X. Christopher Yu, Oleg V. Borisov, Melissa Alvarez, David A. Michels, Yajun J. Wang, Y. [Posted to codon by carlesjavierna on 2009-11-13 18:32:23 as ★★](#)

Everyone's Tags
 Most active tags on CiteULike

Filter:

[\[Display as Cloud\]](#)

review	320
evolution	142
model	111
cancer	106
network	106
statistics	102
theory	94

Тэг (Tag) – это ключевое слово, которое характеризует тематику статьи.
 Тэг присваивается тем человеком, который вносит статью в базу данных.
 Если вести поиск по тэгу, можно извлечь статьи одинаковой тематики.

В правой колонке перечислены тэги, которые особенно часто встречаются в базе данных (библиотеке).

Структура полной записи

Precedence-based speech segregation in a virtual auditory environment **Export**
 by: Douglas S. Brunqart, Brian D. Simpson, Richard L. Freyman
 The Journal of the Acoustical Society of America, Vol. 118, No. 5. (2005), pp. 3241-3251.

Citation Format
 Plain
[\[Posts\]](#)

View FullText article
 Scitation (AIP), Scitation (AIP) (Permanent Link)

▶ **Reviews** [\[Write a review of this article\]](#)
 ▼ **Find related articles from these CiteULike users**
 asterix77
 ▼ **Find related articles with these CiteULike tags**
 intelligibility, precedenceeffect, psych, segregation
 ▶ **Posting History**
 ▼ **Abstract**

When a masking sound is spatially separated from a target speech signal, substantial releases from masking typically occur both for speech and noise maskers. However, when a delayed copy of the masker is also presented at the location of the target speech (a condition that has been referred to as the front target, right-front masker or F-RF configuration), the advantages of spatial separation vanish for noise maskers but remain substantial for speech maskers. This effect has been attributed to precedence, which introduces an apparent spatial separation between the target and masker in the F-RF configuration that helps the listener to segregate the target from a masking voice but not from a masking noise. In this study, virtual synthesis techniques were used to examine variations of the F-RF configuration in an attempt to more fully understand the stimulus parameters that influence the release from masking obtained in that condition. The results show that the release from speech-on-speech masking caused by the addition of the delayed copy of the masker is robust across a wide variety of source locations, masker locations, and masker delay values. This suggests that the speech unmasking that occurs in the F-RF configuration is not dependent on any single spatial cue and may indicate that F-RF speech segregation is only partially based on the apparent left-right location of the RF maskers.

▶ **BibTeX record**
 ▶ **RIS record**

Кнопка, пересылающая библиографическое описание в текстовый файл

Выбор разных стандартов отображения библиографического описания

К первоисточнику

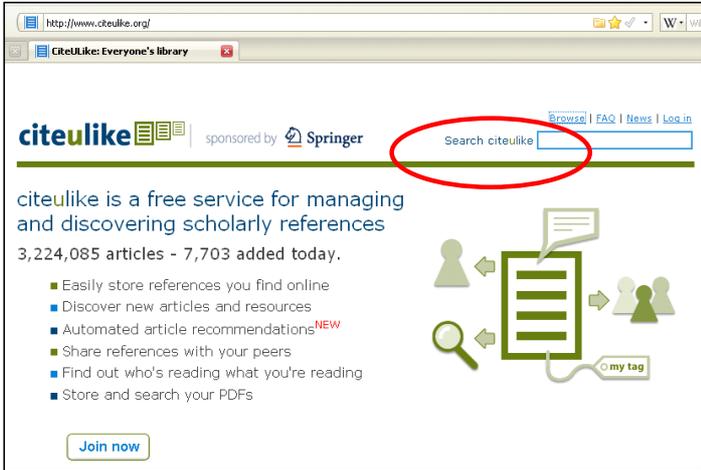
Тэги - указывают тематику статьи

Реферат

Некоторые этапы работы незарегистрированного пользователя в режиме Search показаны на следующих иллюстрациях:

Поисковые бланки – только простейшие

1



2



Результаты поиска

Текст запроса

Search results for: toxic* && nanoparticle* [78 articles]

All articles on CiteULike matching your search criteria

Users interested in: toxic* && nanoparticle*
 monkare faizalm sobolevnmr Marine ml115 Quarem pak benjaminjones anyscientist liyanyn
 ponsamuel calvin Tobenspencer

Groups interested in: toxic* && nanoparticle*

- baker-group
- Molecular dynamics in MSU

Articles discussing: toxic* && nanoparticle*
[Hide details](#)

The toxicological impact of nanoparticles
Nano Today, Vol. 3, No. 1-2. (February 2008), pp. 48-55.
 by [R. Brayner](#)
 posted to [nanoparticle osh toxicology](#) by [faizalm](#) on 2009-03-12 06:59:18 as ★★

