

Общие правила

Базовой единицей оценки является оценка рел+. Релевантный ответ предоставляет решение пользовательской задачи (бывают запросы, для которых задача, по сути, одна, а бывают запросы, для которых существует спектр задач или которым соответствует несколько объектов поиска).

Если страница позволяет решить сразу несколько задач (для запросов с несколькими потребностями), информации гораздо больше, чем на среднего уровня релевантной странице, она лучше структурирована, - такой странице стоит поставить оценку полезный. Когда наоборот, страница не полностью решает задачу, слишком сужает её или наоборот, расширяет, если страница посвящена малозначительному объекту - тогда оценка должна быть ниже, это рел- ответ.

Витальный ответ - это официальный ответ на вопрос, он несколько отличается от прочей шкалы. Для многих запросов витального ответа не может существовать в принципе, потому что нет официального владельца информации по теме. Обычно определить витальный довольно просто, но иногда бывают и более запутанные случаи.

Также в шкале присутствуют две близких по смыслу оценки нерел и не про то. Нерелевантный - это документ, который в целом соответствует теме запроса, но на него не отвечает. Не про то - это оценка для случаев, когда документ не просто не отвечает на запрос, но вообще относится к совершенно другой тематике.

Помимо релевантности есть ещё оценка спама. Для документов, где присутствуют признаки спама, надо поставить оценку "спам", остальные - "не спам".

Документам с оценкой рел+ и полезный также надо ставить срок устаревания (чаще всего - никогда, но некоторые документы релевантны запросу в момент оценки, а со временем теряют релевантность).

Философия оценки

Во время выполнения заданий у вас могут возникнуть разные вопросы по работе, например: в какой части инструкции искать правила к моему запросу? Как понять, где граница между релевантным и нерелевантным ответом? Может ли тут быть витальный ответ? Что делать, если задание в корне противоречит здравому смыслу? Как не скатиться в формализм и не забыть про здравый смысл? Как не потерять веру в человечество после всего этого?

В первую очередь важно понимать, что автоматизировать систему оценки невозможно, требуются именно человеческие оценки, а не ответ от робота. С другой стороны, оценки разных людей (экспертов) должны быть согласованы и проставлены по единой схеме, иначе вместо оценки получится сомнительного качества опрос общественного мнения.

Чтобы экспертные оценки обладали достаточным качеством, была разработана инструкция ассессоров, которая постоянно обновляется и улучшается. Предусмотреть все возможные комбинации запрос-ответ невозможно, но самые распространённые варианты в ней описаны. Если ваш случай однозначно попадает в описание из инструкции, надо оценивать страницы из задания согласно этому описанию. Если же прямого ответа на свой вопрос в инструкции вы не нашли, нужно ориентироваться на здравый смысл.

Часто запросы бывают однозначные сами по себе, например [бесплатные новые игры винкс ТОЛЬКО БРОДИЛКИ винкс только про винкс и больше не чего только бродилки]. Логично

предположить, что если ответ не содержит игр-бродилок про Winx, релевантность у него нулевая, так что формальный подход к запросам такого рода вполне уместен. Или [как раки относятся к своей прошлой любви?]: тут тоже подойдут только страницы, на которых рассматривается отношение людей, родившихся под знаком Рака, к своим бывшим вторым половинкам. Прочие знаки зодиака и иные психологические особенности Раков совершенно точно не соответствуют запросу. Предположение о том, что речь идёт о членистоногих животных, хотя и забавно, но лишено разумных оснований. Многие запросы сформулированы более общо и им соответствуют несколько типов ответов, например [мангалы из кирпича своими руками]. К такому запросу подойдут тексты (желательно с картинками) с описанием процесса или видеоуроки.

Определить, какая именно информация больше нужна помогает не только здравый смысл, но и статистика. Например, логично предположить, что если что-то можно скачать бесплатно, то предложение покупки этого объекта гораздо менее интересно людям, и статистика это подтверждает. Самые типичные запросы с несколькими разными релевантными ответами это запросы, которые состоят только из названия песни или фильма, имени известного человека или коллектива. Также иногда попадаются запросы, для которых существует несколько подходящих объектов. Без статистики определить, какие именно объекты релевантны, может быть довольно затруднительно. В таких случаях стоит ориентироваться на аннотацию. Чтобы понять, может ли у запроса вообще быть витальный ответ, надо ответить на два вопроса: является ли запрос навигационным (то есть пользователю, скорее всего, нужна единственная страница), например [вконтакте], [скачать песню кино перемен с зайцев нет]? Если запрос навигационный, то та самая главная искомая страница будет витальной. Если нет, то важно понимать, есть ли объекта поиска официальный источник или правообладатель? Если объект поиска - товар, объект авторского права (песня, книга, фильм, программа,..), закон или у объекта поиска существует официальный сайт, то витальный ответ в принципе возможен. Во всех прочих случаях витального ответа не бывает. Например, для запроса [погода в томске сегодня] не может быть витального, поскольку не существует производителя или правообладателя погоды. С другой стороны, для запроса [погода в томске сегодня гисметео] витальный существует, потому что это навигационный запрос, пользователя интересует не просто информация о погоде, но информация на определённом ресурсе.

Определить, какой ответ релевантный, а какой - нет, в общем случае несложно. Но иногда возникают непредвиденные ситуации. Предугадать всё заранее невозможно, поэтому если вы столкнулись с ситуацией, для которой нет явного правила, постарайтесь найти разумную аналогию из инструкции. Если и этого не получается, надо действовать исходя из здравого смысла. Важно удерживать баланс между формальным подходом и субъективным мнением по ситуации. Иногда это не получается, но расстраиваться по этому поводу не стоит: работать со стопроцентной точностью невозможно, какие-то ошибки будут всегда.

Иногда непредвиденные сложности возникают в простых, казалось бы, ситуациях. Например, легко представить себе, какие картинки релевантны запросу [фото белой собаки], а какие - нет. Если на фотографии изображена белая собака - фото релевантно, а если, скажем, чёрная собака или белая кошка - то точно нет. Что делать, если собак несколько, а белая только одна? Ну это вроде несложно, релевантность ниже, чем в первом случае, но не ноль всё-таки, белая собака же есть. А что делать, если у собаки маленькое чёрное пятнышко на лапе? Она формально не совсем белая, но так ли это важно? А что делать, если она какая-то светлая, но не особо понятно, это всё ещё белая или это уже другой цвет? Крайние точки тут несложно обозначить: если собака - далматин, она уже не белая, а вполне пятнистая. Но одно маленькое пятно, например, не достаточно "портит" белую собаку. Но где провести границу? Считать площадь не белого? Может, учитывать местоположение пятен (например, пятно на лапе не важно, а на голове - важно) или их количество, а не площадь? В какой момент светлая

шерсть перестаёт называться белой? Внимательный читатель уже наверняка понял, что все эти вопросы - верный путь себя запутать и поставить вообще бог знает что вместо адекватной оценки. В таких случаях надо полагаться на здравый смысл, а формализм только навредит.

В целом, стоит помнить, что правила в инструкции создавались на основе здравого смысла и известной статистики, причём в них описывается максимально общий подход, то есть правила написаны таким образом, чтобы составленная на их основании выдача поисковой системы удовлетворила потребности максимально возможного числа людей. Угодить совсем уж всем и сразу кажется невозможным, но мы очень стараемся подобраться к этому идеалу. Поэтому оценивайте страницы, исходя из инструкции, но не доводите следование ей до абсурда.

May the force be with you!

Товарный запрос

(Релевантность)

Если запрос товарный, то важно обращать внимание не только на содержание страниц, но и на регионы магазинов и сайта производителя. Рассмотрим, для примера, запрос [нокиа п 86] //(Томск)//. Запрос без уточнения, какого рода информация об этом телефоне требуется, поэтому релевантных ответов будет много.

Поскольку у товара есть производитель, страница с сайта производителя про этот телефон будет витальной:

http://www.nokia.ru/products/phones/nokia_n86_8mp/specifications

Данная модель телефона больше не выпускается, поэтому на сайте производителя осталась только такая страница. Заглавная страница раздела о телефоне, который ещё выпускается, выглядит чуть иначе.

По общему запросу будут релевантны хорошие обзоры телефона:

<http://zoom.cnews.ru/publication/item/17880>

На этой странице представлен качественный обзор (много текста, иллюстрации), есть комментарии пользователей. Подобная страница заслуживает оценки рел+. Пример ещё одного хорошего развёрнутого обзора:

<http://zoom.cnews.ru/publication/item/17880>

<http://www.mobile-review.com/review/nokia-n86.shtml>

По общим запросам магазины также могут быть релевантны.

www.sotmarket.ru/product/nokia_n86_8mp.html

Здесь предлагается та самая модель телефона, однако регион магазина не совпадает регионом запроса (пользователя). Надо проверить, как обстоят дела с доставкой в Томск.

Во-первых, имеет смысл проверить наличие Томска в поле "ваш регион":
Список довольно большой, но Томска там нет. Надо посмотреть, что пишут в разделе "Доставка и оплата":

По ссылке попадаем на страницу, посвящённую доставке

<http://www.sotmarket.ru/showparam/magaz-deliver.html>.

Там про Томск ничего явно не прописано, но можно уточнить: Видно, что доставка в регион пользователя есть и она платная: Таким образом, карточка товара (урл из задания) заслуживает оценки рел-, поскольку магазин находится за пределами региона пользователя и есть платная доставка. Положительная оценка будет только у магазина из региона пользователя (или из другого региона, в котором есть бесплатная доставка в нужный). По товарным запросам иногда попадаете Яндекс Маркет. Регион у него определяется как ваш, то есть чтобы понять, что же видел пользователь из Томска (в данном случае), надо поменять регион:

В Томске, оказывается, купить телефон не предлагают, но характеристики, фотографии, а главное, большая коллекция отзывов никуда не пропадают: Таким образом, маркету тоже стоит поставить рел+.

Могут попасться страницы, содержащие только отзывы, например:

http://www.e-katalog.ru/opinion_ie.php?resolved_Name=NOKIA-N86

На этой странице представлена очень большая подборка отзывов пользователей (экранов десять), такая страница также заслуживает оценки рел+.

Мультиобъектные запросы

(Релевантность)

Иногда попадают запросы, которым соответствуют несколько значений. Например, слово запроса - омоним и само по себе многозначно. Или есть несколько разных объектов, названных одинаково и все их пользователи достаточно активно ищут. Как правило, все эти значения описаны в аннотации. Если вам попался какой-то странный объект, вроде бы релевантный запросу, но не описанный в аннотации и вы не понимаете, как его оценивать, смело пишите разметчику. В целом, если страница заслуживает положительной оценки для какого-то одного из указанных значений запроса, значит, так её и надо оценивать.

Посмотрим, для примера, на запрос [перехват]. В аннотации сказано, что запросу соответствуют несколько объектов: фильм, системы и программы с таким названием, передачи на радио и ТВ с таким названием (витального при этом нет).

<http://www.autoreview.ru/archive/2008/16/vzлом/>

На странице представлен обзор противоугонного устройства "Перехват". Это один из возможных объектов, причём обзор хороший (и не зависит от региона), так что эта страница заслуживает оценки рел+.

На следующей странице мы видим уже другой объект из возможных, передачу на радио с названием "Перехват":

<http://echo.msk.ru/programs/interceotion/758209-echo/>

На этой странице представлена передача за определённое число, это только один выпуск еженедельной Передачи. Такому урлу стоит поставить рел- или даже не рел, если страница совсем уж старая. Положительно следует оценить раздел сайта "Эха Москвы", посвящённый этой передаче:

<http://www.echo.msk.ru/programs/interception/>

На этой странице представлен архив передач, самую последнюю можно прослушать прямо отсюда, а к более старым выпускам - перейти по ссылкам. Эта страница уже будет рел+.

<http://www.bnti.ru/showart.asp?lvl=&aid=66>

На странице описано несколько способов перехвата, статья большая (но несколько экранов текста мелким шрифтом неинтересно вставлять в скриншот), такую статью также уместно оценить положительно.

Помимо устройств это ещё и фильм:

<http://www.fast-torrent.ru/film/zhestokij-zahvat.html>

На этой странице есть совсем уж краткая информация, совершенно не содержательные комментарии пользователей и возможность скачать фильм (через торрент). Вот за то, что скачать всё-таки можно, странице стоит поставить положительную оценку.

Мультиинтентные запросы

По многим запросам релевантными могут быть разнообразные типы ответов. Например, для общих запросов про музыкантов и их творения, фильмы, товары. Рассмотрим запрос [белое солнце пустыни]. Это запрос про кинофильм, в нём не указан тип информации, который пользователь хотел найти. Поэтому здесь возможно несколько релевантных типов ответов. Например, хорошая информация про сам фильм, как часто бывает на википедии:

http://ru.wikipedia.org/wiki/WhiteSun_of_the_Desert

Страницы, на которых можно посмотреть фильм по общему запросу заслуживают положительной оценки:

<http://kinobanda.net/modules.php?name=News&file=view&news%20id=5118>

Для всех таких страниц надо проверять, что видео играет (конечно, смотреть фильм целиком не требуется).

Может попасться страница, на которой есть ссылка на скачивание фильма, например:

<http://vagonkino.ru/fighters/2170-beloe-solnce-pustvni-1970-dvdrip.html>

Видео не играет, но есть ссылка на скачивание. Перед тем, как поставить оценку релевантности, надо её поверить:

Скачать бесплатно можно, размер файла похож на разумный для длинного фильма. Поэтому

исходной странице надо поставить рел+ (за то, что можно скачать желанное кино). Хорошие страницы на которых находится много разнообразной информации, могут получать даже оценку полезный, но если информации недостаточно, всё равно будет только рел+:

<http://www.kinopoisk.ru/level/1/film/46066/>

Хотя на кинопоиске и нельзя посмотреть кино или скачать его, практически всё остальное, что только может заинтересовать пользователя, тут есть. Тем не менее, именно для "Белого солнца пустыни" контента не очень много и оценивать его полезным не стоит (в этом можно убедиться, если походить по вкладкам).

Если же фильм предлагают купить, оценка такой странице может быть не выше рел-, потому что если что-то можно бесплатно скачать, покупать файл пользователи не хотят:

<http://www.ozon.ru/context/detail/id/64141/>

Так же могут попасться страницы про какие-то иные объекты с таким названием, например, ресторан:

<http://www.restore.ru/?id=172&pkevl=0000200001>

Чтобы оценить эту страницу, можно обратиться к аннотации (в ней написано только про фильм). Однако, может оказаться, что какой-то вполне релевантный объект оказался пропущен при разметке. В нашем примере, впрочем, и без аннотации видно, что страница нерелевантна: запрос из Кирова, а ресторан - в Москве.

Однозначный запрос

Помимо общих запросов, которым соответствуют несколько объектов или несколько типов разных ответов, регулярно попадаются чётко сформулированные запросы. За счёт того, что формулировка запроса более конкретна, спектр возможных релевантных объектов сильно сужается. Посмотрим, например, на запрос [A-Style - Remember me текст]. Единственным релевантным ответом будет только текст этой песни.

<http://muzprivet.ru/lyrics/QStTdHlsZSslMiZhbXAIM01rQXNvQTPSZWllbWJcittZTo2MiQIMDI3.html>

Страница с текстом - ровно то, что искал пользователь. Это релевантный ответ.

<http://www.zaycev.net/pages/7655/765542.html>

На этой странице можно только скачать песню (ссылка на текст уводит на другой сайт).- Это нерелевантно запросу, поэтому странице следует поставить оценку нерел.

Статьи про песню без искомого текста, как и возможность скачать, будут нерелевантными: Возможность послушать будет тоже нерелевантной запросу про текст:

<http://livemusic.org.ua/song/381442.html>

Страница с видеоклипом и текстом будет релевантна, потому что на ней есть текст. Наличие видеоролика "в нагрузку" не снижает релевантность:

<http://www.probelov.net/16763.html>

Информационный запрос

Один из существенных моментов а работе - понимание разницы между оценками рел+ и полезный. Хотя эта разница и не влияет на зарплату ассессора (ошибка не значима), оценка "полезный" сама по себе очень важна. К ней надо относиться ответственно и не ставить за всё подряд, слегка отличающееся от среднего релевантного ответа. Рассмотрим для примера запрос [Айрон-Мейдон] (речь о британской рок-группе Iron Maiden).

<http://iron-maiden.ru/>

На главной странице видны только новости о группе. В случае, когда по широкому запросу нашёлся раздел сайта или главная страница тематического сайта, стоит ориентироваться не только на контент этой заглавной страницы, но и fi осмотреть на содержание сайта в целом. Например, раздел "Дискография":

<http://iron-maiden.ru/Discography.html>

По любому диску можно кликнуть, перейти на описание альбома, а оттуда - на тексты песен, например:

<http://iron-maiden.ru/Diskografija/No-Prayer-For-The-Dying-Lyrics/Fates-Warning.html>

В разделе "Аудиотека" представлены ссылки на скачивание альбомов, при этом ссылки рабочие. Например:

<http://iron-maiden.ru/Audio/Audioteka/1980-Entire-Exile.html>

<http://iron-maiden.ru/Galereya.html>

Есть большая коллекция клипов, рассказы про группу, биографии участников, форум. Сайт при этом аккуратный и даже без рекламы. Настолько качественный сайт по теме (много разноразной информации, нет рекламы) заслуживает оценки "полезный".

Впрочем, не любой фан-сайт достаточно хорош, чтобы быть полезным. Вот, скажем, такой:

<http://ironmaidenfans.ru/>

На сайте есть некоторое количество текстовой информации, не особо большая фотогалерея:

В общем-то, сайт приличный, но совсем нет треков и видео, отсутствует площадка для общения. Оценки рел+ он, безусловно, заслуживает, но полезным его назвать уже нельзя. Конечно же, по подобным запросам регулярно попадаете википедия:

http://ru.wikipedia.org/wiki/Iron_Maiden

Статья содержательная, есть обширная историческая ссылка, некоторое количество фотографий, список альбомов и даже график участников. Но тут совсем нет медиаконтента, а в запросе про группу аудио и видео, несомненно, важны. Хотя страница и релевантна запросу, ставить ей полезный будет ошибкой, правильная оценка этой статье - рел+. Качественный ответ по одной прагматике (то есть ответ определённого типа) может быть не только текстом, но, например, только прослушиванием, как на нашем Музыкальный колдунщик яндекса, как бы ни был мил нашему сердцу, всё-таки заслуживает только оценки рел+. По запросу про группу может попасться не подборка, а какая-то одна песня, например:

<http://www.lastfm.ru/music/Iron+Maiden/> /Iron+Maiden

Хотя тут и можно посмотреть клип на песню, но песня всего одна, такая страница заслуживает только оценки рел-.

Запрос про организацию

Возьмём, для примера, запрос [банк советский санкт-петербург] //(Санкт-Петербург)//. Во-первых, конечно же, у банка есть официальный сайт (он указан в аннотации):

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

Проблем с определением витального в таком случае возникнуть не должно. А вот разобраться в том, какие урлы будут рел+, а какие - полезными, несколько сложнее.

<http://www.banki.ru/banks/bank/?ID=68763>

Перед вами страница с популярного сервиса banki.ru. На ней можно увидеть контакты, краткую информацию о рейтинге (составленном экспертами) и посмотреть список филиалов. Ниже на странице находится справка о компании, далее - информация по вкладам кредитам: Также можно посмотреть дизайн карт и народный рейтинг, почитать отзывы о банке: Эта визитка содержит исчерпывающее количество информации, она структурирована и удобна для использования. Эта страница является полезной.

Не столь насыщенные разнообразной информацией визитки уже не будут получать оценку полезный, например:

<http://account.spb.ru/banks/items/14>

На этой странице представлены общие контакты банка, отделения нанесены на карту (также можно посмотреть список во вкладке "отделения") и можно найти некоторую информацию по услугам. Однако информация неполная, подробного описания вкладов и кредитов тут нет. Для рел+ информации на этой визитке уже достаточно, но ставить ей полезный - ошибка.

<http://www.bank-klient.ru/spr/spbb/sovetski-bank-spb.html>

Такая визитка ещё более примитивна, однако содержит список контактов отделений. Это

тоже рел+ страница.

Положительной оценки заслуживают и разделы витального сайта:

<http://www.sovbank.ru/private/deposit/>

На странице представлены только вклады для частных лиц. Этот раздел сайта достаточно содержательный, чтобы быть рел+, но полезный ему ставить совершенно не за что. Совсем уж примитивные визитки без развёрнутой информации получают оценку рел-, например, вот такие:

<http://www.bank24.ru/info/banks/info/?ID=271147>

и

<http://www.mk-kadar.ru/partners/sovetskw.htm>

На двух последних страницах недостаточно информации для положительной оценки, однако они не являются совсем уж нерелевантными, их правильно оценить рел-.

Запросный спам

(Релевантность)

Запросный спам встречается как на релевантных страницах, так и на нерелевантных. На оценку релевантности он не влияет. Однако, если он присутствует на странице, ей необходимо поставить оценку "спам".

Например, для запроса [скачать dirt 2 yf rjvgn.nth торрент] //(Москва)// в оценке оказался такой url:

<http://www.best-plays.ru/simuiator/arcade-simuiator/419-colin-mcrae-dirt-2.html>

С этого сайта можно скачать искомый торрент, поэтому оценка релевантности - рел+, однако на странице присутствует и запросный спам, поэтому по шкале спамности страница получает оценку спам.

Типичный пример запросного спама, как выше - блок с названием "у нас нашли", "у нас ищут", "последние запросы" и подобными, сами запросы снабжены фавиконками поисковиков, из которых эти запросы задавали. Спам ставится не за название блока, а за то, что это часть внутренней статистики, которую вебмастера нарочно выставляют наружу, чтобы повысить свои позиции в поиске. Часто таким образом накручивается релевантность и по запросам, для которых страница нерелевантна. За подобные махинации мы ставим оценку "спам".

