

Основные алгоритмы Google

В этой главе...

- Диагностика причин потери трафика сайта
- На что нацелены алгоритмы Google
- Как влияет на сайт его слабый контент
- Алгоритмы Panda и Penguin
- Алгоритм RankBrain

В поисковой машине Google используется много разных поисковых и ранжирующих алгоритмов. Но для веб-мастеров самыми важными являются алгоритмы-фильтры Panda и Penguin, которые могут оштрафовать сайт, в результате чего снизится органический трафик этого сайта. Эти алгоритмы, а также новый алгоритм RankBrain, и являются основной темой этой главы.

Потеря трафика и диагностика причин

Поисковые системы для оценки качества результатов проведенного поиска используют различные методы. Одни из таких методов — это алгоритмические штрафы, накладываемые на те или иные сайты различными алгоритмами поисковой системы. Другие — это ручные штрафы (пенальти), накладываемые на сайт сотрудниками поисковой машины, третьи — это, конечно же, алгоритмы ранжирования результатов поиска.

В некоторых случаях эти алгоритмы и наложенные штрафы могут оказать существенное влияние на естественный (органический) трафик вашего сайта. Все, оказывается, достаточно просто: ваш сайт был оштрафован (весь или какие-то его отдельные страницы) и в результате “выпал” из индекса поисковой машины. Следовательно, количество посетителей вашего сайта уменьшится ровно настолько, сколько их приходило от этой поисковой машины ранее.

На рис. 12.1 показано, что такие провалы трафика могут оказаться довольно существенными. Так, еще 18 января сайт имел посещаемость на уровне 581 уникальный посетитель в сутки, а уже 23 января количество его уникальных посетителей сократилось до 162. Однако печальная тенденция сохранилась — и 14.02 посетителей на сайте было уже всего 87.

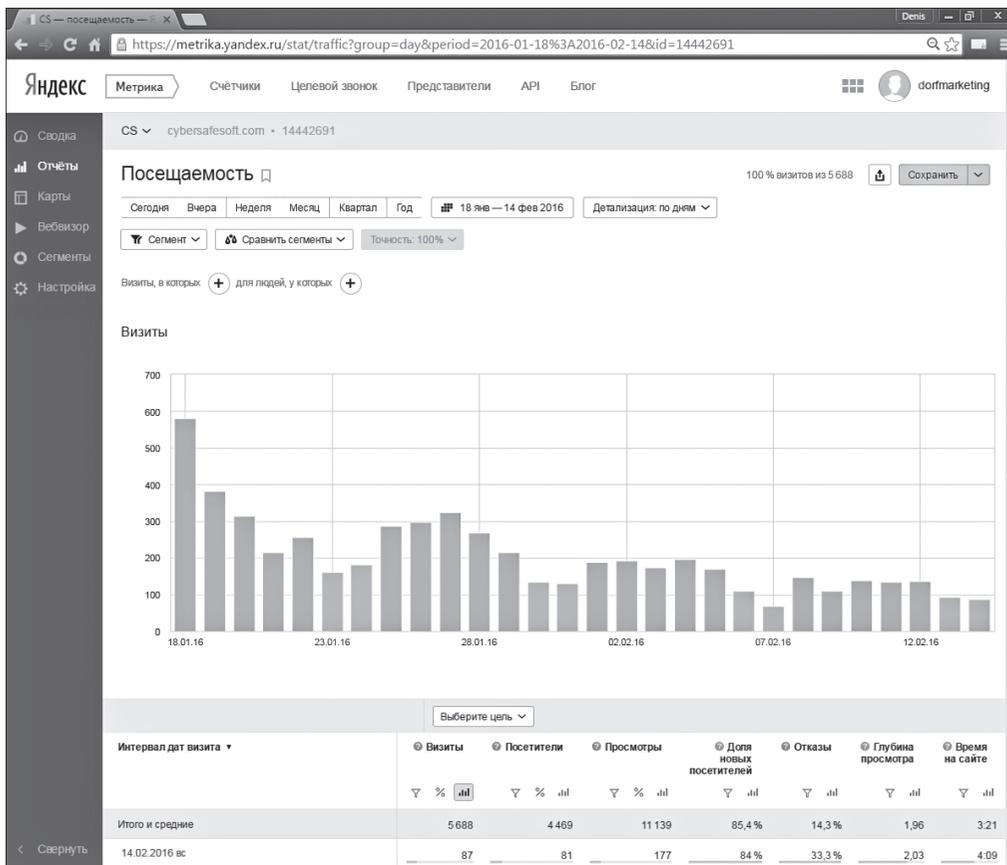


Рис. 12.1. Сокращение трафика на сайт

В чем причина такого провала? Безусловно, причины провала могут быть и естественными, например сезон отпусков (если, конечно, речь идет не о сайте туристической фирмы), новогодние праздники и т.д. Но в нашем примере новогодние праздники уже закончились, а количество посетителей вдруг стало резко сокращаться. Значит, причину этих неприятностей следует искать в чем-то другом.



