



Web Effector



Факторы ранжирования Яндекса

Сергей ЛЮДКЕВИЧ, начальник отдела аналитики

ФОРМУЛЫ РЕЛЕВАНТНОСТИ



Отдельные функции релевантности:

- Москва
- Санкт-Петербург + Екатеринбург + 16 городов России релиза «Арзамас+16»: Новосибирск, Самара, Воронеж и др.
- Общероссийская + все остальные города России + Украина + Белоруссия + Казахстан + остальные зарубежные страны + выдача без учета региона (Земля)

ОБЩАЯ ИНФОРМАЦИЯ



Число факторов – 420+ (Ден Расковалов, руководитель отдела качества поиска Яндекса, апрель 2010, КИБ+РИФ)

В формуле используются комбинированные факторы – мономы (Илья Сегалович, дискуссии на ya.ru и habrahabr.ru: *«В данном случае мы нашли ошибку в комбинированном факторе (мономе), сочетающем анализ текста и ссылок, внесли в него исправление...»*)

КЛАССИФИКАЦИЯ ФАКТОРОВ



По виду зависимости от запроса

1. Статические (не зависящие от запроса)
2. Динамические (зависящие от запроса)
3. Запросные (характеристики запроса)

По виду зависимости от документа

1. Документозависимые
2. Сайтозависимые

КЛАССИФИКАЦИЯ ФАКТОРОВ



По характеру принимаемых значений

1. Бинарные
2. Непрерывные

По источнику

1. Внутренние
2. Внешние
3. Пользовательские (поведенческие)
4. Запросные

ЗАПРОСНЫЕ ФАКТОРЫ



- длина запроса в словах;
 - язык запроса;
 - популярность запроса;
 - коммерциализированность запроса;
 - порнографичность запроса
- и др.

ВНУТРЕННИЕ ФАКТОРЫ



1. Текстовые
2. Свойства документа/сайта
3. Антиспамовые

Антиспамовые

- содержит ли попандеры или кликандеры;
 - естественность текста;
 - уникальность контента
- и др.

ТЕКСТОВЫЕ ФАКТОРЫ



Зоны документа:

- доменное имя;
- адрес страницы;
- текст документа;
- тег title;
- элементы оформления текста.

Учет характера вхождений в текст

- точное вхождение;
- лемма;
- СИНОНИМЫ.

СВОЙСТВА ДОКУМЕНТА/САЙТА



- возраст документа (дата первой индексации);
 - тип документа;
 - язык документа;
 - доменная зона;
 - это важный сайт (Wikipedia);
 - это главная страница сайта;
 - близость тематики документа/сайта и запроса;
 - совпадение региона выдачи и сайта;
 - число страниц на сайте;
 - порнографичность контента
- и др.

ССЫЛОЧНЫЕ ФАКТОРЫ



- статические (не зависящие от запроса);
- динамические (зависящие от запроса).

Ссылочные статические

- PageRank;
 - ТИЦ;
 - количество ссылок на документ
- и др.

ССЫЛОЧНЫЕ ДИНАМИЧЕСКИЕ ФАКТОРЫ



- региональность (совпадение региона выдачи и донора)
 - тематичность (близость тематик запроса и донора)
 - «продажность» ссылок
 - возраст ссылок
 - качество ссылок (деление ссылок на «хорошие» и «плохие»)
- и др.

ПОЛЬЗОВАТЕЛЬСКИЕ ФАКТОРЫ → → → → →



- посещаемость сайта;
 - кликабельность документа в выдаче;
 - поведение пользователя на сайте
- и др.

Спасибо за внимание! **Web Effector** 

Для продолжения темы посетите

www.webeffector.ru

115280, Россия, Москва,
ул. Ленинская Слобода, д. 19, стр. 1, бизнес-
центр «Омега Плаза»

Телефон: (495) 772-97-91 (многоканальный)