Приведённый пример - один из характерных, на деле запросный спам бывает и хитрее. Но подобные яркие примеры должны быть оценены как спам обязательно. Пропустить оценку "спам" в случае, аналогичном приведённому - это совершенно точно ошибка, свести к спорным её не получится.

Запросный спам также ставится в случаях, когда на странице очень много поисковых запросов (необязательно в виде протянутой наружу внутренней статистики), например по запросу [drjynfrnt] //(Киев)// попался такой url:

http://www.flcd.ru/companies/vkontakteru_vkontakte_vkontakteru_ooo_vkontakte/

Здесь название сайта во всевозможных написаниях повторяется сильно больше, чем того требует нормальный текст (даже в урле три раза вписано). Важно ещё, что название повторяется кириллицей и латиницей, а ещё есть строчка "Часто/ошибочно эту компанию называют: В Контакте, drjynfrnt, d rjynfrnt, d контакте, В Контакте ру, dКОНТаКТе, В контакт, В конт, v kontakte, v kontakte ru, В контакте ru", которая сама уже достаточное основание поставить оценку "запросный спам", даже не учитывая текст выше. За перечисление названия вконтакта всеми способами, включая опечатки, обязательно надо поставить оценку "спам".

Ещё примеры:

<http://www.warez999.ru/serialv/6541-ne-rodix-krasivov-vse-200-seriv-2005dvdrip788-gb.html>

<http://psi-journal.ru/1176-vse-tainye-zhelaniya-muzhchiny.html>

Партнерка

(Релевантность)

Партнёрки - это один из видов спама. Если вы определили какую-то страницу как партнёрку, ей надо поставить оценку по релевантности и оценку спам. Например, партнёрка какого-либо магазина - это тоже магазин (в смысле, сам сайт будет сайтом магазина!);, но по сути покупка осуществляется только в основном магазине. Когда какой-то сайт плодит партнёрки, он преследует цель занять как можно больше позиций в выдаче по подходящему запросу (например, по названию товара, которым он торгует). Тогда пользователь может кликнуть на оригинальный магазин или на партнёрку, но в результате всё равно попадёт в основной магазин. Таким образом оригинальный магазин получает посетителей больше, чем получил бы в условиях честной конкуренции. Сайты-пустышки, которые ведут на первоначальный сайт, получают деньги за рекламу и переход пользователей (создатель партнёрской сети платит за привлечение клиентов). Таким образом сайтовладельцы получают клиентов (а с ними и дохода) больше, а пользователи лишаются возможности выбрать наиболее подходящий для себя магазин. Сам оригинал спамом не является, а вот все прочие сайты из партнёрской сетки - спам, они не несут никакой пользы для пользователей и снижают конкуренцию в выдаче. Если не бороться с партнёрками, может получиться, что вся выдача будет заполнена одним магазином (но выглядеть на первый взгляд будет так, словно их несколько). Поэтому партнёрки - зло.

Чтобы поверить, является ли магазин партнёркой или нет, надо начать совершение покупки, дойти до стадии оформления заказа. На самом деле оформлять ничего не надо, естественно. Но если в процессе вас перекидывает на какой-то другой магазин - это партнёрка. Если перекидывает на какую-то систему оплаты (Paypal, Webmoney и тп) - это, естественно, не имеет никакого отношения к партнёркам.

Например, запрос [Л Нильсон Ребенок родился! Чудо зарождения новой жизни скачать] //(Петрозаводск)// и вот такой урл:

<http://www.03portal.kz/index.php?option=com/content&task=view&id=21226>

Это магазин, скачать с него нельзя, поэтому оценка релевантности - нерел. Но надо ещё проверить на предмет, а не партнёрка ли это (да и вообще вид сайта намекает на то, что что-то тут нечисто). Надо попробовать оформить покупку, нажав на ссылку "купить". В

норме, это должно привести к тому, что появится товар в корзине. Но в этом случае пользователя перебрасывает на другой сайт

<http://www.flip.kz/catalog?prod=6316>

Таким образом, сайт из оценки - партнёрка, а значит, спам.

Сателлит

(Релевантность)

Сателлиты, как и партнёрки, являются в некотором роде вспомогательными сайтами, которые искусственно увеличивают выгоду основного сайта. Но есть и отличия. Партнёрскими сайтами (которые перенаправляют людей на оригинальный сайт), как правило, владеют какие-то люди, не связанные с владельцем исходного сайта, связанные с ним только договором о работе партнёрской сети. Сателлиты же создают сами хозяева основных сайтов, они выглядят как сайт другой компании (иногда даже с другими контактами), но организация за основным сайтом и его сателлитами стоит одна. Цель всё та же - заполнить выдачу "своими" результатами, плюс умелый вебмастер может манипулировать ссылками между своими сайтами ради искусственного повышения релевантности какого-то из них. В общем, сателлит в выдаче выгоден его создателям, а пользователям - нет.

Если вы определили какой-то сайт как сателлит, ему надо поставить оценку релевантности, которую он заслуживает и оценку "спам".

Рассмотрим, например, запрос [молоковоз на шасси Камаз 65115] //(Алматы)//. В оценку попал следующий урл:

<http://www.spectechnikagaz.ru/molokovozy/>

Вроде бы сайт называется АвтоСпецТехника, но урл - СпецТехникаГаз, как будто этот домен для части большего общего целого. Дизайн шаблонный без изысков. Это ещё не повод поставить спам, конечно, но должно насторожить, то есть при беглом взгляде вам должно стать понятно, что сайт подозрительный и надо его проверить. Можно просто кликнуть по логотипу "АвтоСпецТехника" и увидеть, что он перенаправляет на основной сайт, а можно и повнимательнее посмотреть на страницу из оценки:

Адрес электронной почты указан для другого домена, того самого, на который ведёт ссылка из логотипа.

В данном случае владельцы сетки сателлитов не потрудились сделать им существенно разный дизайн и даже указали контакты основного сайта на второстепенном. Это довольно простой случай, но иногда удаётся опознать и более качественно сделанные сателлиты, их хорошо видно, когда сателлит и основной сайт попадают в одном задании. Однако особенно искусно сделанные сателлиты и вовсе почти невозможно обнаружить без основного сайта в том же задании (без сравнения).

Псевдосайт

Псевдосайты очень похожи на нормальные сайты, но на самом деле являются просто подделкой под сайт организации или бессмысленным набором данных, они создаются для того, чтобы зарабатывать на рекламе и продажных ссылках. За счёт схожести на нормальные сайты и наличие второконтента, который сам по себе вполне нормальный, такие сайты могут оказываться на высоких позициях в поиске.

Например, для запроса [консультация логопедаонлайн] //(Москва)// в оценке оказался урл

<http://jenkoncult.ucoz.ru/news/2009-01-06-100>

Он немного похож по конструкции на нормальную консультацию. Однако нет никаких контактов, совершенно неясно, что будет, если отправить письмо через форму и вообще на сайт сколь-нибудь адекватной компании это непохоже.

Эта страница - псевдосайт, она выдаёт себя за онлайн-консультацию, на деле таковой не являясь.

Ссылки на псевдообменники

Примечание: антивирус Доктор Веб с некоторой точностью умеет определять наличие этих вредоносных ссылок и сообщает о том, что страница может представлять опасность для компьютера. В целом, это, конечно, верно. Однако ставить таким страницам оценку "вирус" некорректно, надо ставить оценку релевантности и спам.

Ссылки на ненастоящие файлообменники по текущей инструкции являются один из видов спамконтента. Многие вебмастера грешат их использованием и этот вид мошенничества довольно регулярно попадает в запросах про скачивание или Просмотр чего-то онлайн, поэтому заслуживает отдельного рассмотрения в силу своей популярности.

Проверять, что же за файл скачивается с вредоносной ссылки, ассессорам не надо. Однако, чтобы подтвердить подозрения, мы Пробовали качать файлы с таких обманных ссылок, когда они вообще есть (иногда ссылка ведёт в какое-нибудь невразумительное место, даже без скачивания). В большинстве случаев качается архив, который просит смску за открытие. ЭТО уже подозрительно.

Но даже если архив удаётся распаковать (например, потеряв некоторую сумму на экспериментах с смсками), в нём, в лучшем случае, находится Инструкция по работе с нормальными и файлообменниками и торрентами (не подробная, конечно). В худшем ещё и какой-нибудь вирус сидит. В общем, качать это всё, естественно, не надо. Однако пользователи по ошибке могут и скачать что-то с этих ссылок. При этом желанного контента они не получают в любом случае. Поэтому урлам с такими нехорошими ссылками, помимо оценки релевантности необходимо поставить "спам", потому что на странице присутствует обман. При этом на релевантность обманные ссылки влияния не оказывают.

Вот например, в оценке запроса [скачать торрент бесплатно на русском языке без регистрации] //(Москва)// попался урл

<http://iuxvri.biz/download/%D1%81%D0%BA%D0%B0%D1%87%D0%B0%D1%82%D1%8C+%D0%B1%D0%B5%D1%81%D0%BF%D0%BB%D0%B0%D1%82%D0%BD%D0%BE+torrent+1.8.5+%D0%BD%D0%B0+%D1%80%D1%83%D1%81%D1%81%D0%BA%D0%BE%D0%BC+%D1%8F%PQ%B7%D1%8B%D0%BA%D0%B5+%D0%BD%P0%B5+zip.html>

Как правило, эти обещания в этих ссылках повторяют контент страницы, но тут сама страница вообще о другом, а с запросом связана только эта спамная реклама. При оценке нет необходимости углубляться в изучение вредоносных ссылок, важно поставить оценку "спам" за их наличие (если вам нужно углубиться в изучение, чтобы понять, что ссылка спамная, то придётся таки углубиться). Посмотрим на примере этого сайта, почему же такой блок ссылок - обман, а значит и спам.

Первая ссылка ведёт на скачивание файла с сомнительным названием в домене .tk (в этом домене примерно 99% сайтов спам, утверждать, что любой сайт в этом домене - спам, нельзя, но подозрение вызывает существенное). Чтобы убедиться, что это какой-то обман, можно, например, зайти на этот сайт, однако настоятельно рекомендуется этого не делать, на таких сайтах почти всегда находятся вирусы (на этом сайте, например, троян есть):

При скачивании файла по ссылке из нехорошего блока пользователь получит только текст этой самой "оферты", возможно, снабжённый вирусом (в рассматриваемом случае вирус будет). Это проверено экспериментально, ассессорам перепроверять не стоит, ибо шанс подцепить вирус велик (в примере шанс и вовсе стопроцентный). В некоторых случаях эти ссылки вместо скачивания мусорного файла (замаскированного под искомый) уводят пользователя на псевдофайлообменники. Например, для запроса [Лучшие книги серии метро 2033] //(Москва)// в оценке оказался такой урл:

<http://dbbook.ru/xbook/fantastika/apocalipsis/metro/3249-metro-2033-bezvmvanka--audiokn.html>

Спамная ссылка ведёт на такую страницу:

<http://deposit.stimulnmall.com/results.php?notds=l&query=%CC%E5%F2%F0%EE%202033.%20%C1%E5%E7%FB%0EC%FF%ED%EA%E0%20%28%CQ%F3%E4%E8%EE%EA%ED%E8%E3%E0%29%20%F0%E0%E7%EC%E5%F0:59Q%EC%E1>

Это псевдообменник, который маскируется под Deposit Files.

Ещё несколько примеров спамных блоков:

<http://multikonline.ru/935-zviozdnve-sobaki-belka-i-strelka.html>

<http://www.sefon.ru/mp3/%D0%90%D0%BD%D0%BD%D0%B0%D0%A0%D0%B5%D0%B7%D0%BD%D0%B8%P0%BA%D0%BE%D0%B2%D0%B0/%P0%9F%D1%80%D0%BE%D1%89%D0%B0%D0%BB%D1%8C%D0%BP%D1%8B%P1%85%P1%81%P0%BB%P0%BE%D0%B2%D0%BD%D0%B5%D0%B3%D0%BE%D0%B2%D0%BE%D1%80%D0%B8>

<http://xxmuz.ru/shanson/3400-armeiskie-pesni-pod-gitaru-2-2010-bonus.html>

<http://uploader.ws/music/106027-armeiskie-pesni-pod-gitaru-mp3.html>

<http://nvlq.com/index.php?newsid=947>

<http://tabulorasa.info/67546-bokovava-panel-d-lva-xr-gadzhetv-sidebar-for.html>

http://www.filmbom.ru/film_load/serial_load/16953-ne-rodix-krasivoy-vse-200-seriy-2005dvdrip788-gb.html

<http://www.kino-line.biz/musik/10593-top-100-zaicevnet-2011.html>

<http://www.videoanons.ru/film/comedi/1775-bitidzhus-beetle-iuice-1988-dvdrip.html>

http://sonic002.ucoz.ru/news/skachatbitldzhus_1988_bdrip_besplatno/2011-01-05-24364

<http://buenos.ru/176260-policeiiskaia-akademiia-police-academy-1-7-1984.html>

<http://www.fataH.org/6083-prikolnve-ringtonv-dlia-mobily..html>

<http://kinosovetnik.ru/index.php?do=tags&tag=%D4%EE%F0%F1%E0%E6%205>

Все приведённые страницы - спам (по любому запросу!), из-за наличия блоков с обманными ссылками. Дизайн блоков варьируется, но суть остаётся общей. Все такие страницы отправляются в спам.

В некоторых ситуациях, как в примере с книгой про метро 2033, нормальная ссылка и обманная могут выглядеть похожим образом. Если вы не уверены, спам перед вами или нормальная ссылка, проверьте её.

Поиск организаций

(Релевантность)

Рассматриваем запросы вида [название организации], [сайт организации].

Разницы между такими запросами нет, в обоих случаях пользователя в первую очередь интересует

Официальный сайт организации. (Витальный)

Но страница с результатами поиска показывает 10 ссылок, а не одну. Поэтому, при оценке прочих документов стоит задуматься над тем, как должна выглядеть выдача на позициях 2-10: что целесообразно показать пользователю после ссылки на Витальный документ.

Какая информация может быть полезна помимо официального сайта?

Та, которой нет на официальном сайте

Та, которая есть на официальном сайте, но может быть представлена в более удобной и компактной форме.

К первой группе можно отнести следующую информацию:

Альтернативные источники информации: отзывы, статьи, рецензии и пр.

Информация и сервисы, связанные с деятельностью организации: расписания, афиши, продажа билетов и тд.

Новости об организации

Общие «исторические» сведения

Ко второй группе можно отнести

развернутые визитки организаций, содержащие большое количество информации о компании.

Хорошие страницы с подобным содержанием заслуживают положительных оценок.

Сайты подразделений организации, полезные внутренние страницы витального сайта

Запрос ГСПбГУ1: www.spbu.ru - Витальный, а сайты факультетов - Полезны.

При этом страницы филиалов организации не в регионе запроса оцениваются как Рел-.

Запрос [ГУ ВШЭ] (Москва): www.hse.ru - Витальна, сайт магистратуры в ВУЗе

<http://ma.hse.ru/index.html> - Полезен, а сайт Нижегородского филиала www.hse.nnov.ru - Рел-
тк не соответствует региону запроса.

Информация и сервисы, связанные с деятельностью организации: расписания, афиши, продажа билетов и пр

Запрос [Большой театр]

<http://www.bilety-bolshoi.ru/> - портал, целиком посвященный продаже билетов в Большой театр, есть афиша, план зрительного зала, новости и пр. Страница Полезна.

<http://xwt.ru/index.php/menu/place/2> - продажа билетов в Большой на тематическом портале. Оценка - Рел+.

Запрос [Казанский вокзал]

http://www.tutu.ru/poezda/vokzal_02.php - расписания поездов, заказ билетов. Страница Полезна.

Развернутые подробные визитки организаций: содержат нужную пользователю информацию, которую можно найти и на официальном сайте, но здесь она представлена в более компактной/удобной форме. В такой визитке должна быть представлена полная контактная информация (адрес+телефон+ ссылка на официальный сайт), а также дополнительная полезная информация.

Запрос [банк ВТБ]

<http://www.banki.ru/banks/bank/?ID=327> - хорошая, полноценная визитка - со списком контактов, подборкой новостей и Прочей полезной информацией. Оценка - Полезный.
Запрос [Naf Naf] (Москва)

<http://www.4shopping.ru/brands.php?id=76> - на странице представлены последние коллекции одежды этой марки, ссылка на официальный сайт, а также приведен список адресов фирменных магазинов в регионе. Оценка - Полезный.

Для такого запроса страница с адресами магазинов в другом регионе, например, <http://www.4shopping.ru/brands.php?id=76> была бы Рел+, тк помимо контактов здесь представлена релевантная информация - актуальные коллекции одежды.

Что считаем Релевантным+?

Подборки новостей по теме, а также самые свежие новости, даже если они одиночные
Альтернативные источники информации: полезные отзывы, статьи, рецензии и Ир.

Запрос [Auto Global Usa]:

Страница <http://wwwboards.auto.ru/blacklist/73568.html> -Рел+.

Это альтернативное мнение о компании, такая информация пользователю будет действительно ценна, если он решит воспользоваться услугами данной компании. Если бы на странице были приведены свежие отзывы, можно было бы поставить и оценку Полезный. Общая, историческая информация об организации/предприятии. Такую информацию не стоит путать с бесполезными рекламными текстами, которые на заслуживают оценку выше рел-.

[Стадион Metallurg Самара]

<http://www.redwhite.ru/stadiums/pleague/samara.php> - представлена общая информация о стадионе и контакты, оценка Рел+

[При этом очень хорошие материалы общего характера можно отметить оценкой Полезный. Страницы, которые могли бы быть оценены как Полезные по содержанию, но по каким-то причинам оценку им хочется понизить. Например:](http://ru.wikipedia.org/wil<i/%D0%9C%D0%B5%D1%32%DO%B0%D0o/oBB%D0%BB%D1%83%D1%80%PQ%B3 (%D1%81%D1%82%D0%B0%D0/oB4%D0%B8%D0%BE%D0%BD. %D0%A1%D0%B0a/oD0%BC%D0%B0%D1%80%D0%B0) - Рел+</p></div><div data-bbox=)

Отзывы, рецензии, статьи и тд, но в недостаточном количестве для оценки Полезный.

При этом не стоит забывать, что устаревшие статьи, новости и тд, а также рекламные неинформативные тексты не заслуживают положительной оценки.

Если контакты в визитке неполные (нет ссылки на официальный сайт/ адреса/телефона) или не соответствуют региону запроса, но при этом страница содержит развернутую, релевантную дополнительную информацию (новости, подробные общие сведения, отзывы), оценка Рел+.

[Стадион Metallurg Самара]

<http://fratria.ru/cgi-bin/iTable/runbroker.cgi?page=2&id=41&sport%20type=type1%20&tournament=tour16&year=2002> - схема расположения мест, подробное описание того, как добраться до стадиона + подборка фотографий. Оценка Рел+.

Запрос [Центробувь]

<http://www.fshops.ru/?NID=206> -хорошая визитка с полным списком адресов магазинов, заслуживает положительной оценки, но тк другой дополнительной информации нет, на Полезный не тянет. Оценка Рел+.

Что не оцениваем положительно?

Примитивные визитки, содержащие только контакты организации (даже если приведена минимальная дополнительная информация - карта, схема проезда) - не заслуживают оценки выше Рел-.

<http://www.metizmetall.ru/company/656.html>

<http://www.center-s.ru/firmsdescript/firm4778.htm>

<http://www.yondi.ru/firmcard/frmid/8396/>

<http://rusbase.ru/firma/Feniks-Grupp-ooo>

<http://www.rb.ru/biz/companies/show/537917>

http://www.pulset.ru/usl/find13_1u.php?id=18292&typeusl=10

Подобные визитки заслуживают оценки Рел- даже в том случае, если ищут малоизвестную компанию, у которой в принципе нет официального сайта. Можно предположить, что более релевантной информации о такой компаний (новостей, статей и тд) в интернете все равно не найдется, поэтому страница с оценкой Рел- займет заслуженное место вверху выдачи.

Неинформативные рекламные тексты о компании, а также устаревшие новости, неактуальные статьи, малоинформативные одиночные отзывы о компании типа «Плохо/Хорошо»

Например, <http://www.infoorel.ru/firma.php?id=802>,

<http://prov.nag.ru/prov/root/region65/settlement114/providers/provider914/>

контакты неполные, а отзывы нельзя считать информативными. Оценка Рел-

Если иное не следует из переформулировок, вакансии компаний заслуживают оценки Рел-, только если они не являются одним из видов дополнительной информации, приведенной на хорошей, полезной визитке.

Запрос [Союзпромпласт]

<http://hh.ua/employer/536740> - рел-

(Релевантность)

Можно выделить несколько основных типов запросов про работу:

Общие - работа, ищу работу, поиск работы, найти работу

Гео - работа в Москве, работа в нижнем Новгороде

Специальности, разделы - вакансии водитель.

Работа в конкретной организации

"Тип работы (работа для студентов, работа на Дому)

Конкретные вакансии, узкая специализация

В целом, принципы оценки запросов этого типа вписываются в общие правила релевантности, но надо не забывать о некоторых особенностях:

Локализуемость:

Для запроса из областного центра:

Релевантны+ - предложения только в данном городе (например, для запроса из Москвы - только московские вакансии релевантны);

Релевантны-вакансии в пределах области ;

Остальные - Нерелевантны;

Для запроса из областного города:

Релевантны+ - как предложения из данного города, так и из областного центра;

Релевантны- - предложения из других городов данной области;

Остальные - Нерелевантны.