Inhalation Toxicity and Lung Toxicokinetics of C60 Fullerene Nanoparticles and Microparticles
Toxicol. Sci., Vol. 101, No. 1. (1 January 2008), pp. 122-131.
 by [Gregory L. Baker](#), [Amit Gupta](#), [Mark L. Clark](#), [et al.](#)
 posted by [2 people](#)

Эти пользователи вносили в базу данных информацию по теме запроса

Группы пользователей, имеющие в своих библиотеках информацию по теме запроса

Список обнаруженных научных публикаций

Алгоритм работы поисковой программы специфичен, но для рутинного поиска пригоден следующий минимум правил:

- По умолчанию, поиск проводится во всех полях, включая реферат.
- Для поиска по избранному полю следует перед термином указать название поля в формате **поле:термин**, например:
tag:phosphate
author:paulechka или **author:"lesnikovich ai"**
- Шаблон * заменяет любое количество символов.
- Шаблон ? заменяет один символ.
- Оператор по умолчанию: *OR*.
- Оператор *and* отображается символами &&, например:
title:coating* && title:phosphate*
- Оператор *not* отображается знаком ! (слитно с термином), например:
coating* !phosphate*

Проверочная работа.

Извлеките список публикаций по хемоинформатике (*chemoinformatics, cheminformatics, chemiinformatics*).

Самостоятельная работа 2

Настройка браузера для работы в качестве зарегистрированного пользователя.

Операции по настройке браузера показаны на следующих иллюстрациях:

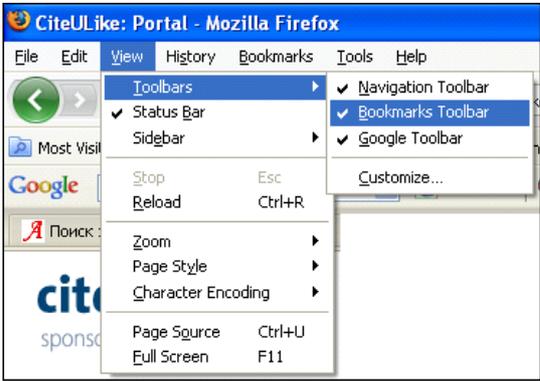
Настройка браузера для работы с CiteULike

1. Удостоверимся, что среди панелей инструментов браузера присутствует панель для закладок.

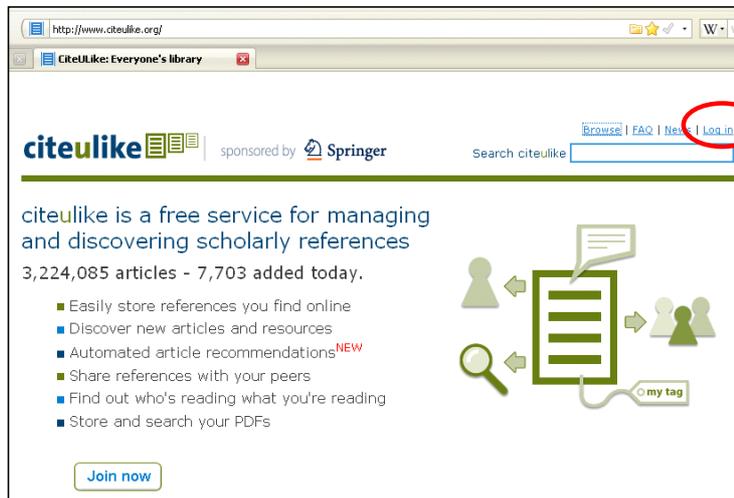
Если отсутствует – выведем ее, выполнив последовательность команд:
Вид → Панели инструментов → Панель закладок

Панель закладок

Наша задача – поместить на эту панель специальную кнопку



2. На сайте *CiteULike* пройдем авторизацию (ссылка Log in)



в любом месте

Username и Password
сообщит преподаватель



3. Находим ссылку, ведущую к нужной нам кнопке

При успешной авторизации на экран выводится страница с двумя закладками:
CiteULike и MyCiteULike.

*Терминологическое
пояснение.*

CiteULike –
раздел, общий для всех
пользователей Интернета.

