

**С.И. Двойников, И.А. Фомушкина,
Э.О. Костюкова, В.И. Фомушкин**

ОРГАНИЗАЦИОННО- АНАЛИТИЧЕСКАЯ ДЕЯТЕЛЬНОСТЬ

Под редакцией С.И. Двойникова

УЧЕБНИК ДЛЯ МЕДИЦИНСКИХ УЧИЛИЩ И КОЛЛЕДЖЕЙ

Министерство образования и науки РФ

Рекомендовано ГБОУ ВПО «Первый Московский государственный
медицинский университет имени И.М. Сеченова» в качестве
учебника для студентов образовательных организаций среднего
профессионального образования, обучающихся по специальности
060101 «Лечебное дело» в рамках профессионального модуля
«Организационно-аналитическая деятельность»

Регистрационный номер рецензии 027 от 03 февраля 2015 г.
ФГАУ «Федеральный институт развития образования»



Москва
ИЗДАТЕЛЬСКАЯ ГРУППА
«ГЭОТАР-Медиа»
2017

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ

9.1. ТЕХНОЛОГИИ ПОИСКА ТЕМАТИЧЕСКОЙ (ПРОФЕССИОНАЛЬНОЙ) ИНФОРМАЦИИ В СЕТИ INTERNET

9.1.1. ПРИНЦИПЫ КОНТЕКСТНОГО ПОИСКА

В современном мире поток информации ежедневно увеличивается лавинообразно. Электронные версии нормативно-правовых документов, методических рекомендаций и указаний, последние номера профессиональных журналов стали доступны медицинским работникам в самых отдаленных регионах при наличии доступа к сети Интернет. Однако при кажущейся простоте получения информации ее качество и возможность использования в работе напрямую зависят от полноты и адекватности формулировки контекстного поиска.

Контекстный поиск — запрос информации в компьютерной системе с использованием набора ключевых слов. Существуют несколько видов поиска (рис. 9.1).

Для эффективного поиска современные алгоритмы заранее формируют **полнотекстовый индекс** — словарь, в котором перечислены все слова и указано, в каких местах они встречаются. При наличии такого индекса достаточно осуществить поиск нужных слов в нем, и тогда сразу же будет получен список документов, в которых они встречаются. Технологии полнотекстового поиска, с использованием индек-

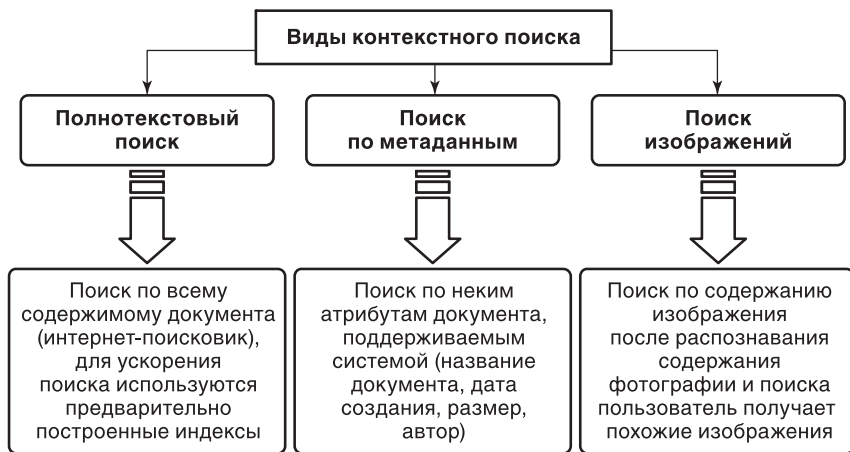


Рис. 9.1. Классификация видов поиска

сирования, широко применяют в современных поисковых системах сети Интернет.

Поисковая система представляет программно-аппаратный комплекс с веб-интерфейсом (сайт, на котором размещен интерфейс системы). Программной частью поисковой системы служит **поисковая машина (поисковый движок)** — комплекс программ, который обеспечивает функциональность поисковой системы. Чаще всего структуру поискового движка считают коммерческой тайной компании — разработчика поисковой системы.

Индексирование в поисковых системах — процесс добавления сведений о сайте роботом поисковой машины в базу данных, впоследствии используемую для поиска информации на проиндексированных сайтах. **Поисковый робот** — программа, служащая составной частью поисковой системы и предназначенная для перебора страниц Интернета с целью занесения информации о них в базу данных поисковика. Поисковый робот анализирует содержимое страницы, сохраняет его в некотором специальном виде на сервере поисковой машины, которой принадлежит, и отправляется по ссылкам на следующие страницы. Владельцы поисковых машин нередко ограничивают глубину проникновения робота внутрь сайта и максимальный размер сканируемого текста. Именно поэтому чересчур большие сайты могут оказаться не полностью проиндексированными поисковой машиной.

Почти все поисковые машины в сети Интернет используют системы контекстной рекламы для получения прибыли. Эти системы позволяют размещать рекламу как на страницах с результатами поиска

по определенным ключевым словам, так и на сайтах, установивших блоки контекстной рекламы на своих страницах.

Большинство поисковых систем ищут информацию на сайтах Всемирной паутины. Однако существуют также системы, способные искать файлы на FTP-серверах, товары в интернет-магазинах и др.