Актуальность:

Некоторые вакансии, такие как кассир в Пятерочке, могут не устаревать годами - текучка. Но дата обновления на нормальных сайтах для любой вакансии должна быть не более месяца.

Т.е. если дата создания старая, а даты обновления нет - объявление Нерелевантно. Если дата обновления есть, и она не старше месяца, то можно считать что вакансия актуальна.

Специфика Один-Много:

Вот моменты, которые заимствованы из Один-Много:

Одиночные предложения всегда оцениваются не выше чем Релевантны-

Если в списке объявлений встречаются вперемешку релевантные и нерелевантные (по региону, актуальности, тематике), для положительной оценки релевантных предложений должно быть больше половины

НО

Если по общему запросу вы встречаете подборку вакансий по предприятию или специальности, это никогда не заслуживает выше, чем рел-

Если по общему запросу попадают более крупные категории (например,

<http://volqograd.job.ru/office/>

по запросу «работа» (Волгоград)), их нужно оценивать в соответствии со здравым смыслом: популярные категории получают рел+, а непопулярные - рел- (например,

<http://volqograd.job.ru/design/> по тому же запросу)

Релевантны или Полезны?

Для оценки Релевантны+ достаточно, чтобы страница удовлетворяла перечисленным выше условиям: количеству релевантных вакансий, региону, актуальности. Однако, если база вакансий плохо структурирована, в ней трудно ориентироваться, то оценивать её выше рел+ не стоит. Полезным можно назвать такой ресурс, где набор релевантных вакансий не только обширный, но и удобный к использованию (есть фильтры, сортировка и тп). Вакансии могут быть отсортированы по фирме (если запрос достаточно широк, что релевантны вакансии разных фирм), по ЗП, по региону, по типу работы и тд.

Например:

по запросу [Поиск работы] региональный поддомен портала <http://hh.ru/> можно оценить как Полезны;

по запросу [Работа для студентов] можно оценить как Полезную страницу <http://www.iob.ru/students>, так как даёт большой список вакансий, который можно фильтровать (справа «фильтры результатов»).

Но по запросу [Работа в фирме АйТи] из Москвы страница

<http://www.superjob.ru/clients/52018/aiti-informacionnye-tehnologii> не дотягивает до Полезной, так как здесь релевантная информация перемешана с нерелевантной (не совпадающей по региону).

Витальный

При запросах о работе в Конкретной фирме раздел с вакансиями на официальном сайте

данной фирмы будет витальным.

Например: по запросу [работа в Газпром Банке] страница

<http://www.gazprombank.ru/vacancy/list/> будет Витальна.

По более общим запросам ([ищу работу], [работа для студентов] и тп) витальных страниц не будет.

Сужение/Обобщение

Страницам, дающим слишком общий или слишком узкий ответ на запрос, стоит понижать оценку. Например: по запросу [ищу работу] региональный поддомен портала <http://hh.ru/> можно назвать Полезным, но по более узкому запросу [работа программист] он заслуживает оценки Релевантный-, так как слишком расширяет область поиска. В то же время, тематический раздел с вакансиями только программистов

<http://www.rabota.ru/vacancy/%D0%9F%D1%80%D0%BE%P0%B3%D1%80%D0%B0%D0%BC%D0%BC%D0%B8%D1%81%P1%82/>

по широкому запросу [ищу работу] будет Релевантным-.

Морды больших порталов по общим запросам

Если на портале отсутствует достойная подборка вакансий по запросу в нужном регионе, он всегда оценивается отрицательно

Большинство крупных порталов с вакансиями содержат предложения для разных регионов. Положительной оценки заслуживает только раздел, соответствующий региону запроса (это может быть домен второго уровня, раздел сайта, либо глобальный url, но когда регион определяется автоматически). В каждом случае нужно понимать, какому региону соответствует оцениваемая страничка и ставить оценку, учитывая его
ВАЖНО: Чтобы верно оценивать редиректы на поддомен, привязанный к региону (вашему/запроса/третьему), нужно перед началом работы с заданием почистить куки, и затем повторять процесс перед оценкой каждой новой страницы того же сайта.

Помимо морды, предлагающей, как правило, не только поиск работы, на многих сайтах обычно есть более узконаправленные страницы /например.

<http://volaoorari.suneriob.rii/var,anr,v/searrh/> регион - Волгоград).

Это не должно являться поводом для снижения оценки морды

Что касается методов навигации на сайтах, которые нас устраивают, ограничений нет. Это может быть подборка по тэгам, либо незаполненная (напоминаю, что в этом параграфе рассматриваются общие запросы) плашка поиска, либо другие способы, дающие на выходе искомую информацию. Проверяйте лишь, что её действительно можно получить: нужный регион, если он не выбран до сих пор, доступен для поиска; если же был выбран - градация должна сохраняться

Тонкости оценки:

Q: надо ли снижать оценкам сайтам, которые не показывают всех контактов работодателей и требуют регистрации, чтобы подать заявку/ "apply job"? (запрос "myrtle beach jobs")

<http://regionalhelpwanted.com/myrtle-beach-jobs/> <http://www.snaqaiob.com/iobs/SouthCarolina/MvrtleBeachjobs.html>

A: нет не нужно, схема откликнуться на вакансию, или ответить на вакансию резюме — очень распространена
(Релевантность)

Согласно законодательств стран все законы, законодательные акты, постановления правительств подлежат официальному опубликованию в средствах массовой информации в полном соответствии с подписанными подлинниками в официальных изданиях. Все общегосударственные постановления, акты и т.д. издаются в аккредитованных изданиях. В свою очередь, постановления министерств и ведомств, законы субъектов федерации, Сената и Мажилиса (Казахстан), издаются на порталах министерств и ведомств, субъектов федерации.

Для официальных документов: Витальными мы считаем страницы на сайте соответствующего ведомства.

Язык официальной публикации: Беларусь, Казахстан, субъекты РФ - на русском и национальном языках; Украина - на украинском.
Для указов, постановлений и других оф. документов (кроме законов) - витальными являются тексты на страницах профильных органов исполнительной власти (Министерств) или президента, если это президентский указ.

//РОССИЯ//

Согласно Закону России "О порядке Опубликования и вступления в силу федеральных конституционных законов, федеральных законов, актов палат Федерального Собрания" федеральные конституционные законы, федеральные законы подлежат официальному опубликованию в течение семи дней после дня их подписания Президентом Российской Федерации.

Официальным опубликованием актов Президента Российской Федерации и актов Правительства Российской Федерации считается публикация их текстов в "Российской газете" или в Собрании законодательства Российской Федерации.

Официальными являются также тексты актов Президента Российской Федерации и актов Правительства Российской Федерации, распространяемые в машиночитаемом виде научно-техническим центром правовой информации "Система", который создает и поддерживает эталонный банк правовых актов высших органов государственной власти Российской Федерации, а также издает и распространяет официальные издания "Собрание законодательства Российской Федерации" и "Бюллетень нормативных актов федеральных органов исполнительной власти" Таким образом, витальными источниками законов России являются: <http://www.rg.ru/>, <http://www.kremlin.ru/> и <http://www.szrf.ru/>

//ДОПОЛНЕНИЯ К ЗАКОНАМ РФ//

Согласно главе 3 Конституции РФ, «Федеративное устройство», в составе Российской Федерации находятся субъекты Российской Федерации.

Официальным опубликованием законов субъектов Федерации, правовых актов Государственных Советов, правовых актов Глав правительств, нормативных правовых актов министерств и иных органов исполнительной власти, а также договоров, соглашений органов государственной власти субъектов Федерации считается первая публикация полного текста этих актов в средствах массовой информации, являющихся официальными изданиями в пределах субъекта.

Например, официальными изданиями Республики Коми, согласно Закона республики «О порядке опубликования и вступления в силу законов Республики Коми и иных правовых актов органов государственной власти Республики Коми» в статье 8 сказано:

Официальным опубликованием считается первая публикация полного текста законов и актов в газетах "Республика", "Коми му" либо в журнале "Ведомости нормативных актов органов государственной власти Республики Коми" на государственных языках Республики Коми
20.01.2012 2:11

//УКРАИНА//

С целью упорядочивания официального обнародования нормативно-правовых актов, Указом Президента «О порядке официального обнародования нормативно-правовых актов и вступления их в силу» N 503/97 от 10.06.1997 установлено, что законы Украины, другие акты Верховной Рады Украины, акты Президента Украины, Кабинета Министров Украины не позднее как в пятнадцатидневный срок после их принятия в установленном Порядке и подписания подлежат обнародованию в официальных печатных изданиях.

Официальными печатными изданиями Украины являются:

Бюллетень «Официальный вестник Украины»;

Газета «Правительственный курьер»;

Газета "Голос Украины";

Газета "Ведомости Верховной Рады Украины".

Витальными электронными источниками являются: <http://www.golos.com.ua/>

<http://ukurier.gov.ua/uk/>

<http://zakon1.rada.gov.ua/laws>

Витальные источники Для всех официальных документов (законы, указы, постановления и пр.) - докумены в «Базе Законодательства» на портале Верховной Рады

<http://zakon1.rada.gov.ua/laws/main>

// Б Е Л А Р У С Ь //

В соответствии с Декретом Президента от 10 декабря 1998 года «Об утверждении Положения об официальном опубликовании и вступлении в силу правовых актов Республики Беларусь», решения республиканских референдумов, декреты Президента Республики Беларусь и законы Республики Беларусь после подписания Президентом, указы Президента, постановления Совета министров нормативного характера подлежат немедленному и обязательному официальному опубликованию.

Иные правовые акты могут быть официально опубликованы после включения их в Национальный реестр правовых актов Республики Беларусь.

Органы официального опубликования:

издание Национального реестра правовых актов Республики Беларусь и его электронная версия;

газета «Советская Белоруссия»;

газета «Звезда»;

«Народная газета»;

газета «Рэспублжа»

Официальное опубликование правового акта осуществляется, как правило, на белорусском и русском языках в соответствии с подписанным подлинником.

Витальными источниками являются:

<http://www.pravo.by/>

<http://www.sb.by/>

<http://www.zvyazda.minsk.by/ru/main/>

<http://www.nq.by/>

<http://www.respublika.info/>

//КАЗАХСТАН//

Согласно Закону Республики Казахстан «О нормативных правовых актах» от 24 марта 1998 года официальное опубликование законов и других нормативных правовых актов является обязательным условием их применения. (Глава 6, статьи 31, 32, 33, 34, 35, 35-1)
Официальными изданиями являются Ведомости Парламента Республики Казахстан, Собрание актов Президента и Правительства Республики Казахстан, Собрание актов центральных исполнительных и иных центральных государственных органов Республики Казахстан, а также периодические печатные издания, получившие такое право на конкурсной основе.

Право официального опубликования нормативных правовых актов на республиканском уровне принадлежит следующим печатным периодическим изданиям:

АО «Республиканская газета «Егемен Казахстан»;
АО «Республиканская газета «Казахстанская правда»;
ТОО Медиа-Корпорация «Зан»; 1

ЗАО «Издательский дом «БИКО»/Официальная газета Витальными источниками являются:

www.parlam.kz/Leaacts.aspx?lan=ru-RU

<http://www.ru.government.kz/>

<http://www.egemen.kz/>

<http://www.kazpravda.kz/>

(Релевантность)

Локализуемость:

Для запроса из областного центра:

Релевантны+ - предложения только в данном городе (например, для запроса из Москвы - только московские предложения релевантны);

Релевантны- - предложения в пределах области ; Остальные - Нерелевантны;

Для запроса из областного города:

Релевантны+ - как предложения из данного города, так и из областного центра;

Релевантны- - предложения из других городов данной области; Остальные - Нерелевантны.

Для запроса с явно указанным городом:

Релевантны+ - предложения только из этого города; Остальные - Нерелевантны.

Актуальность объявлений:

Релевантные*- предложения должны располагаться на "живых" страницах. Чтобы убедиться, что объявление еще "живое", можно посмотреть:

а)на дату его последнего обновления или создания объявления (для запросов типа сниму/сдам в аренду срок последнего обновления не должен превышать 1 месяца);(для запросов про земельные участки/коттеджи срок может быть больше - до полугода.)

б)если даты обновления нет, стоит посмотреть на сам сайт - если это каталог/доска объявлений, то как часто на нем обновляются предложения; если это сайт отдельной фирмы - стоит проверить, что сайт "живой" (на нем регулярно обновляются новости и тп)

ВНИМАНИЕ обратите внимание на галочки актуальности. Не забывайте, что разделы досок объявлений по теме запроса, которые регулярно обновляются автоматически, являются динамическим контентом и не устаревают.

Например <http://irr.ru/real-estate/out-of-town/lands/> по запросу "купить участок в Подмосковье" стоит оценить как Рел+ и Никогда не устареет.

Но если нам попалось предложение от какой-то одной компании (типа "распродаем последние 3 участка"), такой ответ может иметь срок устаревания (например, полгода). В то же время ответ <http://www.dolina-imeniy.ru/> скорее Никогда не устареет, тк вряд ли все предлагаемые на сайте коттеджные поселки построят и продадут за 1 год, а больших сроков устаревания нашей оценкой не предусмотрено.

Один-Много:

Одиночные предложения всегда оцениваются не выше чем Релевантный-;
Оценку Релевантный+ и выше могут заслужить только списки из нескольких (?2) релевантных объявлений;

Если в списке объявлений встречаются вперемешку релевантные и нерелевантные (по региону, актуальности, тематике), для положительной оценки релевантных предложений должно быть строго больше половины.

Релевантный или Полезный?.

Для оценки Релевантный+ достаточно, чтобы страница удовлетворяла перечисленным выше условиям: количеству релевантных предложений, региону, актуальности. Однако, если база плохо структурирована, в ней трудно ориентироваться, то оценивать её выше рел+ не стоит. Полезным можно назвать такой ресурс, где набор релевантных объявлений не только обширный, но и удобный к использованию (есть фильтры, сортировка и тп). Предложения могут быть отсортированы по району, по цене, по типу сделки, по типу помещений и тп.

Сужение/Обобщение

Страницам, дающим слишком общий или слишком узкий ответ на запрос, стоит понижать оценку. Например, по широкому запросу "Недвижимость" положительной оценки заслуживают лишь большие информационные порталы или обширные каталоги объявлений, а отдельные предложения купить квартиру в каком-то конкретном доме будут слишком частным ответом на такой запрос и должны оцениваться как Релевантный-.

Сдам/Сниму

По запросам типа "сниму (квартиру/комнату и тп)" Релевантны+ предложения квартир в аренду Объявления типа "я хочу снять квартиру в этом районе" (т.е. спрос на квартиры) Релевантны- (но не Нерел!), тк пользователь, задавая запрос, скорее хочет найти для себя квартиру, а не узнать, что кто-то еще тоже хочет снять квартиру.

По запросам типа "сдам (квартиру/комнату и тп)" Релевантны+ предложения недвижимости в аренду

Тонкости

Если в запросе сказано, что ищут предложения о посуточной аренде недвижимости, Релевантны+ будут только предложения, в которых в явном виде написано, что она сдается посуточно. Если ничего на этот счет не сказано - предложение Релевантно-. Если явно сказано, что недвижимость сдается только на длительный срок, такое предложение по такому

запросу Нерелевантно.

(Релевантность)

Не секрет, что пользователи часто ищут в интернете порно. Поэтому и на оценку часто приходят запросы, по которым требуется эротический или порнографический контент. Если ассессор не хочет оценивать порнографию, можно от неё Отказаться. Администратор закрывает доступ к соответствующего рода запросам.

Нижеследующая инструкция предназначена для смельчаков, готовых работать с "горячим" контентом :)

Теория

Что мы называем "порнографией"

Спор о том, что является порнографией, а что эротикой, Старинный и бесплодный. Мы в него не встречаем :) Порнографией мы считаем всё то, что не стоит видеть детям.

Не считаем порнографией мы научно-учебные и медицинские материалы, а также произведения изобразительного искусства. При этом стоит всё же придерживаться здорового консерватизма и работы, подобные, к примеру, работам классика Фотоискусства Роберта Мэпплторпа оценивать как порнографию в тех случаях, когда они того заслуживают.

Типы запросов

Что касается отношения к порнографии, то в этом аспекте запросы бывают "белые", "серые" и "чёрные".

"Белый" запрос не имеет отношения к порнографии, и в выдаче по такому запросу никакой порнографии быть не должно. Однако оптимизаторы порносайтов - люди изобретательные, и могут порой показать порнографию даже по запросу "экскаватор". Порно-ответ по "белым запросам" может быть формально релевантен (например, <http://fotki.yandex.ru/users/elfim-vladimir/view/98865?page=4&page=4> по запросу "гифка пиво"), но должен получить отрицательную оценку, как в картинках, так и в релевантности.

"Серый" запрос - это запрос, по которому есть и заметный "нормальный", и заметный порноинтерес. Подробнее позиция изложена по ссылке. Формально порно-ответ на "серый" запрос будет хорошим и ожидаемым - но мы его всё равно показывать не хотим. Поэтому по запросу "cream pie" мы ищем только рецепты, по запросу "беременные" медицинскую информацию, а по запросу "аниме" не хотим видеть хентай.

"Чёрный" запрос - это запрос, явно указывающий на поиск порнографии. Только по таким запросам она будет релевантна.

Примечания:

Иногда в силу разных случайностей к вам могут попасть запросы на нелегальную порнографию или экстремальные перверсии: детское порно, зоофилия, копрофилия, гуру-хентай и т.п. Вы не обязаны их оценивать. Само содержание такого запроса - достаточный аргумент для отказа.

Многие "чёрные" порнозапросы представляют собой не столько ясное указание на то, что ищет пользователь, сколько его, пользователя, эротические фантазии. При оценке подобных запросов считаем, что вся порнография - постановочная. Никаких "настоящих" съёмов скрытой камерой, изнасилований, инцестов, лишений девственности и т.д. и т.п. искать не нужно, достаточно того, что заявляется в подписи при ролике или фотосессии. "Шлюхи", "сучки", "развратницы" = "любые порноактрисы". Определения "жёсткая" или "извращённая"

тоже можно игнорировать (например, по запросу "извращённое анальное порно" годится любое анальное порно). Действительно, есть много типов порнографии, и порой в запросе чётко указывается нужный тип. Как поступить, если релевантность ответа неочевидна? (релевантна ли грудь третьего размера по запросу "большие сиськи"?) В таких случаях считаем, что годится всё, кроме того, что откровенно нерелевантно (т.е. верим, что девушке, выданной по запросу "студентки", действительно двадцать лет - если она не явно ближе к сорока).

Шкала оценки

Порнозапросы мы оцениваем по такой же шкале, что и обычные, как в картинках, так и в релевантности.

В релевантности ресурс можно оценить как Полезный, если он предоставляет много бесплатного качественного контента и удобен в использовании (не замусорен рекламой). Не забывайте проверять документ на спам.

В релевантности сохраняются обычные правила относительно возможности скачать или посмотреть онлайн. Т.е. требования заплатить или зарегистрироваться снижают оценку. Однако на порносайтах есть ещё один, специфический способ затруднить доступ к видео онлайн: искусственное замедление воспроизведения. Пример: vuku.ru За такое оценку тоже нужно снизить.

Практика

Порнозапросы встречаются и в релевантности, и в картинках. В релевантности положительные оценки - Релевантный+ и Полезный, а в картинках - Релевантный+ и Релевантный-. Чтобы избежать путаницы, будем называть оценки "положительными" и "отрицательными".

Примеры конкретных запросов

По запросам «порно» или «xxx» положительную оценку получают фото, видеоролики, фильмы и прочие материалы, в которых содержится «непристойная, вульгарно-натуралистическая, циничная фиксация сцен полового акта и самоцельная детализированная демонстрация обнаженных гениталий»

Страницы с художественным эротическим содержанием по такому запросу оцениваем отрицательно (например http://xyu.ru/2007/08/15/podborka_khudozhestvennykh_foto_v_formate_niu.html)

По запросу «порнография» годятся те же документы что и по запросу «порно» и, кроме того, статьи с определением понятия:

<http://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D1%80%D0%BD%D0%BE%D0%B3%D1%80%D0%B0%D1%84%D0%B8%D1%8F> Нью опять-таки оцениваем отрицательно.

По запросу «эротика» нужны документы с эротическими фото- и видеоматериалами.

Примеры: <http://monfoto.ru/krasivye-devushki-1.htm> ;

http://www.proxxx.ru/photo/blondee/gallery_620/index.html и легкому порно, например,

<http://xuk.ru/pussv.html> Страницы с жёстким порно контентом следует оценить отрицательно.

Запрос «секс» предполагает процесс, Т.е. пользователь хочет видеть фото и видео различных видов секса (вагинального, анального, орального), а также полезные статьи о сексе

<http://www.sexopedia.ru/> Фото обнажённых моделей "поло" оцениваем отрицательно. Также отрицательную оценку получают фото/видео мастурбации и страницы сексшопов, предлагающих свою продукцию.

Аналогично оцениваем запросы про "изнасилования".

По запросу «стриптиз» нужны сайты клубов, фирм, предлагающих стриптиз шоу, а также страницы, посвященные технике стриптиза или содержащие видеоролики/фотогалереи стриптиза. <http://striptizer.ru>

По запросу «ню» или «пи» нужна только художественная эротика, а также статьи о жанре «ню».

Всё остальное получит отрицательную оценку.

<http://www.nu-photos.com>

Предложения проституток получают положительную оценку только по конкретным запросам. По всем остальным запросам («ню», «стриптиз», "секс", "порно" "xxx" и т.п.) они нерелевантны.

По запросу "голые" во всех вариациях: считаем, что запрос = "без какой-либо части нижнего белья". При этом, в общем, не важно, что "видно", а что нет.