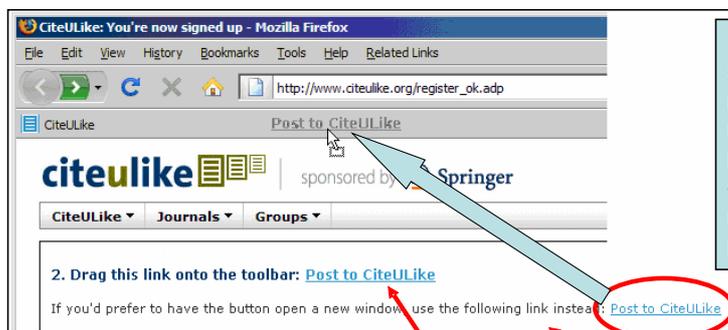
MyCiteULike –
личный раздел данного
пользователя; именно
в этом разделе будет
формироваться
личная база данных
(My Library)

а) Наводим курсор
на текст CiteULike -
открывается меню.



б) Идем по ссылке
Browser Button
и просматриваем
полученную страницу

4. Устанавливаем кнопку Post to CiteULike



Находим на странице ссылку Post to CiteULike, наводим на нее курсор, нажимаем левую клавишу мышки и, не отпуская клавишу, перетаскиваем ссылку на панель инструментов.

Примечание. Которую из двух имеющихся ссылок перетаскивать? Ответ: вторую, с ней будет удобнее работать.

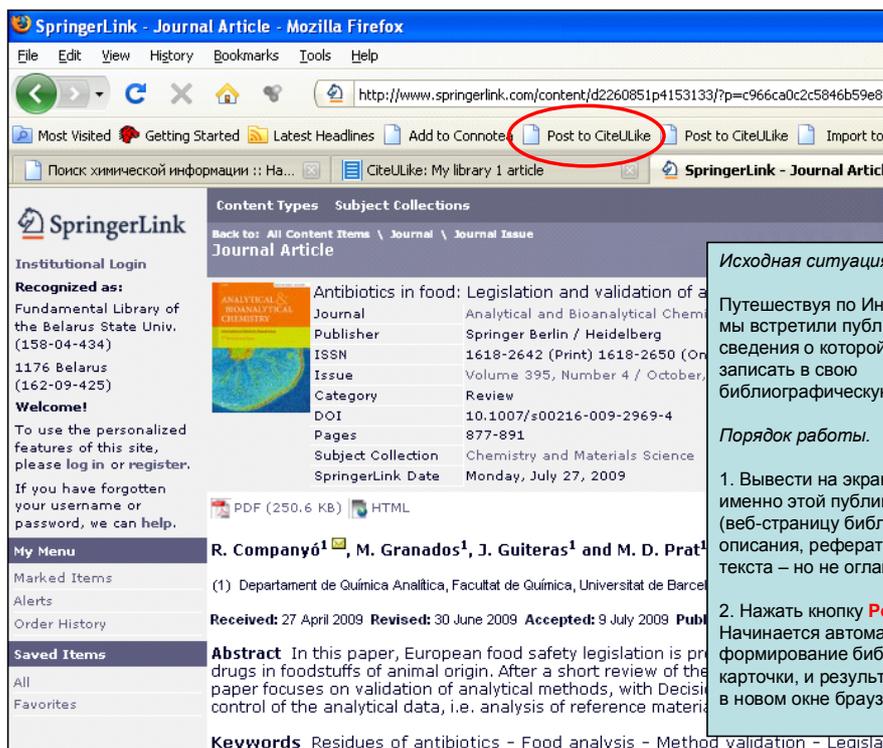
Результат:



Операция выполнена успешно

Заполнение базы данных проиллюстрировано на следующих рисунках:

1. Копирование информации



Исходная ситуация.

Путешествуя по Интернету, мы встретили публикацию, сведения о которой стоило бы записать в свою библиографическую базу данных.

Порядок работы.

1. Вывести на экран веб-страницу именно этой публикации (веб-страницу библиографического описания, реферата или полного текста – но не оглавления журнала).
2. Нажать кнопку **Post to CiteULike**. Начинается автоматическое формирование библиографической карточки, и результат выводится в новом окне браузера.

Контрольная работа

Вся студенческая группа будет коллективно формировать базу данных по тематике "Свойства химических элементов и их соединений".

Ваш химический элемент: ... (задание получите у преподавателя).

Каждому участнику необходимо внести в базу данных по три записи:

- две из журналов двух ведущих издательств (любых из следующего списка: *ACS*, *RSC*, *Elsevier*, *Springer*, *Wiley*),
- одну русскоязычную из ресурсов *eLIBRARY.RU*.

Напоминание:

Тэги следует записывать через пробелы.

Тэги можно записывать только латинскими буквами.

Внимание! Кроме тематических тэгов, вам необходимо будет вписать цифровой идентификационный тэг, по которому можно будет быстро извлечь ваши собственные записи.

Идентификационный тэг сформируем по шаблону:

Номер_группыНомер_студента_по_журналу_преподавателя,
например, *101* или *4104*

(код согласуйте с преподавателем).