Улучшение поиска — одна из приоритетных задач современного Интернета. Существует ряд перспективных направлений развития поиска (например, поиск в графических изображениях, аудио- и видеофайлах и др.), однако прогресс в этом направлении упирается в фундаментальную проблему создания искусственного интеллекта.

9.1.2. ПОИСКОВЫЕ СИСТЕМЫ РУССКОЯЗЫЧНОГО ПРОСТРАНСТВА ИНТЕРНЕТА

Для удовлетворения запросов миллионов пользователей существует достаточно большое количество поисковых систем, как «всеязычных», так и русскоязычных. Международная аналитическая компания *NetMarketShare* ежемесячно составляет рейтинг самых популярных поисковиков. По данным 2013 г., места в непростой гонке распределились следующим образом:

- *Google-Global* — 68,1%;
- *Baidu* — 18,4%;
- *Yahoo-Global* — 6,4%;
- *Bing* — 5,59%.

При всем многообразии выбора неоспоримым лидером в глобальном интернет-поиске считают *Google*. Этот гигант-поисковик обрабатывает примерно 118 млрд запросов ежемесячно и может находить информацию на 195 языках.

Русскоязычные поисковые системы, в отличие от «всеязычных», индексируют ресурсы, где доминирует русский язык. Именно поэтому и расстановка сил в русскоязычном пространстве Интернета (Рунета) несколько иная. По данным *LiveInternet.ru* (www.liveinternet.ru/stat/ru/), в июле 2013 г. поисковые предпочтения российских пользователей выглядели так:

- Яндекс — 53,8%;
- *Google* — 34,2%;
- поиск *Mail.ru* — 9,4%;
- Рамблер — 1,2%;
- остальные поисковики — менее 1%.

Подобный разброс результатов рейтинга и их ежемесячные изменения связаны с постоянными нововведениями конкурирующих поисковых систем, заманчивыми предложениями, облегчающими поиск.

Однако основным критерием качества работы поисковой машины служит **релевантность** — степень соответствия запроса и найденного, т.е. уместность результата. В итоге при поиске первыми в открываемом списке оказываются те веб-страницы, которые наиболее соответствуют запросу пользователя. Однако каждая из поисковых систем использует различные критерии ранжирования документов, т.е. результаты поиска посредством Яндекс и Google по одному и тому же запросу могут существенно отличаться. Естественно, что пользователь будет больше доверять тому поисковику, который представил наиболее релевантные результаты.

9.1.3. ПРИНЦИПЫ ФОРМИРОВАНИЯ ЗАПРОСОВ ДЛЯ ПОИСКОВОЙ СИСТЕМЫ

Поисковую систему можно образно представить как огромную базу ключевых слов, каждое из которых связано с *Web*-страницами, на которых его встречали. Вся совокупность подобных связей по своей сути и является Всемирной паутиной (*вольный перевод английского словосочетания World Wide Web*). Технология WWW позволяет создавать ссылки (их также называют гиперссылками), которые реализуют переходы на веб-страницах.

Вводя в поле поисковой системы ключевое слово или фразу, пользователь посылает поисковой системе запрос. Результаты поиска выдаются в виде списка адресов *Web*-страниц, на которых встретились эти слова.

Чем лучше составлен запрос к поисковой машине, тем меньше общее количество найденных документов и больше из них документов, соответствующих тому, что требуется (релевантные страницы). Если запрос некорректен или составлен неграмотно, то поисковая машина может выдать несколько миллионов результатов.

Для построения оптимального запроса рекомендуют придерживаться нескольких правил:

- Используйте несколько ключевых слов одновременно.
- Выбирайте только самые важные ключевые слова, касающиеся рассматриваемой темы. Считают, что эффективный запрос может сформироваться из 3–4 ключевых слов, но их может быть и больше.
- Используйте функции расширенного поиска поисковых систем, если вам нужны уточнения по датам, географии, языку, формату файла.
- Набирайте в запросе ключевые слова преимущественно строчными буквами. Заглавные буквы необходимы для точного

отграничения в поиске имен собственных, наименований торговых марок, предприятий, географических названий.

- Поисковый запрос необходимо оформлять с соблюдением орфографических норм. Если поисковая система поправляет ваше написание, обратите на это внимание, уточните запрос с учетом орфографии, иначе поисковик выдаст миллионы ненужных страниц с такими же ошибками.
- Используйте разные поисковые системы. Формируйте запрос в разных поисковых системах.
- Не прекращайте поиск на первой десятке сайтов.
- Профессиональная публикация должна иметь автора, ссылки на использованные источники. Доверяйте данным с проверенных профессиональных сайтов.
- Используйте так называемые операторы при составлении запроса (табл. 9.1).

Таблица 9.1. Виды операторов для формирования поисковых запросов для Yandex

Оператор	Использование	Пример
«...» — кавычки	Поиск словосочетания, предложения без изменения порядка слов	«Лечебная физкультура»
! — восклицательный знак	Из поиска будут исключены слова других морфологических форм (употребляемые в других падежах, числе и др.)	!диспансеризация
— вертикальная черта	Поиск всех синонимов (расширение запроса)	Дети младенец ребенок (задает поиск документов, содержащих хотя бы одно из перечисленных слов)
+ — знак «плюс»	Обязательное присутствие слов в найденных документах	+ лечебная гимнастика физкультура
«—» — знак «минус»	Исключение слова из результата поиска	+ аугментин + (инструкция по применению) — аптека — цена (будут показаны страницы только с инструкцией по применению, без цен и предложений аптек)