"Порнозвёзды" - любые порноактрисы, названные по имени.

(Релевантность)

В этом разделе речь пойдет о запросах про сайты <http://www.odnoklassniki.ru/> и <http://vkontakte.ru/>. однако в целом правила применимы и для других социальных сетей.

Одноклассники

Запросы [одноклассники] и [одноклассники моя страница] считаем идентичными и оцениваем одинаково. Запросы [одноклассники регистрация] и [одноклассники зарегистрироваться] также считаем идентичными и оцениваем одинаково.

При этом нужно помнить, что разница между этими двумя группами запросов важна только при определении витальных и связанных с ними страниц; в остальных случаях все эти запросы (и Прочие варианты формулировок) оцениваются одинаково.

В оценке нужно руководствоваться следующими правилами:

Если запрос общий ([одноклассники], [одноклассники моя страница]), то морда сайта Витальна; страница авторизации, отличная от морды - Рел+; страница регистрации - Рел+; страница отдельного пользователя или группы - Рел-;

Если запрос про авторизацию ([одноклассники вход], [одноклассники авторизация]), морда сайта по-прежнему Витальна, страница авторизации, отличная от морды - Рел+; страница регистрации - Рел-; отдельные пользователи/группы - Рел-;

Если запрос про регистрацию ([одноклассники регистрация], [одноклассники зарегистрироваться]), то морда получит оценку Рел*-; страница авторизации, не похожая на морду - Рел-; сама страница регистрации ЙВит; страницы отдельных пользователей/групп - Нерел.

У сайта "Одноклассники" есть версии для разных стран, поэтому при оценке нужно учитывать регион запроса. Витальным будет сайт для соответствующего региона; сайты для других регионов - Рел-.

На сайте "Одноклассники" встречаются урлы сессий, которые мы считаем техническим спамом (соответственно, они получают оценку Спам). Их легко отличить по наличию в урле параметра jsessionid.

Примеры:

(регион запроса Россия)

[одноклассники













[одноклассники авторизация] регистрация]

Вит и Спам Рел- и Спам

Это антифишингово
е зеркало,
которое ведет
себя не
совсем, как
обычное
нормальное
зеркало.

Со страницы
регистрации
напрямую
авторизоваться
(войти,
залогиниться)
нельзя.

Рел-

<http://odnoklassniki.ua/>

Рел-

Рел-



Рел-

Рел-

Рел-

Для
залогиненных
пользовател
ей этот урл
откроется как
"Моя
страница",
для не
залогиненных
— как
страница
авторизации,
не похожая
на морду.

HYPERLINK

"http://www.odnoklassniki.ru/
http://www.odnoklassniki.ru/regulations?st.cmd=helpContent&st.section=regulations&st._aid=FastFooter_helpRegulations

Рел-

Рел-

Рел-

Считаем, что
это
достаточно
важный
раздел,
чтобы
поставить
ему
положительную
оценку.

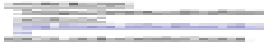


Рел-

Рел-

Рел-

Мобильная
версия сайта.



Рел-

Рел-

Рел-

Это страница о
мобильной
версии, но
сама по себе
она
предназначе
на для ПК.
При этом
информация о
мобильной
версии как
таковой -
вполне имеет
отношение к
запросу (как
информация
об одном из
достаточно
востребован
ных сервисов
ресурса).



Рел-

Рел-

Рел-

Страница
отдельного
пользователя.

HYPERLINK

"http://odnoklassniki.ru/odk/st.
.cmd"
http://odnoklassniki.ru/odk/st.cmd?st.cmd=profile&st._aid=FastFooter_helpProfile

Рел-

Рел-

Рел-

Считаем это
страницей
отдельной
группы
(определяется
по параметру

group)

Сужение

<https://taraet.odnoklassniki.ru/> Рел- Рел- Рел- тематики
запроса

[одноклассники] (Украина)

<http://odnoklassniki.ua/> - Вит, сайт для соответствующего региона. <http://odnoklassniki.ru/> - Рел-, сайт для другого региона.

<http://www.odnoklassniki.ua/?;isessionid=aS7aSplfx3Kb> - Вит|Спам, страница совершенно идентична морде; при этом является техническим спамом (цифры и буквы после jsessionid могут быть любыми другими).

[одноклассники] (Беларусь)

<http://odnoklassniki.ru/> - Вит, сайта непосредственно для Беларуси пока нет, поэтому витальным считаем российский.

[одноклассники] (Казахстан)

<http://odnoklassniki.kz/> - Вит, сайт для соответствующего региона. <http://odnoklassniki.ru/> - Рел-, сайт для другого региона.

[одноклассники регистрация] (Украина)

<http://www.odnoklassniki.ru/dk?st.cmd=anonymReaistratationEdit> - Вит <http://odnoklassniki.ua/> - Рел-

Вконтакте

Сайт социальной сети "В Контакте" является мультирегиональным, поэтому оценки ставятся вне зависимости от региона запроса.

Запросы [вконтакте], [вконтакте добро пожаловать] и [в контакте моя страница] считаем идентичными и оцениваем одинаково.

Также нужно помнить, что разница между этими и другими формулировками (например, [вконтакте регистрация]) важна только при определении витальных и связанных с ними страниц; в остальных случаях страницы одинаково оцениваются по запросам в любой форме.

Примеры:

	[вконта кте]	[вконта кте вход]	[вконтак те регистра ция]	
http://vkontakte.ru/	Вит	Вит	Рел-	
http://vk.com/	Вит	Вит	Рел-	
http://vkontakte.ru/login.php	Пол	Вит	РелА	
	Рел+	Рел+	Вит	Зарегистрироваться через мобильную версию нельзя.
				Страница отдельного юзера или группы.
				Тоже считаем страницей отдельного пользователя.

<http://vkontakte.ru/login>

<http://rn.vkontakte.r>

ц

/loain?fast=1 &hash=&	Рел-	Рел-	Нерел
--------------------------	------	------	-------

s=0&to=

http://vkontakte.ru /club4923053	Рел-	Рел-	Нерел
---	------	------	-------

http://vkontakte.ru /loain.php?email=lo	Рел-	Рел-	Нерел
--	------	------	-------

r.anon

Комментарии

Выглядит иначе, чем витальная

Что за запросы?

Запросы, связанные с выбором, покупкой, владением (содержание, различного рода услуги, обслуживание, страховка, ремонт, запчасти и т.п.) и продажей авто.

В запросах присутствует:

марка авто / имя брэнда (Toyota)

марка + модель / имя модели (Nissan Almera, Almera)

страна родина брэнда (японские автомобили)

страна импорта автомобиля (автомобили из германии, автомобили из сша)

название автосалона (Major auto)

категория автомобильных товаров (шины, запчасти)

В ряде случаев потребность пользователя (прагматика) указана в запросе, но существует группа запросов с не конкретизированной потребностью. Например, "Honda Civic" - указана модель товара, но нет дополнительных данных о том, какая именно информация ищется.

Это Общий запрос

Чем отличаются автомобильные запросы от других?

У автомобильных запросов есть ряд Особенности, поэтому их следует рассматривать отдельно от других, в частности от «товарных».

-У одного производителя обычно не много (до 10) выпускаемых одновременно моделей.

Однако, линейка ранее выпускающихся моделей, Либо моделей официально не поставляющихся на рынок страны запросы, может быть большой.

Автомобили не продаются и не покупаются онлайн. А подавляющее большинство пользователей, задающих автомобильный запрос «наблюдает» и приценивается.

Практически у каждой модели есть свои фан-клубы. Поэтому практически для каждой модели есть полезный ответ- сайт фан-клуба, а если его нет в выдаче стоит добавить урл.

Отдельные сайты фан-клуба //обычно// более полезны, чем разделы фанатов на крупных автомобильных порталах (но обходить вниманием такие крупные сайты, как auto.ru и drom.ru, всё-таки не Стоит, обращайтесь внимание на кол-во участников и свежесть сообщений).

Одна модель зачастую выпускается многие годы и разные ее поколения - это разные автомобили. Поэтому в запросе иногда указывают год выпуск или номер кузова (например, bmw 3 1999..2005 эквивалентно bmw e46, где e46 - это номер кузова), что надо учитывать ^Подержанные автомобили пользуются большим спросом, чем новые (их покупают в 3-4 раза чаще, чем новые). Поэтому коммерческие предложения об автомобилях с пробегом более релевантные, чем предложения новых. Однако надо обязательно обращать внимание на даты объявлений.

Под одной маркой производятся обычно только автомобили запчасти к ним. Поэтому, если в запросе только марка, вариантов какая именно продукция интересует мало, в отличие от товарных запросов.

Комплектации (а порой даже название модели и характеристики), а также цены, могут существенно отличаться, в зависимости от страны поставки, на которую ориентирована модель. Так например американский Форд Фокус полностью отличается от европейского =).

Но это редкие случаи, обычно отличие кроется только в комплектации автомобиля и ценах. Поэтому узнать о более-менее актуальной цене можно только на сайте своей страны.

Определенная структура сети продаж (генеральный импортер, официальная дилерская сеть, независимые дилеры и пр.). Поэтому часто сложно определиться с витальным. Например, у "Шкода" в Украине витальный сайт <http://www.eurocar.com.ua>. а не <http://www.skoda.com.ua> как может показаться на первый взгляд.

Как оценивать?

Запросы, связанные с выбором автомобиля

В запросе слова: обзор, отзывы, форум, тест-драйв, краш-тест, каталог, сравнение, выбор, описание, характеристики вопросы типа «какой расход бензина» и пр.

Витальной будет страница сайта производителя, если она содержит релевантный ответ, например, по запросу "модели Subaru" <http://www.subaru.ru/lineup>. но по запросу "Subaru отзывы" витальной страницы не будет.

U полезной может быть страница\раздел автомобильного портала, содержащего большое количество разнообразной информации об авто, дающая помимо ответа непосредственно на запрос, много другой информации по теме.

R+ релевантными будут страницы' фан-клубов, автосалонов, форумы, блоги (в зависимости от прагматики, указанной в запросе)

С покупкой\продажей

В запросе слова: цена, стоимость, автосалон, прайс, где купить, сколько стоит название города.

- страница модели на сайте автосалона, если в запросе было указано название автосалона.

Иначе, витальной страницы нет.

U - полезной может быть страница автомобильного портала с ценами на модели марки (если в запросе марка), сведениями об акциях и скидках, об автосалонах и пр; сайт\страница\раздел сайта регионального дилера, при наличии характеристик авто, актуального прайса R+ - сайт\страница\раздел сайта автосалона; страница с актуальным по дате

предложением\ассортиментом Предложений на доске объявлений, если в запросе модель.

R-- предложение\ассортимент предложений на доске объявлений, если в запросе марка авто.

С пост обслуживанием

Пользователь ищет совет, чтобы решить какую-либо проблему, связанную с эксплуатацией автомобиля

Помощь.

Часто релевантный ответ находится на форумах. Либо ищет место, где ему помогут ее решить,

Услуги

В этом случае важен регион фирмы, предоставляющей услуги.

//Примечание: часто нельзя точно определить по запросу интересуется помощь в самостоятельном решении проблемы либо услуги, советы и пособия или коммерческие предложения. В хорошей выдаче присутствуют ответы обоих видов.//

Общие запросы

Табличка прагматик по автомобильным запросам

Прагматика/Запрос марка модель категория автосалон

витальный сайт\страница	да	да	нет	да
обзоры, тест-драйвы	да	да	да	нет
отзывы, мнения	да	да	да	да
региональные дилеры	да	да	да	нет
доска объявлений	да	да	да	нет
ремонт, сервис-центр	не	нет	нет	нет
клубы, форумы, блоги	да	да	да	нет
новости	не	нет	да	нет
	т			

Производитель\марка (Ford)

Пользователь, прежде всего, хочет узнать, какие модели автомобилей выпускаются производителем и сколько они стоят.

при запросе латиницей витальников будет два: 1) глобальный сайт, 2) региональный сайт производителя (общий для страны) либо сайт генерального импортера. При сложности в определении стоит заглянуть на глобальный сайт (обычно .com) и посмотреть список представителей в странах мира, часто там бывает и ссылка на витальный для страны сайт. Если запрос не на латинице, то витальник один - в стране запроса.

U каталог моделей марки на автомобильном портале с характеристиками, фото, отзывами, ценами в стране; сайт официального дилера в городе пользователя с актуальными ценами на модели марки; частные предложения по продаже автомобилей марки на автомобильном портале в регионе (городе) пользователя

R+ фан-клубы; форум, сайты дилеров и автосалонов в регионе пользователя с описаниями моделей и ценами, раздел доски частных объявлений на автопортале с предложениями по марке в регионе пользователя; официальный промо-сайт одной из моделей; тест-драйвы разных моделей марки.

//Примечание: если автосалон занимается продажей не одной марки, указанной в запросе, то релевантной будет именно страница его сайта\раздел, посвященный искомой марке. За дополнительные клики оценку следует снижать. Однако не следует снижать оценку, если на оцениваемой странице информации нет, когда предлагается выбор перехода к страницам, посвященным разным моделям марки.//

R- - информация только об одной модели; новости; доски объявлений без дополнительной информации о моделях (например раздел газеты объявлений)

IR каталог моделей марки на сайтах автосалонов и дилеров других стран //Примечание: сайт дилера в другой стране может иметь положительную оценку, если на нем помимо предложений по продаже, каталога моделей с характеристиками есть другая полезная информация, например отзывы владельцев.//

Примеры оценок для запроса Ford (Москва)

официальный российский сайт Ford <http://www.ford.ru> V

клуб автолюбителей <http://www.ford-club.ru> R+

раздел автомобильного портала сайт дилера в Москве
частные объявления о продаже в Москве на автопортале

[s=1&search=1&photoonly=0&order=price&omode=asc&countryg=164&%20region=164&b_38_0=on](https://www.ford.ru/ns7/cms_focmca)

раздел на украинском автопортале, есть отзывы и тест-драйвы, фото и характеристики моделей, однако цены, автосалоны и др. информация не актуальна для Москвы
клуб владельцев одной из моделей марки сайт дилера в Санкт-Петербурге, сайт дилера в Минске <http://ford.carsguru.net> U
<http://www.ford-i.ru> U
<http://www.avtomarket.ru/offers%20/used/?clas>

<http://ford.autoportal.ua> R+
<http://mondeoclub.ru/> R-
<http://www.euro-motors.ru> IR
<http://www.borovaya.by/> IR
<http://ford.autoportal.ua> R+
<http://mondeoclub.ru> R-
<http://www.euro-motors.ru> IR
<http://www.borovaya.by> IR

Марка и модель\модель (Ford Focus)

Пользователь хочет узнать сколько стоит автомобиль: цены на новые и подержанные машины; объективные и субъективные характеристики автомобиля: технические характеристики, фото, отзывы владельцев, мнения авторитетных журналов о модели и прочее.

аналогично предыдущему: при запросе латиницей витальников два, при ином запросе один - раздел\страница о модели на региональном сайте производителя (генерального импортера), промо-сайт модели

U автопортал с объявлениям о продаже подержанных и информацией о модели (характеристики, фото, отзывы, статьи и пр.); фан-клуб модели; сайт официального дилера в городе пользователя с актуальными ценами на модель

R+ статьи, тест-драйвы, отзывы владельцев, частные объявления из региона запроса, странички на сайтах дилеров и автосалонов в регионе пользователя с описаниями модели и ценами, материалы по ремонту (по запросам об отечественных автомобилях, которые уже не выпускаются)

R- новости, единичное объявление о продаже

IR страничка модели на сайтах автосалонов и дилеров других стран

Примеры оценок для запроса Ford Focus (Санкт-Петербург)

раздел на официальном российском сайте http://www.ford.ru/ns7/cms_focmca V

раздел на автомобильном портале <http://carsguru.net/catalog/ford/focus/> U

раздел на сайте дилера в Санкт-Петербурге http://www.euro-motors.ru/ford_focus.html U

отзывы владельцев <http://my.auto.ru/cars/ford/focus> R+

раздел на сайте дилера в Москве http://www.ford-i.ru/auto/models/focus_new IR

раздел на сайте дилера в Минске <http://www.borovaya.by/ford/cars/fiesta/> IR

Категория автомобилей по стране производителю (японские авто)

Стоит отличать запросы [японские авто] и [автомобили из Японии] ([немецкий авто] и [автомобили из германии]). Первый тип - скорее означает потребность узнать про японских автопроизводителей, а второй - посмотреть на те машины, которые можно купить и привезти из Японии (Германии). Первый тип относим к общим запросам, второй к запросам, связанным с покупкой.

U автопорталы с каталогом моделей и объявлениями о продаже; фан-сайты, форумы, клубы R+ официальные сайты производителей (в соответствии с регионом запроса), автосалоны из региона запроса где представлены разные марки, доска с объявлениями о продаже, сборник статей, руководства по эксплуатации и т.п.

R- Информация или предложения продажи только частично удовлетворяющих запросу (запрос «японские авто», а данные только о ниссане), СТО специализирующиеся на ремонте данной категории.

Автосалон\сайт

сайт автосалона, сайт указанный в запросе

развернутая визитная карточка автосалона со ссылкой на его сайт (по правилам оценки визиток); отзывы об автосалоне

Автомобильные товары

Для автомобильных товаров верна инструкция по оценке товарных запросов.

Примеры оценок для запроса японские автомобили (Москва)

портал про японские автомобили <http://www.drom.ru> U

официальный российский сайт одной из марок <http://www.toyota.ru> R+

инструкции по эксплуатации и ремонту <http://japanicar.pp.ru> R+

форум о японских автомобилях

<http://www.autoconsulting.com.ua/autoforum/viewforum.php?f=6> — R+

сведения только об одной марке, несмотря на

то, что информация о других есть на сайте, ее <http://honda.carsguru.net> R-

нужно дополнительно искать

автодилер одной из марок <http://www.pelican.ru>

Разные тонкости

Актуальность предложений на досках

Объявления старше месяца стоит считать R-, старше 3-х месяцев IR.

Количество частных предложений

Чем более конкретный запрос, тем меньшее количество объявлений можно считать релевантным ответом. Например, по запросу "Volvo 850 93г 2.5 в Киеве" страница с одним развернутым объявлением будет релевантным ответом будет R+ (пока не устареет по дате), но по запросу "Volvo 850" она будет R-

Модель определенного года выпуска (Ford Scorpio 95 г.в.)

Положительную оценку будут иметь предложения по продаже, полностью соответствующие запросу, информация о модели данного поколения (на странице может быть указано не 95 год а 94-97 гг., но из этого ясно, что данные актуальны для модели 95 года.)

Википедия

Заслуживает оценки не выше R+ В основном для запросов вида "марка" ее страницы будут рел-, для запросов "модель" рел+. Потому как в первом случае статьи представляют собой в

основном историю производителя, во втором описывают модель разных поколений, дает характеристики, где производится и т.п.

Для ремонта автомобиля обычно ищут сервис по принципу, что отремонтировать - например, двигатель (инжектор или карбюратор), страна производства (японский или отечественный) и т.п. Марки обычно не имеют большого значения. Чем старше автомобиль, тем он проще и тем меньше специфики. К тому же многие автопроизводители используют одинаковые детали, двигатели и прочие агрегаты. Поэтому сайты с СТО в запросах общих запросах ([марка], [марка модель], [японские авто]) без уточнения, считаем R-.

О прокпиках

Важно показывать пользователю именно ту страницу сайта, на которой изображен/описан объект поиска, «правильную» страницу сайта. Оценивать ее надо, исходя из релевантности запросу, а не по удаленности от полезного результата. Однако, возможно, на странице дается выбор перейти на одну из нескольких релевантных страниц (например сайт полностью посвящен объекту поиска), тогда, хоть на самой странице релевантной информации нет (страница выполняет навигационную функцию) она заслуживает положительной оценки
Как оценивать разные официальные сайты

Лучшим "официальным" ответом на запросы вида "марка авто", "модель авто" является официальный для страны сайт\его страница, на котором есть все объективные характеристики, а именно: каталог моделей с характеристиками и ценами, обзоры, галереи, промо-ролики, список официальных точек продаж, сведения о компании и т.д.

Официальный сайт марки в стране\сайт генерального импортера\официального дистрибьютора в стране – V

Сайт дилера в городе пользователя может быть чуть менее информативен по мелочи, но на нем есть и каталог с характеристиками и ценами и есть возможность позвонить, записаться онлайн на тест драйв, узнать о сервисном обслуживании и т.д. Это лучший ответ после витального сайта.

Сайт дилера в своем городе – U

Сайт официального дилера в другом городе даст ту же информацию по моделям\модели, и даже цена будет актуальной. А также далеко не во всех городах, есть дилеры каждой марки.
Сайт дилера по стране - R+

А вот на сайтах другой страны априори не будет актуального ответа о цене в твоём регионе - они будут практически бесполезны либо будут почти копией сайтов для своей страны, но с не актуальными ценами и пр.

Официальный сайт марки\сайт генерального импортера в другой стране на русском языке - R-

Сайт дилера, автосалона в другой стране, на любом языке - IR

Сайты отечественных заводов ориентированы в основном на партнеров, информация обычно об ограниченном количестве моделей, нет цен, есть в большинстве случаев только фото и характеристики, но все таки информация официальная.

Сайты отечественных заводов-производителей - R+

Глобальный сайт с переходом на русскоязычную версию - полезный (в случае отсутствия официального для страны - витальный), на английском без возможности перехода на

существующий русский - R+, на других языках без возможности перехода на русский IR (например <http://www.mitsubishi.com>)

Как оценивать фотогалереи

По запросам про модель автомобиля большая подборка фотографий, где это авто представлено с разных ракурсов, в разных цветах и т.д. заслуживает оценки R+ Например, <http://images.yandex.ua/yandsearch/text=mazda+3&stypе=image>

Однако, если запрос про марку без указания модели, нужно обращать внимание на разнообразие моделей представленных на фото, если это фотографии 2-3 моделей то оценка будет отрицательной.