6. ПАТЕНТНЫЕ БАЗЫ ДАННЫХ (сокращенная версия)

6.1. Введение

По некоторым оценкам, патенты содержат до 90 % производимой в мире технологической информации, причем лишь 5–10 % этих сведений публикуется в иной научной литературе. Понятно, что патентный поиск является обязательным этапом любого прикладного исследования.

В *WWW* есть несколько источников, предоставляющих достоверные сведения о патентах, опубликованных некоторыми национальными и международными патентными бюро. Мы ознакомимся с сайтом **USPTO Patent Full-Text Databases**, содержащим систематичную патентную информацию о патентах США за очень большой период времени.

Тем, кто желает получить большой объем информации по данной теме, рекомендуем обратиться к имеющемуся на нашем сайте онлайн-методическому руководству "[Патентные базы данных в Интернете](http://www.abc.chemistry.bsu.by/patent/default.htm)" (<http://www.abc.chemistry.bsu.by/patent/default.htm>).

6.2. US Patent and Trademark Office (USPTO)

[Patent Full-Text Databases \(http://patft.uspto.gov/\)](http://patft.uspto.gov/)

Ведомство патентов и торговых марок США — USPTO — предоставляет свободный доступ к американским патентам, опубликованным с 1790 г. по настоящее время.

Сайт содержит две автономные базы данных:

- **USPTO Patent Full-Text and Image Database (PatFT)** — патенты
- **USPTO Patent Application Full-Text and Image Database (AppFT)** — патентные заявки

Каждая из этих двух баз данных имеет собственную поисковую систему. На [Главной странице сайта \(http://patft.uspto.gov/\)](http://patft.uspto.gov/) размещены две группы гиперсвязей к поисковым бланкам:

- **Quick Search** — для простейших заданий.
- **Advanced Search** — для заданий любой степени сложности.
- **Number Search** — только для извлечения патента (патентной заявки) по его/ее регистрационному номеру.

Приемы поисковой работы одинаковы в обеих базах данных.

База данных **USPTO Patent Full-Text and Image Database (PatFT)**

С точки зрения пользователя, патентная база данных **PatFT: Patents** состоит из двух частей:

- **с 1976 г. по настоящее время**. Патенты этого периода хранятся и в текстовом (HTML) формате, и в графическом (TIF). Материал патента распределен по трем десяткам полей, что позволяет проводить прецизионный поиск информации.

Поисковая программа выполняет следующие операции:

- поиск патентов по критерию присутствия/отсутствия термина (комбинации терминов) в отдельном поле (группе полей; патенте в целом);
 - поиск патентов, которые цитируются в извлеченном патенте;
 - поиск патентов, в которых цитируется извлеченный патент.
- **до 1976 г.** Эти патенты хранятся только в графическом формате. Поисковая программа способна извлекать патенты только по их номеру или по классификационному коду.

Познакомимся с самым простым бланком базы данных *PatFT Quick Search* (<http://patft.uspto.gov/netahtml/PTO/search-bool.html>).

Некоторые принципы формулирования поискового задания

- Поисковое задание может состоять либо из одного термина, либо из двух терминов, объединенных оператором (*AND, OR, ANDNOT*).
- Шаблон \$ (*обратите внимание на нестандартность обозначения*) заменяет произвольное количество символов в конце слова.
- Точная фраза обозначается двойными кавычками.
- Поиск по слову может проводиться либо в любом из текстовых полей, либо во всем патенте в целом.

Упражнение

Задача. Извлечь патенты, опубликованные начиная с 1976 года, в названии содержащие слова *фосфат* (*фосфаты*) и *покрытие* (*покрытия*).

Порядок работы. Заполните бланк:

- В графе *Term 1* запишите первый поисковый термин — *phosphate\$* ;
- В меню *Field 1* выберите поле поиска первого термина — *Title* ;
- В графе *Term 2* запишите второй поисковый термин — *coating\$* ;
- В меню *Field 2* выберите поле поиска второго термина — *Title* ;
- В меню выберите оператор, объединяющий оба термина — *AND* ;
- В меню *Select years* выберите временной интервал поиска — *1976 to present*;
- Приступите к поиску — нажмите кнопку *Search*.

Примечание. Если бы поисковое задание состояло из единственного термина, графа *Term 2* осталась бы пустой, и тогда не имело бы значения, какие пункты указаны в меню *Field 2* и *Operator* .

Результаты поиска

На экран выводится список обнаруженных документов, составленный в обратном хронологическом порядке. Гиперсвязь, начинающаяся от регистрационного номера и названия патента, ведет к полному тексту данного патента.

Патент состоит из следующих структурных частей: титульной страницы, патентной формулы (*Claims*) и описания (*Description*).

