1. Опишите процесс составление индекса поисковой машиной.

Индекс – Своеобразный словарь, в котором основы упорядочены по алфавиту, а при каждой основе записано, с какой страницы она взята ( номер страницы) и на каком месте на этой странице данная основа стояла (номер вхождения)

1. Почему индекс поисковой машины называется инвертированным?

Если в обычном тексте мы идем от страницы к словам, то в индексе поисковая машина идет от слов к страницам. Поэтому **индекс поисковой машины называется инвертированным, или инверсным, то есть обращенным, перевернутым**.

1. Какие поисковые системы хранят проиндексированные страницы полностью?

Например, Google имеет у себя текстовую копию всего мирового Интернета (в том объеме, до какого смог добраться его «паук»), а «Яндекс» — копию всего Рунета.

1. Нужно ли индексировать стоп-слова и цифры?

Из-за того, что пользователи довольно часто запрашивают стоп-слова и цифры, их нужно индексировать

1. Как поисковая система справляется с высокой флективностью русского языка?

Решением проблемы изменчивости слов является хранение только основы слов, а окончания и чередования – отбрасывать и хранить отдельно. Если все формы слова сводятся к его основе на этапе создания индекса, то потом, при задании любой словоформы слова в запросе, она также будет сведена к оснве и будут найдены все вхождения этой основы в тексте

1. На какие категории можно разделить многообразие запросов пользователя?

* Навигационные – найти конкретное место в сети;
* Информационные – поиск информации;
* Транзакционные – купить, заказать, скачать;
* Нечеткие – не относящиеся ни к одной из вышеперечисленных категорий.

1. Какие факторы оказывают значительное влияние на индексацию страниц сайта и дальнейшее ранжирование ссылок на документы в результатах поиска на поисковых системах?

Адреса страниц

Работа серверных скриптов

1. Какой вид должны иметь адресе страниц для лучшего поискового ранжирования?

Если необходимо использовать в имени страницы несколько слов, то их можно соединить: Написать слитно, разделить дефисом или знаком подчеркивания. Если несколько слов разделены знаком подчеркивания, то робот увидит одно большое неизвестное ему слово. Так что лучше называть через дефис (realty-broker-service.html).

1. Какими способами можно влиять на индексацию сайта?

Для того, чтобы максимально ускорить индексацию сайта, необходимо правильно настроить взаимодействие сайта с поисковыми работами и предельно четко показывать, что нужно сканировать

1. Составьте план действий на сайте перед добавлением его в поисковые машины используя <http://www.ashmanov.com/pap/ivsprep.phtml>.

* Настроить титульную фразу с помощью тега <title>
* Описать страницу с помощью тега <meta name = “description” content=”>
* Настроить позиции найденных слов в титульной фразе
* Настроить «вес» найденного слова и близость его к началу документа
* Настроить наличие найденных слов в заголовках (h1, h2 и т.д.) и фрагментах, выделенных жирным шрифтом (b, strong)
* Настроить близость найденных слов друг к другу (компактность цитаты)
* По возможности распространить ссылку на свою страницу на других страницах
* Настроить специальный метатег <META name=”ROBOTS”> и файл исключений robot.txt