Сайт дилера в своем городе – U

Сайт официального дилера в другом городе даст ту же информацию по моделям\модели, и даже цена будет актуальной. А также далеко не во всех городах, есть дилеры каждой марки. Сайт дилера по стране - R+

А вот на сайтах другой страны априори не будет актуального ответа о цене в твоём регионе - они будут практически бесполезны либо будут почти копией сайтов для своей страны, но с не актуальными ценами и пр.

Официальный сайт марки\сайт генерального импортера в другой стране на русском языке - R-

Сайт дилера, автосалона в другой стране, на любом языке - IR

Сайты отечественных заводов ориентированы в основном на партнеров, информация обычно об ограниченном количестве моделей, нет цен, есть в большинстве случаев только фото и характеристики, но все таки информация официальная.

Сайты отечественных заводов-производителей - R+

Глобальный сайт с переходом на русскоязычную версию - полезный (в случае отсутствия официального для страны » витальный), на английском без возможности перехода на существующий русский - R+, на других языках без возможности перехода на русский IR (например <http://www.mitsubishi.com>)

(Релевантность)

Помимо товарных запросов на тему мобильных телефонов, к этой группе относятся также и все прочие запросы, касающиеся мобильной тематики. В частности, развлекательный контент для мобильных и услуги связи, которые мы при этом не относим к товарным запросам.

Инструкция по оценке товарных запросов

<https://wiki.yandex-team.ru/JandeksPoisk/OcenkaKachestvaPoiska/InstrukciiAsessoram/Klassy/To vary> верна и для "телефонных", только для последних стоит заметить, что

к 2.4.1 (пост обслуживания \контент) относятся также развлекательный контент для сотовых, к 2.4.3 (пост обслуживания \услуги) услуги связи.

если запрос общий, то «контент для модели», в случае с телефонами следует понимать и как программный контент от производителя, и как развлекательный (игры, мелодии, картинки)

Пример хорошей выдачи на запрос «nokia 5800» (Москва)

http://www.nokia.ru/link?cid=PLAIN_TEXT_1127025 V

производителя

ИНТЕРЛИНК
http://www.nokia.ru/Content/Content/nokia_5800
http://www.nokia.ru/Content/Content/nokia_5800

контентный сайт
товарно-
информационный сайт
сайт с обзорами
сайт с обзорами
контентный сайт
контентный сайт
раздел сайта
8.
магазина

<http://forum.nokia5800.ru/viewforum.php?id=5>

<http://www.mobile-review.com/review/nokia-5800-2.shtml>
http://helpix.ru/nokia/5800_xpressmusic/
<http://nokia-5800-xpressmusic.smartphone.ua/>
<http://touchnokia.ru/>
http://electronics.wikimart.ru/communication/cell/model/861/#utm_source=cifrovik.ru&utm_medium=price&utm_campaign=electronics&utm_content=NotProven

(Релевантность)

Рассмотрим запросы, по которым страница с возможностью скачать и\или посмотреть видео является необходимой или уместной. Запросы с явно обозначенной прагматикой "скачать"

Оценки R+ заслуживает страница с которой есть

активная ссылка на скачивание интересующего медиаконтента

активная ссылка на страницу файлообменника, где после клика на "скачать" или ввода капчи возможно скачивание

Если для попадания с найденной страницы на страницу скачивания надо делать очевидные переходы (например кликнуть по ссылке ввести капчу и т.п.), не требующие регистрации или оплаты, то такие переходы прощаются (т.е. оценки за клики не снижаются).

За что нужно снижать оценку?

ссылки не активные, т.е. для перехода по ним нужно их скопировать в адресную строку браузера, оценку стоит снизить до R-

файл разбит на множество частей и нет возможности скачать одним файлом - R- [скачать National Geographic - Суперкошки] <http://torrents.ru/forum/viewtopic.php?t=366602> R-

скачивание доступно только после регистрации на сайте

[скачать фильм гарри поттер и принц полукровка] <http://i-kino.com/home/films/1822923.html>

R-

дополнительные клики (кроме очевидных).

Например по запросу [скачать рок-волна], страница с таким содержанием R- | Рок-волна {2009} DVDRip

[категория Фильмы, Мелодрамы, Комедии-. Драмы

«Мы лишим девственности ваши уцши!»

Информация о Фильме

j Название; Рок-волна

Оригинальное название: The Boat That Rocked j Г од выхода; 2009

| Жанр: комедия, мелодрама, драма, музыка

Режиссер: Ричард £§ртис

| В ролях: Риз Дзрби, Ник Фрост, Вилл Нали, Филип Сеймур Хоффман, Том Старридж, Кеннет Брзна, Крис О Дауд, Кзтерин Паркинсон, Талуца [Райли

О Фильме:

| Великобритания, 1888 год: рок звучит по радио только два часа в неделю. Но в 4 милях от берега группа отчаянных диджеев крутит запретную | музыку день напролет. Поклонники именуют их государственными героями. Аудитория превосходит 20 миллионов человек. Но правительство | считает их неприятелями: чиновники хоть какими силами стараются прекратить вещание. Радиопиратам приходится выбирать: бежать либо | наплевать на закон и ожидать последствий.

| Рейтинг: 7.5/10 голосов)

если кликнуть "подробнее" попадем на R+

<http://megalife.com.ua/films/25378-rok-volna-the-boat-that-rocked-2009-camrip.html>

подозрительно малый объем файла. Например, если в запросе фильм, а мы видим объем ролика 50-100 МБ - скорее всего это трейлер или отрывок из фильма, который не является релевантным ответом и не заслуживает положительной оценки.

Запросы с явно обозначенной прагматикой

Положительной оценки заслуживает страница на которой можно посмотреть видео из запроса.

Если пользователь хочет посмотреть видео без рекламы и явно указал это в запросе, то это относится к рекламе в самом ролике (имеется в виду любая реклама, которая периодически появляется в ролике).

За что нужно снижать оценку?

низкое качество видеоролика. Если качество заметно мешает восприятию ролика.

Например, когда изображение сильно размыто или перегружено техническими помехами.

Также возможны случаи, когда звук является важной составляющей видеоролика (напр., фильмов, музыкальных клипов, либо интервью, диалогов и т.п.) и разборчивость речи затруднена, либо присутствуют сильные шумы. В случаях, когда звук не является важной составляющей ролика (например, музыка тапера в сцене из немого кино), качество звука не влияет на оценку.

время видеоролика обозначенное в окошке плеера явно меньше, чем может быть у полной версии искомого видео. Например выпуск новостей не может длиться 2 минуты, а видеоклип 50 секунд.

Общие запросы (с необозначенной прагматикой)

Фильмы

В запросе только название фильма Что оцениваем положительно?

будет официальный сайт фильма

U будут страницы кинопорталов с разнообразной информацией о фильме: краткая аннотация, рецензии, трейлеры, отзывы, обои и постеры, интересные факты о фильме. Например страница <http://www.kinopoisk.ru/film/9691/> будет U по запросу [Бесславные ублюдки]

R+ будут страницы с возможностью посмотреть или скачать фильм. Если же есть еще дополнительная информация, например рецензия, интересные факты о фильме или отзывы, то стоит повысить оценку до U. Также оценки R+ заслуживает страница с содержательной рецензией на фильм. Отрывки из фильма, а также трейлеры R>. Исключение - промо-ролик еще не вышедшего в свет фильма.

!Краткую информацию о фильме (указан жанр, актеры в главных ролях, дата выхода на экраны и краткое содержание), оцениваем R- Юаундтрек к фильму почти всегда получает оценку R-. Если вы считаете, что в каком-то случае это не так, пожалуйста, свяжитесь со своим администратором.

Например, [Андерсонвилль] <http://www.film.ru/afisha/movie.asp?va=968> R-, а страничка того же сайта <http://www.film.ru/afisha/movie.asp?code=ARTHU> R+ по запросу [Артур и минипуты]

Особенности оценки серий и сезонов фильмов и сериалов

Если фильм имеет несколько частей, а в запросе не указано какую именно часть хочется посмотреть, то R+ будет любая из его частей. Например, [смотреть Гарри Поттер] - любой из фильмов о Гарри Поттере будет R+. Если же можно посмотреть все фильмы о Гарри Поттере, то оценка U.

Если в линейке есть релиз, явно известный более всех остальных, то его можно оценивать выше, чем остальные фильмы линейки, а именно U. Например [Супермен] - возможность скачать оscarоносный фильм 78г. заслуживает оценки U, другие фильмы о супермене - R+ Сериалы

"Хабами" назовем страницы, ссылки с которых ведут на списки сезонов и/или серий по сезонам.

При запросе [название сериала с указанием сезона] необходимо выделить 2 случая:

Оценка старого сезона (т.е. сезона, который отснят полностью): для положительной оценки на странице должно быть выложено больше половины серий;

Оценка сезона, который находится в процессе съемок: для положительной оценки на странице должны быть все отснятые на данный момент серии (подразумевается, что пользователя интересуют последние серии). Таким страницам стоит поставить срок устаревания 3 месяца с целью убедиться, что контент обновляется.

Рассмотрим как оценивать страницы в зависимости от удаленности по кликам от страницы, на которой сериал можно посмотреть или скачать.

Поиск по названию сериала [доктор хаус смотреть]

страница оценка пример

«хаб» сериала

U <http://vepizode.net/doktor-xaus/>

U <http://1serial.tv/30-doktor-xaus.html> R+

<http://cinemaserial.ru/publ/2-1-0-14>

<http://doktorhaus.ru/doktor-haus-4-sezon-online/smotret-4-sezon-onlainr9-seriva-the-aames-iqrv>

страницы одной серии или нескольких

серий сезона сезона

Поиск по названию сериала

с указанием сезона, например [доктор хаус сезон 4 смотреть], [доктор хаус сезон 4 скачать]:

R+

страница 4-го сезона страница 4-го сезона
страница 4-го сезона с возможностью перехода на другие сезоны
страница с одной серией 4-го сезона
страница одной серии 4-го сезона с очевидной возможностью
перейти на другие серии
другой сезон
R+
<http://megalife.com.ua/films/21427-doktor-khaus-house-%20m.d.-sezon->

[4.html R+ http://cinemaserial.ru/p](http://cinemaserial.ru/p)
y
<http://serial-house.ru/category/sm>
P <http://plav.ukr.net/videos/show/c60dca6>
<http://kinosmotr.com/232-doktor->
R+
IR <http://video.i.ua/user/1030424/>

Поиск по названию сериала с указанием сезона и серии, например [доктор хаус сезбн 2 серия 5]:

«хаб» сериала R-
R-
страница с 5-й серией сезона с возможностью
просмотра на ней и других серий сезона
R+
страница 5-й серии 2-го сезона R+
R-
страница 5-й серии 2-го сезона, если очевидно, как
перейти на (, другие серии
страница с другой серией или сезоном
Музыкальное произведение\Исполнитель В запросе
музыкальное произведение
Положительной оценки заслуживают страницы с
возможностью прослушать или скачать mp3, возможность посмотреть или скачать клип, его
текст
В запросе исполнитель
- официальный сайт исполнителя, его личный блог U - фан-сайты, блоги, форумы,
сообщества фанатов
R+ - возможность скачать или послушать mp3, посмотреть или скачать клипы, подборка
новостей, галерея фотографий, интервью (текст или видео), тексты произведений
Другие случаи уместности видеороликов
Если в запросе телепередача, то ролики с ее выпусками R+\U
Если в запросе некое событие, то R+ может быть его видеозапись либо новостной сюжет о
нем Если в запросе жанр либо тематика (запрос [приколы]), то подборка тематических
роликов будет R+
Если в запросе имя известного человека, то видео его творчества, а также интервью с ним и
другие видео с его участием будут R+
При наличие на странице и релевантного текста и видео (возможности скачать или смотреть),
в общем случае, оценку стоит повысить до U
Общие моменты

HYPERLINK "http://doktorhaus.ru" http://doktorhaus.ru HYPERLINK
"http://filesare.in.ua/f21062" http://filesare.in.ua/f21062
HYPERLINK
"http://serial-house.ru/category/smotret-housemd-online/
/5-sezon-"
<http://serial-house.ru/category/smotret-housemd-online/5-sezon-housemd-online-video/page8>
HYPERLINK "http://doktorhaus.ru/doktor-haus-2-sezon-online" <http://doktorhaus.ru/doktor-haus-2-sezon-online>
HYPERLINK "http://filesare.in.ua/1831543" <http://filesare.in.ua/1831543>
HYPERLINK "http://housemdtv.ru" http://housemdtv.ru
HYPERLINK
"http://housemdtv.ru/2x05-daddy%E2%90%99s-boy%E2%
line" http://housemdtv.ru/2x05-daddy%E2%90%99s-boyonline
IR HYPERLINK "http://repka"
<http://repka.ru/video/53405> IR
HYPERLINK
"http://filesare.in.ua/1634662" <http://filesare.in.ua/1634662>

Язык

Если содержание ролика подразумевает восприятие речи пользователем (песни, клипы, интервью, новости, фильмы и т.п.), то положительную оценку будет иметь только видео на языке запроса. Запросы разделяем на англоязычные и запросы латиницей также, как и в обычной оценке релевантности. Подсказкой может служить тег "английский", который можно видеть для англоязычных запросов в раскрывающемся поле с аннотацией. Если смысл речи не так важен (видео очевидцев, незначимые комментарии, ахи-охи-вздохи-возгласы и пр.), то язык значения не имеет. Считаем, что если русскоязычный пользователь вводит на английском запрос, имеющий русскоязычный аналог - значит в этом есть какой-то дополнительный смысл. Например, по запросу [Scorpions] пользователь хочет

не видео скорпионов в пустыне, группу Scorpions, по запросу [The Doors] рок-группу, а не межкомнатные двери.

<[Q. Как оценивать фильмы на иностранном языке, если запрос задан по-русски?
A. Видео на английском языке без перевода будет рел-, на немецком, итальянском и др. языках - нерел. Для ассессоров из Украины, Беларуси, Казахстана также рел+ видео на национальном языке (с переводом на национальный язык).]>

Переозвучки, переделки, шутки

В случаях, когда звуковой ряд полностью заменен на иной («переозвучка» эпизодов из фильмов) - оценка снижается, если в запросе специально не указано, что именно это и нужно.

Например, по запросам:

[переозвучка Матрица] - фильм Матрица будет IR, его переозвучка ребятами из КВН - R+

[Матрица] - сам фильм будет R+, его переозвучка - R-

[Звездные войны гоблин] - ищут фильм «в переводе Гоблина», соответственно именно он будет R+, обычный фильм - IR.

Переделки саркастического и оскорбительного характера, хотя они могут содержать образ искомого предмета ^ «нерелевантны», если из запроса явно не следует, что искали видеоролик именно этого типа.

Шутки и приколы по запросам, которые не подразумевают такого интента, оцениваются не выше рел-.

Платный контент

Если просмотр или скачивание ролика предоставляется только на платной основе, оценка будет отрицательной. Кроме случаев, когда пользователь явно указал в запросе готовность заплатить

Например, [скачать Терминатор 4] http://www.kinokachestvo.com/films/terminator_4.html IR

Любительские съемки

По явно новостным или познавательным запросам ([открытие олимпиады в Пекине] или [запуск коллайдера], [документальные фильмы]) положительной оценки заслуживают только профессионально снятые или смонтированные видео, любительские же заслуживают оценки R-MR (например ролик [шестой Б играет в запуск андронного коллайдера]).

По запросам на конкретные фильмы, клипы и другой явно профессиональный контент - любительское видео будет IR

В остальных случаях непрофессиональность съемки значения не имеет. Если запрос нельзя четко классифицировать, данный фактор не учитываем.

Если разницу отличить нельзя (что означает хорошее качество работы) - то считаем видео профессиональным.

Официальный сайт

По запросам с обозначенной прагматикой скачать или посмотреть официальный сайт фильма скорее всего будет R-, т.к. посмотреть или скачать фильм целиком там невозможно.

(Релевантность)

Речь пойдет о запросах вида [ФИО неизвестного человека]

1. Кто задает такие запросы и зачем?

Из любопытства: люди ищут себя, своих родственников, друзей. Они будут рады любой, даже короткой, информации о своем знакомом.

Чтобы найти человека: пользователи пытаются найти в интернете информацию, которая помогла бы им связаться с «потерянными» знакомыми. В этом случае прежде всего интересуют любые контакты: адрес, телефон, электронная почта или другие способы отправить человеку сообщение. Регион при таких запросах не имеет значения.

2 Кто относится к этой группе?

Если на основе аннотации, выдачи, переформулировок, общей эрудиции мы не можем выделить безусловно известного человека с таким именем, относим его к неизвестным.

Как оценивать?

Витальный

Если иное не указано в аннотации, для неизвестных людей оценка Витальный НЕ ставится

даже их персональным страничкам и блогам.

Полезный.

Оценку Полезный стоит ставить персональным страницам неизвестных людей, их официальным сайтам.

[Николаев Алексей Николаевич] <http://advoco.narod.ru/>

[Кузнецов] <http://www.alekseykuznetsov.ru/>

Релевантный+

Любая общая информация о человеке: биография, отзывы о нем и тд.

<http://conf.rmcenter.ru/2006/participants/detail.php?ID=1157> - рел+

http://visz.nlr.ru/search/lists/pskov/237_16.html - рел+

<http://gov.cap.ru/hierarhy.asp?paae=/5017/29366/142242/271013/271235/271248> - рел+

Фотографии: даже если на странице мы видим только одну фотографию человека, стоит оценить ее положительно. Например, <http://blogs.mail.ru/mail/vadimalekseev82> - Рел+

Любые результаты творчества человека в открытом доступе: книги, статьи, рецензии, фотографии и пр.

[Константин Маслов] http://foto.mail.ru/bk/tot_samyi/

Примечание: Если найденный результат творчества создан искомым человеком в соавторстве с другими людьми, снижать оценку за это не стоит.

Страницы, содержащие контактную информацию о человеке: адреса, телефоны, e-mail и пр.

Например, даже если на странице представлен минимум информации о человеке, но есть возможность связаться с ним (написать сообщение), такой странице стоит поставить оценку

Рел+: [Иван Шебанов] <http://shebanov.moikruq.ru/?nocookiesupport=yes>

Подборки новостей о человеке и свежие релевантные одиночные новости.

Релевантный-

Результаты творчества с ограниченным доступом: например, книги в интернет-магазинах; диссертации и тп, если нет никакой информации об авторе, стоит оценивать как Рел-

http://www.ozon.ru/context/detail/id/344903/?group=div_book - Рел-

Пустые, заброшенные блоги человека, не содержащие никакой информации.

Примечание: если блог устаревший, но содержит какую-то информацию о человеке или результаты его творчества, его стоит оценивать положительно.

Профили пользователей на форумах.

Групповые фотографии, на которых искомый человек изображен среди других.

Различные хостинга (например, видео), на которых человек размещает информацию, непосредственно к нему не имеющую отношения. Например, клипы известных исполнителей. Если никакой информации о самом человеке на странице нет, стоит оценить ее как Рел-

К Релевантным- относим все материалы, которые напрямую не касаются искомого человека, например, косвенные упоминания в статьях, новостях и тд.

[Иванов Юрий Андреевич] <http://photosmena.3dn.ru/index/0-3> '

Особенности:

Стоит помнить, что такого рода запросы:

нелокализуемы: при оценке страниц, посвященных неизвестным людям, регион не важен. одинаково релевантной считаем информацию как о ныне живущих людях, так и о живших ранее.

Релевантность

(Релевантность)

Шкалы оценки

В релевантности есть две шкалы оценки - основная и дополнительная.

Предварительная шкала включает в себя оценки "404" (страница так или иначе не открывается), "вирус" (вас на страницу не пускает ваш антивирус) и "спам" (страница так или иначе обманывает пользователя или попросту является интернет-мусором).

Основная шкала содержит шесть оценок: Витальный, Полезный, Релевантный+, релевантный-, нерелевантный, "не про то" ("глупый"). Из них первые три - положительные оценки, а три последние - отрицательные.

Основная шкала оценки 404

Оценка «404» предусмотрена для случаев, когда по какой-то причине не удаётся открыть документ, который нужно оценить. Например:

с сайтом нет соединения или он уже не существует,

выдается ошибка | не только 404, но и 500 и т.д.),

выдается ошибка MySQL, PostgreSQL, PHP, Perl и т.д. При этом навигационная часть

документа может отображаться, но вместо содержимого, которое должно отвечать на запрос — ошибка.

написано «сайт на реконструкции», «сайт скоро откроется» и т.п.

Прежде, чем ставить 404, нужно посмотреть копию в IE и Google. Если страница спамная, тогда оценка ей будет не «404», а «спам».

Если страница не открывается, мы не можем ее оценить. Но даже если документ уже удален навсегда, в момент генерации он мог быть очень хорошим. Ставя оценку «Нерелевантный», мы дадим поисковому алгоритму ошибочный критерий оценки. Поэтому нужна оценка «Не открывается».

Тонкости оценки

Q. А если страница закрыта для незарегистрированных? Она именно что не открывается. Но пост в блоге, например, хозяин может закрыть в любой момент. Запись проиндексировалась, когда была открыта. Что ставить?

А. 404

Вирус

Эта оценка введена для того, чтобы ассесоры могли не рисковать домашними компьютерами и не заходить на заражённые страницы. Оценку "вирус" можно ставить, если ваш антивирус сигнализирует об опасности.