Титульная страница патента США содержит библиографические

сведения, реферат, а также ссылки на другую информацию:

- В разделе *References Cited* — к патентам-прототипам.
- Гиперсвязь *Referenced By* направлена к патентам, в которых цитируется данный патент.

Вверху и внизу страницы в навигационных группах есть кнопки для вызова поисковых бланков (*Quick, Advanced, Pat Num*), для перехода к следующему или предыдущему патентам списка (*Next, Previous*) и к самому списку (*Hit List*), а также к графическим файлам — отсканированным листам патента (*Images*).

Контрольное задание

Для чего в патентованной жевательной резинке (*chewing gum*) предлагается использовать карбамид (*carbamide* и *urea*)?

Составьте рациональный план поиска ответа на поставленный вопрос.

Извлеките минимальное число релевантных документов.

Изучите структуру извлеченных патентов.

Найдите ответы в нужных местах текста.

Прокомментируйте с точки зрения химика.

7. НОРМАТИВНЫЕ ДОКУМЕНТЫ

7.1. Сертификат безопасности материала (MSDS)

Сертификат безопасности материала (*англ.: Material Safety Data Sheet, MSDS*) — это документ, предназначенный для информирования работников предприятий и персонала аварийных служб о процедурах безопасного обращения с конкретными веществами или смесями веществ. Сертификат безопасности содержит сведения о физических характеристиках (температура плавления, температура кипения, температура воспламенения и т.д.), токсичности, влиянии на здоровье, способах оказания первой помощи, химической активности, условиях хранения, переработки, использовании защитного снаряжения и спецсредств, методах обезвреживания.

Изготовление сертификата входит в обязанности той фирмы, которая производит данный химический товар, поэтому на одно и то же вещество можно обнаружить несколько документов *MSDS*. Они могут различаться объемом, детализацией — но в пунктах, относящихся к технике безопасности, смысловые отличия, как правило, не очень существенны.

Сертификат безопасности материала не является первичным информационным источником; он не может служить в качестве авторитетного справочника физико-химических характеристик веществ. В этой части его следует рассматривать лишь как справочное пособие.

Сертификат безопасности материала содержит заслуживающую доверия информацию в той части, которая описывает опасность вещества, правила работы с ним, приемы обезвреживания. Здесь допущенные ошибки были бы чреваты для авторов сертификата судебными разбирательствами.

В Беларуси и странах СНГ аналогичный документ носит название **Паспорт безопасности (вещества или материала)**.

Согласно Государственному стандарту, в *Паспорт безопасности* должна вноситься только достоверная информация.

Сертификаты безопасности, как правило, обязательно присутствуют на сайтах фирм, производящих химические товары, а также в каталогах продавцов реактивов. Кроме того, в Интернете есть немало количество баз данных, в которых такая информация собрана из разных источников.

Хорошей стартовой точкой при поиске информации об опасности/безопасности вещества является метасайт баз данных:

[Where to find MSDS on the Internet \(http://www.ilpi.com/msds/index.html\)](http://www.ilpi.com/msds/index.html)

Для поиска *MSDS* распространенных веществ, имеющих не очень сложные названия, вполне пригодна поисковая система *Google*, запрос для которой можно составить (на английском языке) в форме:

MSDS название_вещества

Для распространенных веществ ссылки на рекомендуемые сертификаты *MSDS* приводятся в соответствующих статьях в английской версии [Википедии \(http://en.wikipedia.org/\)](http://en.wikipedia.org/).

Задание 1

Несколькими способами найдите несколько сертификатов безопасности фенола (фенол = *phenol*).

Проанализируйте, какую структуру имеет типичный сертификат безопасности и в каком объеме в нем представлены блоки информации.

Оцените степень токсичности фенола.

Что означают следующие аббревиатуры: LDLo, LD50 ?

Что могут означать такие текстовые фрагменты в разделе "Сведения о токсичности" сертификата безопасности материала:

"ORL-HMN LDLO 140 mg/kg",

"ORL-RAT LD50 317 mg/kg" ?

Задание 2

Брусника может храниться долго, и объясняют это тем, что в ней содержится консервант *бензойная кислота*.

Абстрагируемся от всех остальных эффектов и примем, что единственным опасным для организма компонентом брусники является бензойная кислота. Допустимо ли съесть 150 г брусники в один прием?

Сделайте оценочный расчет.

Необходимые справочные данные следует брать в (*а вот где именно — предложите сами*).

Сведения о веществах, представляющих опасность для человека, имеются в специализированных базах данных.

[EINECS — European INventory of Existing Commercial chemical Substances \(http://esis.jrc.ec.europa.eu/\)](http://esis.jrc.ec.europa.eu/) — информация о всех химических товарах, производимых в странах ЕС.

Задание 3

Извлеките из базы данных *EINECS* информацию о феноле.

Обратите внимание на ссылку *IUCLID Chemical Data Sheet*.

Выведите на экран количественные характеристики токсичности фенола. В каком месте документа речь идет о токсичности для человека?

Можно ли использовать базу данных как справочник физико-химических величин?

Можно ли считать этот документ нормативным для стран ЕС?

В чем заключается информационная ценность документов, размещенных в базе данных *EINECS*?

[Hazardous Substances Data Bank — HSDB \(http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB\)](http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB) — информация о токсичности веществ (США).

Задание 4

Извлеките из базы данных *HSDB* информацию о феноле.

В каких параграфах документа приведены количественные характеристики токсичности фенола?

Какова предполагаемая степень достоверности информации, содержащейся в *HSDB*?

Можно ли использовать *HSDB* как справочник физико-химических величин? (проанализируйте параграфы "Температура плавления", "Температура кипения", "рН" и др.)

В чем заключается информационная ценность документов, размещенных в базе данных *HSDB*?

7.2. ГОСТ - Государственный стандарт

ГОСТ (Государственный стандарт) — документ, в котором излагаются утвержденные органами стандартизации требования к качеству определенного объекта.

В нашей стране действуют государственные стандарты Беларуси (СТБ), а также ГОСТы СССР и межгосударственные стандарты СНГ, если они еще не заменены белорусскими национальными стандартами.

В целях идентификации каждый стандарт имеет свое *Обозначение* и *Название*.

В *Обозначении* указывается порядковый номер и год разработки, например:
СТБ 1874-2008 (стандарт Беларуси);
ГОСТ 4245-72 (стандарт СССР);
ГОСТ Р 51210-98 (российский стандарт);
СТБ ГОСТ Р 51212-2001 (стандарт Беларуси на основе российского).

Стандарты — ценные источники научной информации.

В этих документах подробно излагаются тщательно проверенные, воспроизводимые методики выполнения разнообразных операций, с которыми приходится сталкиваться химику.

Например, стандарты на реактивы содержат методики химического анализа вещества и количественного определения примесей.

В Сети имеется много баз данных (главным образом, неофициальных), содержащих советские, российские и межгосударственные стандарты. Правда, только часть из этих источников предоставляет бесплатный доступ к полнотекстовой информации.

Google без проблем обнаруживает стандарты, если запрос сформирован по схеме:

**гост номер_стандарта или
гост слова_из_названия_стандарта**

Обычно список результатов поиска получается большим, и пользователь вынужден разыскивать полнотекстовый документ методом перебора.

Поисковый процесс может оказаться более эффективным, если обратиться к соответствующей базе данных напрямую.

Рассмотрим некоторые ресурсы.

Со списком стандартов, действующих в Беларуси, можно ознакомиться в **Национальном фонде технических нормативных правовых актов Республики Беларусь (ТНПА)**:

[Расширенный поиск ТНПА \(http://www.tnpa.by/PoiskTnpaGosOrgan.php\)](http://www.tnpa.by/PoiskTnpaGosOrgan.php)

База данных содержит полные тексты новейших стандартов Беларуси (правда, пока что не всех).

Контрольное задание 1

Определите, в соответствии с каким стандартом в Беларуси следует определять жесткость питьевой воды.

Примечание. С запросом придется поэкспериментировать, поскольку поиск по словоформам здесь не ведется.

Какова доступность этого документа на сайте ТНПА?

Контрольное задание 2

Определите, в соответствии с каким стандартом в Беларуси следует определять наличие ПАВ в питьевой воде.

Какова доступность этого документа на сайте ТНПА?

Пример полезного и удобного российского информационного источника.

GOST Expert (<http://gostexpert.ru/>) — неофициальная база данных, содержащая 25 тыс. российских, межгосударственных и советских ГОСТов.

Работу можно вести в режимах *Browse* и *Search*.

Упражнение

Режим *Browse*.

а) Перейдите по цепочке гиперсвязей *Промышленность* → *Химическая промышленность* → *Аналитическая химия* → *Аналитическая химия в целом*.

На примере документа "ГОСТ 10067-80 Реактивы. Калий фтористый кислый. Технические условия" изучите типичную структуру стандарта; обратите внимание на описание аналитических операций.

б) Перейдите по цепочкам гиперсвязей *Промышленность* → *Химическая промышленность* → *Неорганические химические вещества и Промышленность* → *Химическая промышленность* → *Органические химические вещества*, извлеките списки стандартов, проанализируйте используемую терминологию.

Сформулируйте для себя правило: какие поисковые термины целесообразно использовать, если требуется обнаружить ГОСТ для заданного вещества?