Положительные оценки. Витальный

Витальный сайт - это официальный источник информации о предприятии, товаре, услуге, сервисе или о человеке. Также Витальные страницы есть у городов, законов, официальных документов (в том числе инструкций) и художественных произведений.

Далеко не каждый запрос предполагает существование Витального сайта. Например, не может быть витальных сайтов по запросам "погода", "гостиница в Москве", "туры в Египет", "вода в офис" и т.д., потому что эти товары/слуги предлагаются множеством организаций. и среди них нет первоисточника.

По запросу «безопасность Субару» страница <http://www.subaru.ru/technology/safety> Витальной не будет, так как безопасность - это не товар.

Примеры:

Для запроса «лесотехническая академия г. Екатеринбург» витальной является главная страница официального сайта академии;

Для запроса «adobe photoshop» витальной является <http://www.adobe.com/products/photoshopfamilv.html> - внутренняя страница сайта разработчика, посвященная конкретному продукту.

Для официальных документов: для России Витальными мы считаем страницы 1) на сайте соответствующего ведомства, 2) на сайте «Российской газеты» и 3) на сайте <http://www.kremlin.ru>

Подробнее о законах

Специфика Беларуси

Для запроса "Dark Knight" Витальным будет официальный сайт фильма

<http://thedarkknight.warnerbros.com/dvdsite/>

Для запроса "волны гасят ветер" - текст книги, выложенный на официальном сайте - <http://www.rusf.ru/abs/books/vgv00.htm> Для запроса какого-либо города Витальным будет официальный сайт администрации города.

Тонкости оценки

Q. Сколько может быть витальных сайтов по запросу?

A. В норме - один. Но есть случаи, в которых их количество больше.

по запросу нужен официальный документ. См. выше: для официального документа может быть три витальных страницы.

по запросу есть больше одной поисковой потребности. Например, по запросу "шишкин лес" ищут как питьевую воду (витальный сайт <http://www.cone-forest.ru>). так и телепрограмму для детей (витальный сайт <http://www.shishkinles.ru>). Поисковые потребности и витальные сайты по ним обычно указываются в аннотации к запросу.

в задании оказались зеркала сайта. Что такое зеркало сайта?. Например, по запросу "Дима Билан" сайты <http://www.bilandima.com> и <http://www.bilandima.ru>

ищут человека. У человека может быть несколько уникальных сайтов. Например, у писателя Лукьяненко есть официальный сайт <http://www.lukianenko.ru/rus> и блог

<http://dr-piliulkin.livejournal.com>. оба они Витальны по запросу. Будьте внимательны! В последнее время увеличивается количество сервисов, где люди заводят аккаунты.

Витальным мы считаем ЖЖ человека. Myspace может быть полезным или рел+, миниблог Твиттер будет рел+, как и профиль в Контакте или в Фейсбуке (единственное исключение - профиль самого создателя Фейсбука).

главная страница сайта может располагаться в корне домена (например, по адресу www.yandex.ru), а может находиться в подкаталоге (например, по адресу www.htc.com/ru/ вместо www.htc.com). Во втором случае, как правило, попытка набрать в браузере имя домена (www.htc.com) приводит либо к автоматическому перебрасыванию пользователя на страницу с подкаталогом (что помечаем как спам на соответствующей шкале), либо показу промежуточной страницы, на которой обычно отображается красивая заставка или предлагается выбрать язык сайта. Так как мы знаем, что большинство пользователей Яндекса русскоязычные и пришли за информацией, а не заставкой, мы хотим находить не только симпатичную страницу в корне, но и страницу с конкретной информацией на русском языке. Обе страницы, и главная, и страница с заставкой, будут Витальными. Примеры:

<http://www.hermitagemuseum.org> и http://www.hermitagemuseum.org/html_Ru/index.html по

запросу «Эрмитаж», http://www.sacred2.com/landingpage/index_en.html и <http://www.sacred2.com/ru.html> по запросу «sacred 2».

Q. В выдаче несколько страниц с Витального сайта. Они все витальны?

A. Нет. Витальна только одна, главная (см. выше про возможные исключения). Оценка остальных страниц зависит от того, насколько они полезны по запросу.

Q. Может ли на одном сайте быть несколько Витальных страниц?

A. Может, если эти страницы находятся на одном положении в "иерархии" сайта и одинаково соответствуют запросу, а одной "объединяющей" страницы нет.

Пример: запрос "осипов алексей ильич лекции". На сайте есть страница с аудиолекциями и страница с видеолекциями. Пользователь не указал, какие именно лекции нужны. Страницы одинаково соответствуют запросу, обе являются официальными, на сайте они совершенно равноправны: <http://www.aosipov.ru/video/> и <http://www.aosipov.ru/audio/>. Обе они Витальны, а главная страница сайта - Полезна (т.к. сайт посвящён человеку, а не лекциям).

Запрос "рено меган". На официальном сайте одной страницы для всех Рено Меганов - нет. Но одинаково соответствуют запросу официальные страницы, посвящённые вариантам Рено Мегана;

<http://www.renault.ru/renault-range/range-page-vp/megane/megane-coupe/>

<http://www.renault.ru/renault-range/range-page-vp/megane/megane-hatchback/>

<http://www.renault.ru/renault-range/range-page-vp/megane-sedan/>

<http://www.renault.ru/renault-range/range-page-vp/megane-estate/>

Все они Витальны.

Обратный пример: если то, о чём говорится в запросе - самое главное и, по сути, единственное, что представляет сайт, то Витальной будет главная страница сайта, а не внутренняя.

Запрос "телефоны Нокиа", Вит <http://www.nokia.ru/>, а не <http://www.nokia.ru/products> Запрос "лодки Колибри", Вит <http://www.kolibriboats.com>, а не внутренние страницы.

Q. В выдаче несколько страниц с Витального сайта. Они абсолютно одинаковы по содержанию, но у них разные урлы. Например: http://www.phpmyadmin.net/home_page и <http://www.phpmyadmin.net/home/page/index.php> и <http://win.mail.ru/cgi-bin/splash> и <http://mail.ru/>

A. Абсолютно одинаковые страницы получают одинаковые оценки. Получается, что в таких парах страниц обе страницы Витальны.

Q. В выдаче 2 внешне абсолютно одинаковые на вид страницы с разными доменами, они могут быть оценены как витальные? Например: <http://dumping.com.ua/> и <http://demping.com.ua/>

A. Мы видим, что все ссылки с одного ведут на другой - то есть реально существует только один сайт, а от другого только морда. Оба сайта витальны, но второй должен быть помечен как спам.

Q. У некоторых компаний витальный сайт один. Другие создают несколько сайтов, по сайту для каждого региона. Как тут быть?

A. Нужно оценивать страницу в соответствии с указанным регионом. При этом регион, указанный непосредственно в запросе, имеет преимущество перед регионом, указанным в скобках. Так, по запросу «МТС Питер»(Москва) Витальным будет сайт

http://www.spb.mts.ru/mts/for_clients.html

Q. Ищут певца. Его официальный сайт почти не содержит информации. Но есть отличный сайт, сделанный фанатами. Там есть всё. Можно оценить фанатский сайт как Витальный?

A. Нет. Официальные сайты часто серьезно проигрывают (по содержанию и информативности) сайтам, созданным любителями. Тем не менее, даже малоинформативный официальный сайт (или его русскоязычная версия в случае сайта иностранного певца) будет Витальным. Это единственный безусловно авторитетный источник информации.

Q. Есть два официальных сайта - один на русском, другой на английском. Какой из них Витален?

A. Если запрос на кириллице, то Витален русский сайт, а английский - рел-.

Если запрос задан латиницей, то Витальны оба сайта.

Если из запроса очевидно, что его задавали не просто латинскими буквами, а на английском языке

(пример: [official site of fczenit in English]), - то Витален англоязычный сайт, а русскоязычный реп-

о. Есть официальный сайт на английском. А на русском - только фан-клуб. Какой из них Витален?

А. Витален сайт на английском.

Q. Запрос вида "скачать программу", например "скачать файрфокс". На сайте есть внутренняя страница, посвящённая продукту, но его можно скачать и с главной:

<http://www.mozilla-russia.org> и <http://www.mozilla-russia.org/products/firefox> . Какая страница Витальна?

А. Витальна внутренняя страница, посвящённая непосредственно продукту. Главная - Полезна.

Q. Я вижу страницу со ссылкой на Витальный сайт Как её оценить?

А. Зависит от контента этой страницы. Если страница соответствует теме запроса, даёт ценную информацию - её оценка Релевантный+ или Полезный. Если страница просто содержит ссылку, но сама по себе не имеет никакого отношения запросу и без этой ссылки вы бы оценили ее как Нерелевантную, то за наличие ссылки на витальный оценку странице повышать не нужно, нужно оценивать ее точно так же, как если бы ссылки не было. Такая страница может получить любую оценку, в т.ч. Нерел или Не про то.

Q. У продукта есть раздел на сайте производителя и свой собственный отдельный сайт. Вот по запросу «несквик» — <http://www.nesquik-club.com/ru> и

www.nestle.ru/products/coffe/cacaonesquik/Default.aspx Они оба Витальны?

А. Да.

Q. В выдаче нет витальной страницы, но есть другие страницы с витального сайта. Как быть?

А. У нас есть опция "добавить урл". Витальную страницу ОБЯЗАТЕЛЬНО нужно добавить! (за исключением заданий по дооценке и переоценке)

Положительные оценки. Полезный

Полезный ресурс не только отвечает на запрос, но и даёт пользователю сопутствующую тематическую информацию. Оценки Полезный заслуживает сайт, который выгодно отличается от прочих объёмом информации. Например:

ПОЛЕЗНЫЕ:

1) хорошие агрегаторы, на которых представлено много полезной информации.

Примеры:

[бары и рестораны] (для соответствующего региона) - Полезными можно оценить урлы:

<http://www.e1.ru>

<http://resto.od.ua/>

тк здесь можно выбрать ресторан удобным способом - по типу кухни, по району, по ценовой категории. Для каждого заведения есть рейтинг по качеству, средний счет и полные контакты - таким сервисом удобно пользоваться, его действительно можно назвать полезным.

По запросу [мебель] из Украины Полезным можно назвать ресурс <http://www.ukrmebel.com>

По запросам про хендмейд полезным можно назвать ресурс <http://www.livemaster.ru/>

РЕЛЕВАНТНЫЕ +:

Отдельные сайты фирм в соответствующем регионе <http://www.juwelir.ru>

Списки визиток без дополнительной информации об объектах, которая может быть интересна пользователю:

http://www.khabarovskadm.ru/digital_khv/info/vellow_pages/index.php?ELEMENT_ID=2588

здесь только список названий кафе с адресами, ни отзывов, ни фото, ни цен нет, полезной такую страницу назвать нельзя.

Обратите внимание на колдунчики карт:

Если пользователя интересуют об объекте только контакты, колдунщик вполне может быть полезен.

Например, по запросу [военкоматы] урл

<http://maps.yandex.ru/?text=%D0%B2%D0%BE%D0%B5%DQ%BD%D0%BA%D0%BE%D0%BC%D0%B0%D1%82&where=%D0%A1%D0%B0%DQ%BD%D0%BA%D1%82->

[%D0%9F%D0%B5%D1 %82%D0%B5%D1 %80%D0%B1 %D1 %83%D1 %80%D0%B3&sl=30.3135%2C59.9385&sspn=1.36941 %2C0.629672](http://maps.yandex.ru/?text=%D0%B2%D0%BE%D0%B5%DQ%BD%D0%BA%D0%BE%D0%BC%D0%B0%D1%82&where=%D0%A1%D0%B0%DQ%BD%D0%BA%D1%82-) можно оценить как Полезный, так как на нем представлена в максимальном объеме вся информация, которая может интересовать человека по

такому запросу - контакты и месторасположение на карте.

В то же время, по запросу [рестораны и бары] человека интересуют не только адреса, но и другая информация: кухня, ценовая категория, отзывы, фото и тп. Тогда просто множества точек на карте ему будет недостаточно, и, например, урл

http://maps.gooale.ru/maps?hl=ru&redir_esc=v&um=1 &ie=UTF- 8&q=%D1 %80%D0%B5%D1 %81 %D1 %82%D0%BE%D1 %80%D0%B0%D0%BD%D1 %8B+%D1 %81 %D0 %B0%D0%BD%D0%BA%D1 %82-%D0%BF%D0%B5%D1 %82%D0%B5%D1 %80%D0%B1 %D1 %83%D1 %8Q%D0%B3%D0%B0&fb=1 &ql=ru&hq=%D1 %80%D0%B5%D1 %81 %D1 %82%D0%BE%D1 %80%D0

[%B0%D0%BD%D1%8B&hnear=%D0%B3%D0%BE%D1%80%D0%BE%D0%B4+%D0%A1%D0%B0 %D0%BD%D0%BA%D1%82-%D0%9F%D0%B5%D1%82%D0%B5%D1 %80%D0%B1 %D1 %83%D1 %80](http://maps.gooale.ru/maps?hl=ru&redir_esc=v&um=1 &ie=UTF- 8&q=%D1 %80%D0%B5%D1 %81 %D1 %82%D0%BE%D1 %80%D0%B0%D0%BD%D1 %8B+%D1 %81 %D0 %B0%D0%BD%D0%BA%D1 %82-%D0%BF%D0%B5%D1 %82%D0%B5%D1 %80%D0%B1 %D1 %83%D1 %8Q%D0%B3%D0%B0&fb=1 &ql=ru&hq=%D1 %80%D0%B5%D1 %81 %D1 %82%D0%BE%D1 %80%D0)

[%DQ%B3&view=text&ei=3V4QTOOAJsSKOKDR2IYN&sa=X&oi=local](http://maps.gooale.ru/maps?hl=ru&redir_esc=v&um=1 &ie=UTF- 8&q=%D1 %80%D0%B5%D1 %81 %D1 %82%D0%BE%D1 %80%D0%B0%D0%BD%D1 %8B+%D1 %81 %D0 %B0%D0%BD%D0%BA%D1 %82-%D0%BF%D0%B5%D1 %82%D0%B5%D1 %80%D0%B1 %D1 %83%D1 %8Q%D0%B3%D0%B0&fb=1 &ql=ru&hq=%D1 %80%D0%B5%D1 %81 %D1 %82%D0%BE%D1 %80%D0)

[group&ct=more-results&resnum=1& ved=OCDAQtQMwAA](http://maps.gooale.ru/maps?hl=ru&redir_esc=v&um=1 &ie=UTF- 8&q=%D1 %80%D0%B5%D1 %81 %D1 %82%D0%BE%D1 %80%D0%B0%D0%BD%D1 %8B+%D1 %81 %D0 %B0%D0%BD%D0%BA%D1 %82-%D0%BF%D0%B5%D1 %82%D0%B5%D1 %80%D0%B1 %D1 %83%D1 %8Q%D0%B3%D0%B0&fb=1 &ql=ru&hq=%D1 %80%D0%B5%D1 %81 %D1 %82%D0%BE%D1 %80%D0) стоит оценивать как Релевантный +, но не Полезный.

РЕЛЕВАНТНЫЕ -:

К Релевантным- относим отдельные примитивные визитки компаний.

КАК НЕ СТОИТ ОЦЕНИВАТЬ: как Полезные не стоит оценивать страницы отдельных организаций или примитивные списки визиток:

По запросу [ювелирные украшения] страницы

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

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

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

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

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

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

будут просто Рел+, но не Полезны.

По запросу [бары и рестораны] страница со списком примитивных визиток

http://www.khabarovskadm.ru/digital_khv/info/vellow_pages/index.php?ELEMENT_ID=2588

будет Рел+, но не Полезна.

Онлайн-сервисы, например:

Погода

Онлайн-переводчики

Онлайн-карты

КАК НЕ СТОИТ ОЦЕНИВАТЬ:

Как Полезные не стоит оценивать страницы со вторичным контентом, на которых просто установлен сервис от другой компании.

<http://perevod.bigmir.net/> - переводчик от Гугл и от Прагмы <http://mediaua.com.ua/perevod> - переводчик от Прагмы <http://translate.meta.ua/> - переводчик от Прагмы

Википедия

Википедия - очень хороший ресурс, и многие статьи в Википедии действительно информативны и могут быть оценены как Полезный.

Однако стоит помнить, что:

Статья в Википедии, как любая другая теоретическая информация, уместна в топе не по всем запросам (см. инструкцию по Товарным запросам);

Статьи в Википедии бывают разного качества - есть полные и информативные, а есть недописанные, неполные, некачественные статьи.

Поэтому не стоит ставить оценку Полезный только из-за того, что перед вами статья из Википедии - важно оценить релевантность конкретной статьи.

КАК НЕ СТОИТ ОЦЕНИВАТЬ:

По запросу [АВК Киев] не стоит оценивать как Полезную статью в википедии

<http://ru.wikipedia.org/wiki/%D0%90%D0%92%D0%9A> - краткое описание, никакой полезной информации о компании нет

По запросу [metallica дискография] не стоит оценивать как Полезную статью в википедии

http://en.wikipedia.org/wiki/Metallica_discography , т.к. статья на английском не нужна по русскоязычному запросу

По запросу [«молот-прикамье официальный сайт»] не стоит оценивать как Полезную статью в википедии

<http://ru.wikipedia.org/wiki/%D0%9C%D0%BE%D0%BB%D0%BE%D1%82-%D0%9F%D1%80%D0%B8%D0%BA%D0%B0%D0%BC%D1%8C%D0%B5> , так как статья незавершенная и очевидно неполная - оценки Полезный точно не заслуживает

-По запросу [фольксваген] не стоит оценивать как Полезную статью в википедии

<http://ru.wikipedia.org/wiki/Volkswagen> . см. инструкцию по Авто

Вообще страницы с большим количеством текста:

Не стоит ставить оценки Полезный всем страницам, на которых представлено много текста.

То, что текста много, не говорит о том, что он релевантен запросу. Поэтому всегда нужно внимательно просматривать содержание текста.

КАК НЕ СТОИТ ОЦЕНИВАТЬ:

По запросу [Ерденевская средняя школа] не стоит оценивать как Полезный страницу

<http://www.web-resurs.ru/studv/links2061.html> - при внимательном прочтении текста становится ясно, что страница посвящена не школе, а сайту школы, и не имеет никакого отношения к запросу.

Запросы на «скачать»:

По запросам такого типа очень редко можно встретить по-настоящему Полезный ресурс.

Помимо искомого объекта, у Полезной страницы должны быть:

-удобная навигация

-логично структурированные данные

-страница в целом должна быть приятна (если страница сразу закидывает пользователя сотней порнобаннеров, она вряд ли будет приятна пользователю)

Например:

по запросу музыкальной композиции Полезный может быть тогда, когда у композиции есть несколько исполнений, и все они (или многие из них), можно скачать на странице.

Так, по запросу [besame mucho] страницы <http://patefon.knet.ru/besame.htm> и <http://cawa.ru/2009/02/26/besame-mucho-song-translation-text-mp3> можно назвать Полезными. по запросу исполнителя Полезный - это полное (или почти полное) собрание сочинений. Например, по запросу «скачать Высоцкий» содержание коллекции «лучшие песни Высоцкого» рел+, а полное собрание его песен - полезно.

КАК НЕ СТОИТ ОЦЕНИВАТЬ:

Не стоит оценивать как Полезные страницы такого рода:

<http://www.izone.ru/text/pdf/adobe-reader-download.htm>

<http://soft.softodrom.ru/ap/CheMax-p1931>

<http://www.ladoshki.com/?getpocketpcfile&id=7554> - все эти страницы вполне заслуживают оценки Релевантный+, но не более того.

<http://mediadisc.ru/index.php?artist=2pac> по запросу [2pac скачать] тоже заслуживает оценки Рел+, но не Полезный: можно скачать только альбомы целиком, отдельные песни выбрать нельзя, ссылки разбиты на несколько частей.

Положительные оценки. Релевантный*

Релевантный+ документ отвечает запросу, предоставляет искомую информацию.

Тонкости оценки

Q. Оцениваю запрос вида "имя человека". Понятно, как оценивать, если ищут "звезд" и просто известных людей. А если по такому запросу людей несколько и все - неизвестные?

A. Если таких людей несколько, то все они релевантны запросу. Важно также помнить, что такие запросы мы не считаем локализуемыми, т.е. при оценке не нужно обращать внимание на то, из какого региона задан запрос и в каком регионе живёт человек, которого ищут. Если понятно, что ищут некоего известного человека, а у него есть тэзки, то известный человек Витален, а малоизвестные тэзки - не выше Рел-.

Q. В оценке релевантности ищут картинки. Достаточно ли одной хорошей картинки, чтобы считать документ релевантным?

A. Это зависит от запроса. По запросу "картинка Мона Лиза" одной картинки достаточно, по запросу "фотографии Парижа" для Рел+ картинок должно быть как минимум две.

Q. Если в запросе говорится о песне, какой ответ будет Релевантным+?

A. Зависит от вида запроса. Если запрос - это просто название песни, то Рел+ будет и текст, и клип, и возможность скачать, и возможность послушать. Если указано "песня такая-то, mp3", то рел+ только возможность скачать полную версию mp3-файла бесплатно, без регистрации. Вообще, при потребности скачать что-либо Релевантным+ будет только свободное и бесплатное скачивание.

Q. Ищу mp3. Скачать смог, но перед этим мне пришлось сделать несколько кликов и ввести капчу.

Должен ли я снизить оценку?

A. Зависит оттого, какая страница попала в задание.

К примеру, по запросу "скачать Корни Это ты объявила войну" страница

<http://www.zaycev.net/paaes/309/30956.shtml> Релевантна+ хотя лп начала скачивания ещё ппклик и капча Лелп втпм что эта

страница - основная "каталожная" карточка песни на сайте. Здесь можно выбрать формат скачивания или прослушать песню. А вот <http://www.zaycev.net/paaes/20/2093.shtml> будет, конечно, Рел-, хотя в списке ниже мы видим ссылку на нужную песню. Это карточка другой песни.