Контрольное задание 3

Режим *Search*.

а) Из базы данных *GOST Expert* извлеките стандарт, в котором излагаются требования к реактиву — *карбонату натрия*.

Далее мы увидим, что в ГОСТе указываются мельчайшие подробности методики химического анализа.

б) В ходе количественного определения нерастворимых примесей, содержащихся в карбонате натрия, исследуемый раствор необходимо профильтровать.

Как следует правильно подготовить фильтровальную бумагу для этой операции? Обоснуйте целесообразность именно такой методики.

Обобщим и закрепим.

Контрольное задание 4.

Предположим, вы работаете в метрологической лаборатории.

Неожиданно в самый неподходящий момент времени вам привезли две пробы питьевой воды для определения жесткости воды: одну пробу в стеклянной посуде, а другую — в полиэтиленовой.

Есть ли у вас основание отказаться от внеплановой работы под тем предлогом, что клиент неправильно выбрал материал посуды?

В такой ситуации каким документом следовало бы руководствоваться в белорусской лаборатории? в российской лаборатории?

Официальная информация о российских ГОСТах накапливается на сайте Федерального агентства по техническому регулированию и метрологии в базе данных:

[Национальные стандарты РФ \(http://protect.gost.ru/default.aspx\)](http://protect.gost.ru/default.aspx)

Ознакомьтесь с сайтом, после чего выполните задание.

Контрольное задание 5.

В наших магазинах мы нередко встречаем российские продукты.

Какие фосфаты калия разрешено использовать в РФ в качестве пищевых добавок?

7.3. СанПиН — Санитарные правила и нормы

Санитарные нормы, правила и гигиенические нормативы, утвержденные Министерством здравоохранения Республики Беларусь, находятся на сайте Министерства в подразделе:

[Тексты технических нормативных актов \(http://minzdrav.gov.by/ru/static/acts/tehnicheskije/teksty\)](http://minzdrav.gov.by/ru/static/acts/tehnicheskije/teksty)

Самостоятельная работа.

Ознакомьтесь со структурой раздела "База правовых актов".

Выясните, где размещены полные тексты нормативных актов.

Найдите акты, регламентирующие содержание примесей в питьевой воде, в атмосфере помещений.

Контрольное задание.

а) Определите допустимые значения рН питьевой воды.

б) Чему равна величина ПДК фосфатов в питьевой воде?

8. СПРАВОЧНЫЕ БАЗЫ ДАННЫХ

8.1. Введение

Перечни имеющихся в WWW справочных баз данных можно найти на метасайтах и в тематических каталогах в разделах *Databases*. Мы ознакомимся только с некоторыми основными типами такого рода ресурсов.

Примеры исходных пунктов для поиска справочных баз данных:

[Sixty-Four Free Chemistry Databases](http://depth-first.com/articles/2011/10/12/sixty-four-free-chemistry-databases/)
(<http://depth-first.com/articles/2011/10/12/sixty-four-free-chemistry-databases/>)

Адреса и аннотации десятков особо интересных бесплатных баз данных.

[Chemistry - LibGuides at California Institute of Technology](http://libguides.caltech.edu/chemistry)
(<http://libguides.caltech.edu/chemistry>)

Адреса и аннотации на сайте библиотеки Калифорнийского технологического института.

[Top: Science: Chemistry: Chemical Databases](http://dmoz.org/Science/Chemistry/Chemical_Databases)
(http://dmoz.org/Science/Chemistry/Chemical_Databases/)

Раздел *Chemical Databases* тематического каталога **Open Directory**.

8.2. Базы данных для работы в режиме Browse

Примером электронного справочника, в некоторой степени сохраняющего традиционную структуру печатного издания, является сайт [Номенклатура ИЮПАК](http://www.chem.qmul.ac.uk/iupac/) (<http://www.chem.qmul.ac.uk/iupac/>).

Текстовая часть справочника разбита на разделы и подразделы, содержащие привычные оглавления. Особенности именно электронного издания проявляются в развитой гипертекстовой структуре информационного массива.

Упражнение.

Ознакомьтесь с главой *Glossary of Organic Class Names*.

Более специфичен сайт [WebElements](http://www.webelements.com/) (<http://www.webelements.com/>) (*Свойства элементов и их соединений*).

Этот справочник содержит большое количества иллюстративного материала, в том числе мультимедийного.

Упражнение.

Ознакомьтесь со строением и содержанием раздела "Фосфор".

8.3. Базы данных для работы в режиме Search

Вторая группа - это истинно электронные базы данных, имеющие свои поисковые программы и, в основном, не ориентированные на работу в режиме *Browse*.