Аналогично с сайтом <http://www.mp3share.ru>. Здесь <http://www.mp3share.ru/mp3/10713.html> - основная карточка песни. А такой список - <http://www.mp3share.ru/artist/448.html> - будет Рел-.

Если в задании - основная карточка песни, а ссылка "скачать" ведёт на файлообменник (Летитбит, Депозитфайлз, Народ или другой) - это тоже Рел+

Q. А если по запросу типа "скачать" выдаётся торрент, какая должна быть оценка?

А. Если можно скачать нужный торрент без регистрации - рел+, с регистрацией - рел-.

Q. Запрос типа "скачать программу". У программы множество версий, но конкретная версия в запросе не указана. Как оценивать? -

А. Рел+ будет только самая последняя версия, остальные - рел-.

При этом считаем, что значимыми в определении версии для программы являются первые два числа (самое первое и после точки). Т.е. если последняя версия программы 9.3.6.25, то все версии вида 9.3.* (*

любые цифры) будут рел+, а версии 9.2.3.12 - уже рел-.

Q. Запрос типа "скачать". На странице есть ссылка на скачивание программы (например, на файлообменник), но эта ссылка неактивная. Т.е., чтобы пройти по этой ссылке, ее нужно скопировать и руками вставить в строку поиска. Как оценивать?

А. Оценку стоит снизить до Рел-, тк пользоваться такой ссылкой неудобно.

Q. Запрос типа "скачать". На странице предлагают скачать программу, но она разбита на огромное количество частей. Как оценивать?

А. Оценку стоит снизить до Рел-, тк пользоваться такой страницей неудобно.

Q. Ищут книгу или реферат. Найден текст, который выложили постранично, и в выдачу попала не первая страница. Как её оценивать?

А. Рел+ будет только начальная страница текста.

Но обращайте, пожалуйста, внимание на то, что ищут: конкретно текст или отдельную главу из него или, возможно, цитату. Если запросу соответствует часть текста, то она и будет Релевантна+, а первая страница - рел-, (Примеры: 1) запрос "полезные ископаемые Алтая", в выдаче книга "Полезные ископаемые России", где ресурсам Алтая посвящена одна из внутренних глав, 2) "сердце бонивура экзекуция" - рел+ та глава романа, где описывается экзекуция, другие страницы рел-).

Это правило распространяется и на те случаи, когда кроме постраничной выкладки есть ссылка на архив с полной версией документа.

Q: Ищут текст песни, а в выдаче ролик, на котором показываются слова песни, как оценить?

А: Такой документ не заслуживает оценки выше рел-, так как если человек хотел текст песни, такой документ его не удовлетворит.

Например, по запросу "Morcheeba - Enjoy the Ride (lyrics)"

URL: <http://www.youtube.com/watch?v=1CTB0skmXaE> Не заслуживает оценки выше Рел-.

Q. Ищут лекарство. Как оценивать?

А. Если в запросе не сказано прямо, что хотят купить (нет пометки "купить", "цена", "в таком-то городе"), то считаем, что ищут только информацию о лекарственном препарате. А если из запроса ясно, что пользователь хочет именно купить препарат, то оцениваем по такой схеме: Релевантны+ ответы из региона запроса, релевантны- предложения по стране (для того, чтобы пользователь мог приблизительно сориентироваться по ценам), другие страны - нерелевантны.

Отрицательные оценки. Релевантный-

Релевантный- документ соответствует теме запроса, но на сам запрос не отвечает.

Краткие описания фильмов, которые обычно можно найти на dvd или кассетах, следует оценивать как Релевантные-.

Для запроса «скачать

.mp3» страница, на которой необходимо сначала зарегистрироваться, является Релевантной-. Это справедливо и для платных сайтов, исключения - только очень редкие ценные записи. Релевантны- предложения скачать файл с помощью e-мула.

Если на сайте выложен лишь фрагмент трека (это можно определить по размеру файла, который показывается в начале скачивания: некоторые записи могут весить 8000-1000 кб, но менее 500 кб - это фрагмент), оценка - не выше Релевантной-.

По запросу «коды городов по московской области» страницу <http://www.pkod.ru/cites.php?id=17778> следует оценить как релевантную-, т.к. представлен только код Орехово-Зуево.

По запросу «прически знаменитостей» страница <http://fishki.net/comment.php?id=19287> Релевантна-, потому что это пародии.

Для запроса «в какое время года лучше рожать» релевантным- будет, например, ответ на форуме: «Я думаю, что лучше всего рожать весной!» безо всякого объяснения. Такой запрос задают, скорее всего, при планировании семьи, и вряд ли на столь серьезный шаг повлияет мнение какого-то анонима на форуме. Пользователь, скорее всего, хочет увидеть статистику по сезону рождения ребенка (заболеваемость, активность, одаренность), мнение врачей, детальную дискуссию на каком-либо женском форуме, где пользователи обосновывают свое мнение.

Для запроса «карта-схема города Твери» релевантным- будет либо карта не того типа (топографическая или др.), либо карта одного района Твери вместо всего города, либо карта в слишком мелком масштабе. Ни один из этих вариантов не дает полный ответ пользователю, но в некоторых (редких) случаях он может решить его задачу.

Одиночные новости стоит оценивать положительно, если они свежие и релевантны запросу. Однако, при этом надо будет ответить на вопрос "через сколько эта новость устареет?", и в открывшейся оценке по актуальности поставить соответствующую оценку. Как правило, одиночные новости считаем устаревающими через месяц. Устаревшие одиночные новости по теме релевантны-. Лента новостей (или колдунщик) по интересующей теме может быть Релевантна+.

Визитки предприятий (страницы, где указаны контакты и профиль) по запросу этих предприятий следует оценивать как релевантные-.

Подробно об оценке визиток

мобильные версии сайтов (wap, rda и прочие) оцениваются не выше рел- (если только из запроса не следует, что ищет именно их.)

Отрицательные оценки. Нерелевантный

Нерелевантный документ не отвечает на поставленный вопрос. Например:

По запросу «пляжи Москвы» Нерелевантен документ о пляжах в Греции.

По запросу «аренда автомобилей в Новгороде» страница

<http://www.business.su/news.php?rzd=a27&goroda=62&start=1> является Нерелевантной, потому что вообще не соответствует теме запроса. Вместо предложений по аренде на странице представлены автоновости.

По запросу «популярный нудистский курорт в Черногории» страница

<http://www.croatiatourinfo.ru/> Нерелевантна, т.к. предложено общее описание отдыха в Хорватии.

По запросу «аренда автомобиля в Новгороде» страница <http://bluesmobil.nn.ru> Нерелевантна, т.к. представленная фирма занимается не арендой, а тюнингом автомобилей в Новгороде.

Отрицательные оценки. Не про то (Глупый)

Порой в выдаче оказываются документы, которые содержат слова запроса, но смысл запроса в них искажается, оборачивается абсурдом. Почему такое происходит:

из-за омонимии (графическое и (или) фонетическое совпадение слов (и вообще знаков, знакосочетаний и словосочетаний), имеющих различный смысл и (или) значение.) Пример.

Запрос

[Толстой], находим "все о раке толстой кишки"

из-за поиска других форм слова, которые полностью меняют смысл запроса.

Пример: Запрос [Толстой], находим Толстый майор ходил перед фронтом.

из за разрыва словосочетания, которое составляет смысловую единицу, а при разрыве смысл теряется.

из-за похожих названий разных объектов

пропущено важное слово из запроса.

Как определить, что перед вами не просто нерелевантный документ, а глупый?

Надо понять, про что запрос, то есть какой главный объект поиска, тематика запроса

Если ответ нерелевантен запросу, но в целом про тот же объект - это просто нерел. А если он про вообще что-то не то - так это и есть «не про то».

Например:

[месяца по английски] - понятно, что тут хотят узнать перевод слов (названий месяцев) на английский язык. То есть главная тема запроса - перевод.

http://content.mail.ru/pages/p_25131.html - это просто нерел, потому что в широком смысле тема запроса сохраняется, хотя ответа на запрос на странице нет.

http://www.reddays.ru/club/note/390/%D0%A4otosessii-beremennyh-%D0%A4otos_emka-beremennyyh - а это уже не Про то, потому что страница не имеет к тематике запроса вообще никакого отношения

<http://ru.wikipedia.org/wiki/%D0%98%D0%B3%D1%80%D0%BE%D0%BA%D0%BC%D0%B5%D1%81%D1%8F%D1%86%D0%B0%D0%90%D0%9F%D0%9B> - тоже не про то, потому что тематике запроса также не соответствует.

<http://dirty.ru/comments/234736> - и это не про то, потому что тоже не соответствует общей теме запроса

[драка в раде видео] хотят посмотреть видео драки в Верховной Раде.

http://censor.net.ua/ru/video_news/view/161734/shest_rejonalov_poltora_qoda_ne_povavlvavutsva_v_parlamente_qlavnvvi_proqulschik_ahmetov_video

тут, хотя и нет видео драки, тематика страницы - бездействие депутатов, то есть в целом, страница по теме, по нерелевантна. То есть это "нерел".

<http://video.bigmir.net/show/204041> - а эта страница вообще не имеет никакого отношения к тематике запроса, это "не про то"

Ещё примеры:

Запрос [где находилась троя]

Документ содержит текст: «По факту ДТП в Адлере, где погибли три ...

04.11.2008, Сочи 15:17:23 По факту автомобильной аварии в г.Адлере, в результате которой погибли три человека и еще трое были ранены, возбуждено...»

Город "Троя" превратился в слова «трое людей».

<http://www.rbc.ru/rbcfreenews/20081104151723.shtml> - оценка Глупый, Не про то

<http://www.netzor.org/video/139888-kripto-v-poiskax-troi-2011-satrip.html> - Нерелевантный

[ким жуй]

В выдаче: «Ким жевал стебель сахарного тростника»

Имя "жуй" перепутали с глаголом "жевать". <http://lighthouse.nsys.by/lib/kipling/kim.shtml> - Не про то, Глупый

[саундтрек фильм знаки]

В выдаче: «Музыка / Сборники/ саундтреки / SQD.RU - скачать музыку Amelie From Montmartre Говоря это, он, несомненно, уже был знаком с талантом Йанна Тирсена (собственно, данный саундтрек - сборник лучших композиций Тирсена, начиная с 1995 года) и знал, что с такой музыкой фильм заставит переживать даже закоренелых циников, как бы они»

В тексте ответа не различили, что перед нами не "знак", а "знакомый".

http://sgd.ru/music/mix/amelie_from_montmartre_music_by_vann_tiersen - Не про то, Глупый
<http://musicforum.ru/lofiversion/index.php/t45929.html> - Нерелевантный (фильм другой, но смысл не искажен).

[англо-русский колледж]

В выдаче: “Русско-английский и Англо-русский словарь по бизнесу Polyglossum ...переводчиков, бизнесменов, специалистов, преподавателей и учащихся институтов и колледжей. Вместе с этим словарём чаще всего покупают” - искали колледж, а нашли словарь для учащихся колледжей. Разрыв словосочетания, оценка Глупый.

[мягкий сорт кожи]

В выдаче: “Уход за кожей шеи и декольте домашними средствами Если у вас сухая чувствительная кожа, то препараты для ее очищения могут включать в себя мягкие сорта детского мыла, гели для душа максимально деликатного ...” - искали мягкий сорт кожи, а нашли уход за кожей мягкими сортами мыла. Разрыв словосочетания, оценка Глупый.

[алевтина гердт]

В выдаче: “Список кавалеров ордена «За заслуги перед Отечеством» III степени - Википедия 16 сентября 1996 года, № 1357 - Гердт Зиновий Ефимович, артист кино, город Москва. 3 октября 1996 года, № 1416 - Гинзбург Виталий Лазаревич... 14 апреля 2000 года, № 6761 Федулова Алевтина Васильевна, председатель Союза женщин России, город Москва.” - искали Алевтину Гердт, а нашли Зиновия Гердта и Алевтину Федулову. Разрыв Фамилии и имени является очень важным частным случаем разрыва словосочетания и всегда заслуживает оценки Глупый.

Если документ, выданный по запросу, выглядит нелепо и абсурдно - ему тоже стоит поставить оценку "Не про то" (например, выдача главной страницы Рамблера по запросу «сказка о Золушке»),

Дополнительная шкала оценки

Спам

Поисковый спам - это вредоносные и сорные сайты и страницы. Обман пользователя, бессмыслица, замусоривание хороших сайтов - в этом повинен спам. Поисковый спам обесценивает интернет как источник информации и заставляет поисковые системы тратить силы и время на его искоренение, а не на повышение качества поисковых результатов.

Цель спамера - продвинуть свою страницу в первые строки выдачи. Наша цель - помешать ему в этом. Как оценивать

Существует множество видов спама. Некоторые из них представляют собой спам в "чистом виде", но другие могут находиться на достаточно ценных или просто релевантных запросу страницах. Поэтому оценка "спам" вынесена в дополнительную шкалу.

Перемещение страниц

Если фактический урл отличается оттого, который есть в задании (при этом отличается существенно: происходит смена хоста), то фактический урл в интерфейсе подсвечивается красным цветом.

ВАЖНО: смена адреса сама по себе не есть спам, однако есть некоторые спамные страницы, на которых происходит и редирект тоже. Поэтому в случае смены хоста урл подсвечивается, чтобы обратить внимание асессора на эту смену адреса. То есть красный урл - не всегда спам, а лишь повод посмотреть на сайт внимательнее, на предмет проверки, а не дорвей ли он.

Перемещения страниц бывают разной природы:

дорвей, т.е. сайт или вео-страница, перенаправляющая посетителей на другой сайт. Другой цели существования у дорвея нет: на нём нет никакого ценного контента.

Дорвей бывают двух видов: с автоматическим редиректом и без него.

Дорвей с редиректом примечателен тем, что адрес страницы в адресной строке самопроизвольно меняется. Перенаправление пользователя на другой сайт происходит автоматически. Иногда оно бывает мгновенным.

Дорвей всегда перебрасывает пользователя на какую-то страницу, которую он не искал (часто это порно), на разный мусор, который явно не несёт никакой пользы никому (кроме его создателей). Часто конечной страницей, на которую приводит дорвей, тоже является какой-либо спам. Вообще, дорвей достаточно уродливы для того, чтобы вопроса в их спамности не возникало, когда они попадают в оценку. Однако попадают они редко, поскольку все поисковые системы достаточно успешно с ними борются.

Дорвей без редиректа требует кликнуть по ссылке, чтобы перенаправление произошло.

Например: <http://my.opera.com/seksvkontakte/blog/>

Дорвей - это спам всегда, с редиректом он или без.

Серверный редирект.

Изменение адреса страницы или полное удаление контента с неё может и не являться злонамеренным действием. Крупные интернет-площадки часто манипулируют структурой сайта и даже доменными именами. Аналогично поступают сознательные и опытные вебмастера при переносе контента.

Один из частых примеров такого редиректа - создание промосайтов под актуальный продукт (новый тариф у сотового оператора, новый ноутбук и т.д.) Когда сайт выполнит свою функцию, на нём выставляют редирект.

Пример: <http://somebodyelsesphone.com/> - промосайт Nokia для нового телефона. Как сайт выглядел раньше, можно посмотреть здесь:

http://wklondon.typepad.com/welcome_to_optimism/2008/10/somebodyv-%20elses-phone.html

Теперь длинная цепочка редиректов приводит нас на <http://europe.nokia.com/home>

Встречаются внутренние редиректы там, где страницу переместили. Пример: mail.google.com или <http://www.joblist.ru/> (этот сайт купил холдинг Мейла и теперь он редиректит на сервис работа на сайте mail.ru).

Если случился какой-либо редирект, мы оцениваем эту страницу по релевантности. Спам за саму по себе смену адреса не ставим.

Ещё один вариант - зеркала сайта, которые функционируют неправильно:

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

<http://avp.ru/>

На <http://www.kasperskylab.ru/> главная страница функционирует как зеркало основного сайта, но все ссылки ведут на <http://www.kaspersky.ru/> - в этом отличие таких сайтов от настоящих зеркал, навигация которых полностью повторяет основной сайт и никуда не перекидывает.

Такие страницы получают оценку по релевантности и спам.

Информация для спам-ассессоров

Результаты поиска и выводы по тэгу

Результаты поиска по вебу всегда получают оценку "спам" на предварительной шкале и оценку по релевантности - на основной.

Результаты поиска по сайту получают оценку "спам" только тогда, когда на сайте есть альтернативный

Примеры: ищем аптеки на сайте <http://www.btk-online.ru> Результаты поиска http://www.btk-online.ru/search/?btk_all=1&r=&subr=&txt=%E0%EF%F2%E5%EA%E0&BookID=search&loc= - спам, т.к. те же результаты можно получить с помощью выдачи по тэгу <http://www.btk-online.ru/btk/?subr=169> Ищем дуэт арф на music.privet.ru <http://privet.ru/search/music/?qo=1&aquery=%D0%B4%D1%83%D1%8D%P1%82+%D0%B0%D1%80%P1%84&ss=music&x=0&v=0> - спам, т.к. возможна навигация с помощью тэгов: <http://music.privet.ru/taqs/?taq=%P0%B4%P1%83%P1%8P%P1%82+%P0%B0%P1%80%P1%84>

Ещё примеры:

<http://www.okov.net/index.php?autocom=mp3base&q=%CC%E8%EA%E0+%CD%FC%FE%F2%EE%EP> по запросу "Мика Ньютон - Выше чем любовь- Нежность" = спам + нерел, потому что ни слушать, ни скачать без регистрации нельзя, да и запрос к сайту не совпадает с запросом пользователя.

Если результаты поиска пустые, то оценки будут спам + нерел.

Внимание! Колдуншики. результатами поиска не являются.

Тонкости оценки=====

Q. Как мы относимся к выводам по тэгу?

А. Вывод по тэгу - это страница с сайта, которая должна объединять похожие по смыслу документы. Она является элементом навигации на сайте, но иногда может представлять из себя технический спам. Рассмотрим такие случаи подробнее:

http://www.riliz.ru/tags/mtk_6218?type=link - Спам. Получили документ по тэгу "mtk 6218" сообщение на нём гласит "Нет записей"

<http://habrahabr.ru/taq/mlcx/> - Спам. На сайте в данный момент времени нет ничего с таким тегом. Аналогично относимся к облаку тэгов.

<http://www.warez-box.ru/taq/> - Спам. Вся страница представляет из себя ссылки на существующие тэги, на странице результатов поиска такой документ не нужен.

Если вывод по тэгу не пустой, оцениваем его по релевантности. Примеры таких тэгов:

<http://www.autoblog.com/taq/geneva+motor+show/>

<http://news.kremlin.ru/by-keyword/93>

Q. Как правильно действовать при оценке, чтобы не перепутать спам и не спам?

А. Самое главное - разобраться в навигации на сайте. Как она устроена? Можно запросить рубрику или подкаталог, вывод по тэгу? Если да, то результаты поиска - спам.

Нет ни рубрик, ни подкаталогов, ни тэгов? Результаты поиска - не спам.

а дальше оцениваем по релевантности.

Q. А если на странице есть одновременно и контент, и результаты поиска? Например, в рефератах, как здесь: <http://revolution.allbest.ru/literature/00217649.html> - на странице искомый реферат, а внизу - список "подобные работы"

А. Это не спам. Просто дополнительный полезный сервис сайта.

Информация для спам-ассессоров

Спам-каталоги

В сети много каталогов сайтов. Известны каталоги <http://yaca.yandex.ru/>

Каталог сайтов - это систематизированная подборка ссылок на интернет-ресурсы с описаниями. Каталоги делятся на специализированные (на какую-то одну, конкретную тему) и общие, а ещё на региональные, национальные и глобальные. Каталогами удобно пользоваться, если вы ищете не ответ на конкретный вопрос, а сайты на какую-то тему. Составление хорошего каталога - ручная работа. Человек подбирает ресурсы, классифицирует их и пишет аннотации.

Каталог вовсе не обязательно будет спамом. Но есть каталоги, которые нужно пометить как спам: каталоги, составленные автоматически, не выполняющие своих функций или замусоренные до такой степени, что ими невозможно пользоваться. Такие каталоги создаются не для удобства пользователей, а для открутки рекламы.

Вот некоторые свойства спам-каталогов:

реклама является основным содержанием и занимает большую часть экрана или несколько экранов. Рекламный текст выделен ярким цветом, «основной» контент замаскирован или слабо заметен.

встречаются разделы, в которых описаний сайтов нет, а реклама уже есть.

для каждой ссылки или адреса с контактами создается отдельная страница (открытая для индексации поисковиками). Как правило, эта страница не несет никакой дополнительной информации, необходимой для выбора и оценки ресурса.

подозрительны и неактивные ссылки, хотя сами по себе признаком спама они не являются. дизайн у такого каталога шаблонный, т.к. они обычно создаются сетками. Нередко в контактах одного каталога указан e-mail на другом домене, соответствующем другому автокаталогу. Также подозрителен популярный почтовый ящик типа yandex, gmail, mail, rambler.

описания сайтов и даже адреса могут обрываться в неожиданных местах.

на странице могут присутствовать поисковые запросы.

заявленная тематика не соответствует фактической, как в отношении общей тематики каталога (заявлена тема «Все о мебели», но рубрики посвящены работе, знакомствам, и т.п.), так и внутри рубрик.

Пример: <http://www.milfirm.ru/fc1/firm15727.htm> Что здесь не так:

большую часть страницы занимает реклама

нет собственного поиска по сайту, а рубрикаторм невозможно пользоваться, потому что а)

нет классификации по регионам, б) заявленные разделы не имеют отношения к

действительности (издательский центр находится в рубрике "бизнес и финансы", в рубриках нет внутреннего деления).