Примеры:

[NIST Chemistry WebBook \(http://webbook.nist.gov/chemistry/\)](http://webbook.nist.gov/chemistry/)

Справочник по термодинамическим и термофизическим свойствам веществ и ионов. Материал находится на сайте NIST (*The National Institute of Standards and Technology*), США.

Упражнение.

Перейдите к поисковому бланку *Name* (гиперсвязь находится в блоке *Search Options*) и извлеките имеющиеся в базе данных сведения о метане (*methane*).

[Термические константы веществ](http://www.chem.msu.ru/cgi-bin/tkv.pl?show=welcome.html/welcome.html)

[\(http://www.chem.msu.ru/cgi-bin/tkv.pl?show=welcome.html/welcome.html\)](http://www.chem.msu.ru/cgi-bin/tkv.pl?show=welcome.html/welcome.html)

Электронная версия 10-томного одноименного издания, публиковавшегося в 1965–1982 гг.

Упражнение.

Извлеките имеющиеся в базе данных сведения о метане.

[ChemIDplus \(http://chem.sis.nlm.nih.gov/chemidplus/\)](http://chem.sis.nlm.nih.gov/chemidplus/)

База данных, созданная Национальной медицинской библиотекой США (*National Library of Medicine, NLM*), содержит сведения почти о 400 тыс. органических и (некоторых) неорганических соединений (тривиальные и торговые названия, структура, токсичность, основные физические свойства, ссылки на другие базы данных *NLM*, в которых содержится информация о данном веществе).

Упражнение.

Извлеките имеющиеся в базе данных сведения об ацетоне (*acetone*). Поиск проводите по названию вещества — для этого воспользуйтесь разделом *Substance Identification* поискового бланка.

Тщательно просмотрите таблицу, в которой характеризуются физические свойства ацетона; одна из деталей должна привлечь ваше внимание.

[ChemSpider \(http://www.chemspider.com/\)](http://www.chemspider.com/)

ChemSpider — бесплатная база данных, в которой собраны сведения о свойствах более 26 миллионов химических веществ, главным образом, органических.

ChemSpider — структурная база данных. По замыслу создателей, именно структура химического вещества является первичным идентификатором в конструкции всего информационного массива; все остальные идентификаторы (название, регистрационные номера и т. д.) здесь считаются вторичными и поэтому не всегда точно соответствуют определенному химическому объекту.

Пример несоответствия, на которое мы не всегда обращаем внимание. Термин "поваренная соль" — это не только тривиальное название вещества хлорида натрия, но и название товара, в котором вещество $NaCl$ является основным — но не единственным — компонентом.

Материал в *ChemSpider* аккумулируется из разных источников: из других баз данных — больших и маленьких, научных и учебных, коммерческих и бесплатных,— а также из первичной литературы — и рецензируемой, и не прошедшей рецензирование. В результате сведения о веществе здесь представляют собой скорее компиляцию ранее опубликованной информации, а не массив выверенных численных данных.

Сильная сторона *ChemSpider* — возможность поиска информации по заданной структуре вещества; этим мы займемся во втором семестре.

Теперь же мы проведем текстовый поиск по названию вещества, а также ознакомимся со структурой информационного массива.

В запросе допускается использовать систематическое название вещества, торговое, тривиальное, регистрационный номер из различных баз данных, а также название смеси, в которой данное вещество является одним из основных компонентов.

Примечание. Варианты названий вещества здесь некорректно именуется синонимами (кстати, такая некорректность широко распространена; она присутствует даже в ресурсах *Chemical Abstracts Service*).

Упражнение.

Извлеките имеющиеся в *ChemSpider* сведения о феноле (*phenol*).

Изучите структуру страницы этого химического вещества.

Контрольное задание.

Найдите **достоверное** численное значение физико-химического параметра вещества (индивидуальное задание получите у преподавателя).

Bibliographic Record

Type: Book (Electronic)

Title: Text Search for Scholarly Chemical Information on the Internet : Manual

Author: Rahoisha, Aliaksandr A.

Year: 2011

Description: 64 p. : ill.

Available online: http://www.abc.chemistry.bsu.by/lit/Rahoisha_2011.pdf

Subject: Chemistry

Keywords: chemistry, chemical information, information retrieval, text query

Abstract: Text search is still the main method of obtaining chemical information from online databases. The manual introduces students into the vast domain of scholarly Web resources and guides them in designing appropriate text queries for the effective retrieval of the relevant material. The list of objects under study includes but is not limited to search engines, journals, theses, books, patents, and national standards.

Institution: Belarusian State University

Country: Belarus

Language: Russian

Address for correspondence: Aliaksandr Rahoisha, Chemistry Department, BSU, pr. Nezalezhnasci 4, Minsk 220030, Belarus.

E-mail: arahoisha@gmail.com