К вредоносным каталогам относятся и так называемые линкопомойки - системы

автоматического обмена ссылками, которые подвешивают на вполне нормальных сайтах.

Например:

h	//noni-sok.ru/mvlink
t	s.php
t	
p	
h	//oaleks.ru/links.htm
t	
t	
p	
h	//paleks.ru/links/ind
t	ex.php
t	
p	
h	//paleks.ru/friends.ht
t	ml

t	
p	

и т.п. Т.е. если в выдаче по какому-то запросу мы видим страницу автоматического обмена ссылками, то помечаем её как спам.

Тонкости оценки

Q. "Много рекламы"- это сколько?

A. В целом надо смотреть, мешает реклама пользоваться содержанием каталога или нет. Если нужно долго прокручивать страницу, чтобы увидеть единственную строчку ценной информации - это плохо.

Q. Я увидел одно объявление в неподходящей рубрике. Я могу пометить каталог как спамный?

A. В любом каталоге могут быть ошибки. Стоит оценить каталог с точки зрения пользователя: насколько этот каталог внушает доверие, насколько он удобен. Если в рубриках постоянно попадаются неподходящие по тематике объявления - это плохо. Но

одно-единственное может быть случайностью.

20.01.2012 1:58

одну страницу определённо не для людей. <http://v-zopy.ru/> - Спам, переопт

Откровенный переопт, тут и поспорить не о чем: лингвистические рассуждения о происхождении слов на порно-сайте явно не для людей написаны.

<http://weicq.ru/> - Спам, переопт

Текст - сплошное переливание из пустого в порожнее, повторения, бесконечный перебор синонимов, попытка запихать побольше низкочастотных запросов.

<http://oval.ru/enc/49230.html> - Спам, переопт

Нет даже попыток вставить запрос в текст органично, в результате он совершенно нечитаемый.

Обещают рисунки, а их на деле нет. Спам однозначно.

http://www.molniam-realty.ru/residential_real_estate/house1/ Спам, переопт Большая портянка.

<http://songi.ru/perevod/H/HIM/Heaven%20Tonight/> - Не спам, не переопт

Спамить за строчку вверху: «Перевод песни HIM Heaven Tonight на русский. Слова песни Heaven Tonight исполнителя HIM. Текст песни HIM Heaven Tonight.» - не надо. Это стандартный блок, типа расширенного заголовка, и ещё он совершенно соответствует содержанию страницы. В целом, сайт низкого качества, но спамом не является.

<http://www.millionpodarkov.ru/sms/privatnye-sms.htm> - Не спам, не переопт

Текст подозрительный, но он не оптимизирован, просто рекламный текст с сайта.

<http://funlove.ru/odnoklassniki.html> - Не спам, не переопт

Текст совершенно нормальный, про сайт. Понятно, что упоминание "одноклассников" рефлексивно вызывает подозрение, однако, если отвлечься оттого факта, что это именно они, статья совершенно разумная и не переоптимизированная, спамить её не следует. Есть ещё момент, что там присутствуют разные написания названия сайта. За это тоже не надо спамить. Во-первых, там нет ни одного опечаточного названия (которые суть зло и спам). Во-вторых, те, что есть, разбросаны по тексту и употребляются как синонимы. В случае, когда все эти названия указаны в скобках через запятую, и таким образом ещё и вставлены в текст много-много раз, получается спам. В данном случае никакого злоупотребления нет. Понятно, что этот текст написан для переманивания аудитории одноклассников на другой сайт, однако сделан он качественно и спамить его не следует.

<http://3gp-film.ru/> - Не спам, не переопт

Предлагалось заспамить за блок «популярные фильмы», однако он не спам. Он одинаковый на всём сайте, это просто список популярных страниц с сайта. Слова про скачать присутствуют в заголовках и самих статей тоже, поэтому нельзя считать, что тут есть накачивание релевантности по слову «скачать».

<http://www.yeisknalog.ru/> - Не спам, не переопт

Перечисление ключевых слов под логотипом - некрасиво, но спамить их не стоит.

<http://moscowjob.net/vacsearch.php> - Не спам, не переопт

Подозрительный блок снизу, но недостаточно ужасный, чтобы посчитать его переоптом.

<http://svadba-gid.ru/catalogue/weddingsalon.html> - Не спам, не переопт Текст, конечно, оптимизированный, но не чересчур.

Технический спам

Тонкости оценки

Q. Часто встречаются так называемые развлекательные порталы. Как правило, они дают ссылки на разнородную информацию (музыку, фильмы, книги, картинки, приколы).

Подавляющее большинство из них не содержит оригинальной информации. Как нужно относиться к таким сайтам?

A. Если основной контент портала вторичен, а рекламы столько, что прочитать-посмотреть материалы нормально невозможно, то нужно пометить сайт как спам.

Q. На сайте есть только описание контента (фильмов, музыки и т.п.), а всё скачивание идёт с файлообменников. Помечаем как спам?

A. Нет Ссылки на файлообменники (letitbit, sendspace, rapidshare и другие) признаком спама не являются.

Q. На странице присутствуют странные ссылки на якобы файлообменники, которые не дают

скачать нужный файл (и ещё маскируются под известные обменники по виду). Что с ними делать?

А. Такие ссылки - это обман. Чаще всего ссылки на псевдообменники расположены в рекламном блоке с названием наподобие "скачать на высокой скорости". За наличие подобных ссылок странице надо поставить оценку "спам", вне зависимости от того, присутствуют ли на ней настоящие ссылки. На страницах бывают дополнительные блоки и другого типа, предлагающие не ссылку на псевдофайлообменник, а прямое скачивание файла с хоста, отличного от оцениваемого сайта. Примеры "плохих" и "хороших" ссылок можно посмотреть в учебнике.

Если при этом на странице есть и "хорошие" ссылки, их стоит оценить по релевантности. Например, страница может получить за хорошую рабочую ссылку - Рел+, а за наличие обманной - дополнительную оценку Спам.

Q. Когда нужно помечать спамом товарные агрегаторы (сайты-подборки цен на товары в разных магазинах)?

А. Признак спама для агрегаторов, помимо неудобной рубрикации и избытка рекламы - малое число источников. У спамных агрегаторов обычно 1-2 источника.

Q. Многие сайты с рефератами выглядят мусорными. Когда нужно помечать рефераты спамом?

А. Спамом помечаем те рефераты, которыми невозможно нормально пользоваться (в реферате отсутствуют картинки и таблицы, упомянутые в тексте, реферат нельзя скачать в нормальном виде (скачивается абракадабра какая-нибудь)). Создаются подобные сайты в первую очередь с целью заработка на рекламе.

Ещё признаки спам-рефератов:

Информация не отформатирована (формулы отсутствуют или не читаются, картинки и таблицы отсутствуют), невозможно скачать реферат в правильном формате; зачастую в одну статью собрано несколько текстов из разных источников (опять же без форматирования); Запредельное количество рекламы.

Помните, что на сайте могут быть и спамные, и неспамные страницы. Если вы видите хороший, удобный для пользования реферат, то помечайте его как неспам. Но если есть сомнения и колебания, то рекомендуется заглянуть в раздел "Математика" данного сайта и посмотреть на рефераты, содержащие формулы. Характерно, что на сайтах со спамными рефератами формулы не отображаются нормально. Подобные сайты

<http://bestreferat.com.ua/referat/detail-8269.html> или же вот такой <http://shara.orq.ua/referats/ru/show/1930/%CA%EE%EC%E5%F2%FB%20%E8%20%E8%F5%20%EF%F0%E8%F0%EE%E4%E0%20> (здесь скачивается всё в странной кодировке плюс есть классика жанра <http://shara.orq.ua/referats/ru/show/16721/90%20%D1%82%D1%80%D0%B8%D0%B3%D0%BE%D0%BD%D0%BE%D0%BC%D0%B5%D1%82%D1%80%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D1%85%20%P1%84%D0%BE%D1%80%D0%BC%D1%83%P0%BB%20/>) на основании данного подхода признаются спам-рефератами.

Q. Сайт открыл кучу новых окон с порно! Его пометить как спам?

А. Бывают развлекательные порталы, которые при любой активности пользователя на странице (и даже вовсе без его участия) открывают дополнительную страницу с порно или рекламой. Само по себе это не признак спама, * но только если дополнительные страницы можно без затруднений закрыть. Иногда они не позволяют закрыть себя естественным путём, требуют послать смс-ку и вынуждают пользователя закрывать браузер через диспетчер задач. Это вредные, раздражающие пользователя страницы, и их, ■ конечно, надо пометить как спам.

Скрытый текст

Это примитивный способ обмана поисковой машины. Он всегда помечается как спам (плюс оценка по релевантности). Такой текст «виден» для поискового робота, но имеет разметку, скрывающую его от отображения пользователю в браузере.

Наличие скрытого текста можно заподозрить, когда на странице много свободного места без видимых на то причин (вверху/внизу страницы). Для детектирования попробуйте «выделить всё» (Ctrl-A) или "отобразить страницу без стиля" (в Файрфоксе Вид => Стиль страницы).

Скрытый текст может стать при этом виден.

Приемы скрытия текста:

Цвет шрифта, близкий к цвету фона

Маленький, нечитаемый размер шрифта

Расположение за границами видимой зоны экрана

Партнерки

Существуют сайты, продающие товары и/или услуги другой компании. За привлечение покупателей они получают комиссию От компании - реального продавца. Подобные сайты имеют право на существование, но для нужд поиска необходимо уметь их определять.

Признаки партнёрок:

Описания товаров, их цены, на сайте-партнере повторяют описания на сайте-организаторе партнерской программы.

Если заказать что-либо (оформить заказ), происходит перенаправление на другой сайт, т.е. корзина или оформление заказа происходит на целевом сайте. Возможно отображение корзины и/или оформления заказа в новом окне, адрес которого можно увидеть в «Свойствах» страницы.

Картинки товаров грузятся с другого сайта (обнаруживается просмотром свойств картинки)

На сайте присутствует ссылка «Партнерская программа», которая ведет на целевой сайт организатора.

Отсутствуют оффлайн-координаты, особенно адрес.

В общих случаях мы оцениваем партнёрки по релевантности, а потом помечаем как спам.

Тонкости оценки

Q. По запросу требуется конкретный сайт, и этот сайт - партнерка. Нужно ли пометить его как спам?

А. В тех случаях, когда Витальной по запросу будет партнерка (например love.rambler.ru по запросу "Рамблер-знакомства"), спамом она не будет.

Q. Какие бывают партнерки?

А. Особое внимание следует уделять партнеркам Мамбы и Лавпланет, выглядящим как самостоятельные сайты. Их можно опознать по дизайну. <http://lovemoskva.ru/> - партнерка

<http://loveplanet.ru/> - исходный сайт

<http://mheart.ru/> - партнерка <http://mamba.ru/> - исходный сайт

Вот ещё партнёрки: <http://sota.allprikol.ru/> - партнёрка <http://playfon.ru/> - основной сайт много партнёрок у <http://www.diamondelectric.ru/> (обратите внимание на шапку "Вы можете сделать заказ", она на всех партнёнках выглядит одинаково), <http://www.sotmarket.ru>

Также всё ещё встречаются партнёрки Мерамеда, например,

<http://www.medplaneta.ru/consultation/> хотя основной сайт уже закрыт

А вот сайт <http://neprikasaemye.ru/shop/> - не партнерка, а подборка ссылок, потому что здесь

нет имитации магазина, нет цен и фальшивой корзины.

Q. А если сайты не продают услуги, а предоставляют возможность свободно скачать контент? Но этот контент на нескольких сайтах одинаковый.

A. Тут мы оценку "спам" не применяем.

Q. А если контент продаётся? Например, за деньги предлагают скачать фильм, музыку, порно.

A. Сам по себе платный контент - не спам. Но часто встречаются сайты, которые предлагают платить смс-ками на четырёхзначный номер. Если в выдаче несколько сайтов с платным контентом, нужно сличить эти номера. Одинаковые номера - признак партнёрки. Тогда мы такие сайты пометим как спам.

Сателлиты

Сателлиты - это сайты, создаваемые одним владельцем для продвижения главного сайта. Они нужны либо для наращивания числа ссылок на главный сайт с сайтов той же тематики, либо для того, чтобы занимать несколько позиций в топе выдачи по разным запросам, под которые они и оптимизируются. Например, если сайт предлагает набор услуг/товаров, то под каждую группу услуг/товаров делается отдельный сайт, который наполняется тем же контентом, что и соответствующая рубрика основного сайта. Наиболее распространенными примерами таких сайтов являются сетки сайтов по бронированию гостиниц (под каждую гостиницу создается свой сайт), туристические сайты (создаются сайты по странам, направлениям, базам отдыха), фирмы, предоставляющие услуги обучения (под каждый курс либо направление создается отдельный сайт), интернет-магазины с большим ассортиментом товаров.

Основные признаки:

шаблонный дизайн, обычно довольно бедный, без флеша и наворотов.

дублирование контента на главном сайте и на его сателлитах.

одинаковые контактные данные (часто на сателлитах приведен только телефон).

наличие информации о главном сайте (например в виде ссылки).

Примером подобного спама может служить сетка сайтов курортного магазина.

<http://www.kurortmag.ru/> - основной сайт, сателлиты - <http://www.abxaziya.ru/>

<http://www.krimea.ru/> <http://www.company-tour.ru/> <http://www.magadansochi.ru/>

Как мы поступаем с сателлитами? Если основной сайт хорошо известен (как в случае курортного магазина, например) и есть в выдаче наряду с другими сателлитами, то его спамом пометить не надо. Всё остальное оцениваем как спам.

Если вы не можете разобраться, который сайт главный, то пометайте спамом всё.

Спам-ассесоры разберутся.

Однако есть исключения. По ряду запросов сателлиты бывают полезны. Например, если человек задаёт

<http://cnews.ru/news/top/print.shtml72010/03/25/384233> * версия для печати.

<http://www.cnews.ru/news/top/index.shtml72010/03/25/384233>

12. Проблемы с автоопределением кодировки (aka крокозябры)

Если в документе вместо букв крокозябры, это, скорее всего, вызвано проблемами с автоопределением кодировки. За это документы никогда не получают оценку спам, это даже не является признаком спама. Что нужно сделать, чтобы увидеть читаемый текст? Пойти в меню Firefox (рыжая кнопка в левом верхнем углу) -> Веб-разработка -> Кодировка (либо Вид -> Кодировка, если вы используете старый дизайн). В этой вкладке нужно попробовать выбирать одну за другой все доступные варианты в нижней части (после серой черты, начиная с "Кириллица"), пока текст не обретёт адекватный вид. Когда правильная кодировка будет найдена, нужно оценивать сайт, как если бы он изначально был показан в таком виде. Если этого добиться так и не удалось, расследование считается законченным, других действий предпринимать не надо, верная оценка - 404

Грубые ошибки

Оценка - это мнение экспертов, в котором есть доля субъективности. Поэтому оценки разных

ассессоров одной и той же страницы могут различаться. Разногласия во мнениях не являются грубой ошибкой. Однако возможны в корне неверные оценки, которые мы считаем грубыми ошибками.

Расхождением во мнениях мы считаем:
разницу РелеванТный+/Полезный;
разницу РелеванТный-/НерелеванТный.
разницу НерелеванТный/Глупый
Остальные ошибки являются грубыми.

Тонкости оценки

При оценке страницы стоит в первую очередь обращать внимание на ее соответствие запросу, а не на ее удаленность от релевантной/витальной страницы. Наличие ссылки на витальную не делает текущую страницу более релевантной.

Если текущая страница дает ответ на запрос, ее стоит оценить как Рел+, ссылка на витальный сайт не сделает из страницы Полезную.

Если текущая страница не дает ответа на запрос (или дает неполный/косвенный ответ), но имеет отношение к запросу, ее можно оценить как Рел-. Сюда можно отнести, например, списки, из которых нужно выбрать конкретную страницу:

[Текст песни Максим Дождь], <http://www.sentido.ru/artist.php?artist=Maksim>

Если текущая страница не имеет никакого отношения к запросу, наличие на ней ссылки на релевантную или витальную не делает текущую страницу более релевантной. Пример:

[Ромка Киримов], <http://my.mail.ru/mail/zuaka/> - Не про то. Сама страница посвящена совершенно другому человеку, и тот факт, что в блоке "Друзья" виден некий Ромка Киримов, не делает ее лучше.

Если по запросу есть витальная страница, но в оценку попала другая страница с официального сайта, ее тоже стоит оценивать по релевантности запросу:

она может быть Полезной или Релевантной+, если в дополнение к витальной странице раскрывает тему запроса, может оказаться полезной для пользователя (например, сервисы Яндекса по запросу [Яндекс], сайты факультетов по запросу [ВУЗ] и т. п.)

она может быть Релевантной-, если имеет отношение к теме запроса, но не дает ответа или дает косвенный/слишком узкий/слишком широкий ответ (например, морда Яндекса по запросу

[яндекс.картинки] - слишком широкий ответ; страница

http://panasonic.net/avc/lumix/compact/zs7_tz10/ по запросу [panasonic] - слишком узкий).

~~Аналогично~~ она может быть нерелевантной или не про то, если не имеет никакого отношения к запросу или совершенно бесполезна для пользователя (например, по запросу [сайт Газпрома] новость от 2005 года на офсайте; страница <http://www.belinvestbank.by/index/1758/2609/> по запросу [Белинвестбанк] (Гродно)!!

Другими словами, для страниц витального сайта действуют те же правила, что и для любых других страниц: количество прокликов не должно влиять на оценку.

Q. Насколько важна формулировка запроса?

A. Чем конкретнее формулировка, тем выше требования к найденной странице. Например, по запросу «рецепты» страница <http://www.cooking.ru/> (входная страница отличного кулинарного сайта) будет полезной, несмотря на необходимость сделать несколько кликов для получения собственно рецептов.

Но по запросу «рецепт салата мимоза» оценка для этой страницы будет низкой (не выше релевантной-), так как пользователь конкретно сформулировал свою поисковую задачу, а мы не дали ему адекватного ответа.

По той же причине регион, указанный непосредственно в запросе, важнее региона, указанного в скобках (определенного машиной).

Q. Что повышает или понижает оценку сайта?

A. На оценку влияют:

достоверность информации. Бывают запросы, для которых важна авторитетность источника. Например, запрос «курс валют». Очевидно, что пользователя не удовлетворит высказывание на форуме на тему «что-то евро дорожает». Пользователь ищет конкретные данные. И удовлетворит его только информация на авторитетном ресурсе. В свою очередь, чтобы выяснить, как добраться до пляжа N (запрос «схема проезда к пляжу N»), достаточно развернутого ответа на форуме.

актуальность. Если запрос касается цен, расписаний, адресов или других данных, которые могут устаревать, то документ, где нет указания на дату или дата слишком давняя, менее ценен, чем заведомо свежий документ.

полнота информации, (см. комментарии к оценке Полезный)

Q. Может ли релевантный документ не содержать слов запроса?

A. Язык человека богат, одну и ту же потребность можно выразить разными словами.

Например, для пользователя запросы «гаражное оборудование», «оборудование для автосервиса», «оборудование для авторемонта» означают примерно одно и то же.

Потребность одна, а слова разные.

Поэтому вполне логично, что может существовать документ, хорошо соответствующий запросу, но не содержащий его слов.

Например, по запросу «поисковая система» находятся Полезные сайты yandex.ru, и google.com, но оба они не содержат в тексте документа фразу "поисковая система".

Вывод: релевантные документы часто содержат слова запроса, но это не обязательное условие, чтобы оценить документ как релевантный.

Q. Что делать, если в длинном документе начало не имеет никакого отношения к запросу?

A. Это может означать одно из двух:

- по запросу "А.С. Пушкин" сайт даёт биографию, полное собрание сочинений, критику, анализ.

по запросу «тесты по правилам этикета» сайт предлагает не только тесты, но и описание правил этикета, по возможности, для разных стран и эпох.

по запросу «телефоны Nokia» выдаётся страница с описаниями и фотографиями всех моделей телефонов фирмы Nokia, обзором цен на них и отзывами пользователей (например, http://market.yandex.ua/guru.xml?CMD=-RR=9.0.0.0-PF=1801946%2BEQ%2Bsel%2B1871151-VIS=560-%20CAT_ID=160043-EXC=1-PG=10&hid=91491)

Если ищут институт - Полезны сайты факультетов, подразделений и т.п.

Тонкости оценки

Оценка Полезный - очень важная, она имеет большой вес в наших метриках и мы хотим,

Аннотации ассессоры ставили ее вдумчиво. Мы просмотрели множество страниц, которым были поставлены оценки Полезный, и не все из них кажутся нам справедливыми.

Мы попытались выявить области, в которых чаще всего встречаются ошибки с оценкой Полезный. Ниже приводятся страницы, которые по нашему представлению заслуживают оценки Полезный, а также примеры страниц, которые с нашей точки зрения не стоило оценивать как Полезные,

Организации, услуги, товары и тп ([Ювелирные украшения]; [бары и рестораны])

- в документе есть релевантная часть, которая расположена ближе к окончанию; документ нерелевантен.

Как это проверить?

если релевантный отрывок документа содержит слова запроса (что бывает довольно часто), его можно быстро найти с помощью поиска. Для IE и Mozilla поиск вызывается нажатием комбинации клавиш Ctrl + F, а убирается окошко поиска по нажатию клавиши Esc. Работа с клавиатурой и мышкой быстрее и эффективней, чем работа только с мышкой. При поиске стоит искать не полное слово, а его основу. Например, при поиске «упражнения при плоскостопии» стоит начать поиск по слову