Еще раз посмотрите на рис. 12.1. Если эта компания генерирует практически весь свой доход от органического поискового трафика, то ей был нанесен сокрушительный удар. Произошедшее означает сокращение дохода от сайта более чем в 6 раз, что может привести либо к сокращению штата компании, либо, в наихудшем случае, к полной ликвидации бизнеса. Необходимо понимать причины таких потерь трафика и знать, что нужно делать, чтобы восстановить прежние позиции сайта.

Цель данного раздела состоит в том, чтобы продемонстрировать, как можно выяснить реальные причины потери сайтом трафика. Прежде всего, необходимо проверить

его *аналитические* данные, чтобы уточнить, является ли наблюдаемое снижение трафика прямым следствием потери органического поискового трафика. Если у вас на сайте был установлен сервис **Google Analytics**, сначала убедитесь, что вы *проверяете* ваши источники трафика, а затем выделите трафик от поисковой машины Google.

Если в результате выяснится, что причина снижения трафика заключается в действиях Google, обязательно откройте поисковую консоль Google (**Search Console**) для вашего сайта и проверьте все сообщения, которые в нее поступили. Весьма вероятно, что среди них будут присутствовать сообщения о том, что вы были оштрафованы поисковой машиной Google.

На рис. 12.2 представлены сообщения в окне **Search Console**. Среди них *нет* сообщения о ручном штрафе сайта, однако есть множество других, свидетельствующих о том, что у сайта имеются серьезные проблемы. Одно из сообщений извещает о том, что обнаружены проблемы с мобильным поиском по вашему сайту. А это означает, что в результатах поиска с мобильных устройств рейтинг вашего сайта будет очень низким, это также может быть одной из причин снижения трафика.

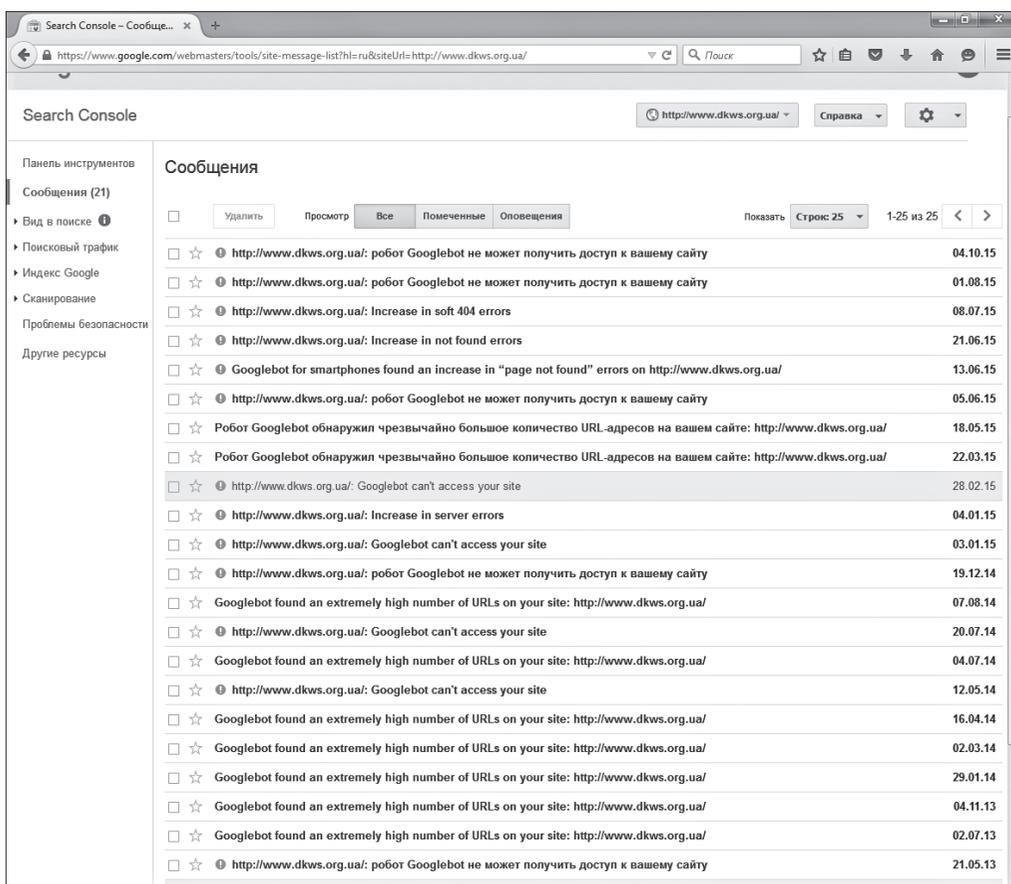


Рис. 12.2. Пример сообщений на панели **Search Console**

Если среди прочих обнаружится и сообщение о наложенном штрафе, то теперь вы точно будете знать, что эта проблема существует. Конечно, само по себе сообщение о наказании сайта — новость очень неприятная, но зато вы теперь точно знаете, что с сайтом есть конкретная проблема, а осознание проблемы — это первый шаг к ее решению.

Как ни странно, хуже, если подобного сообщения не обнаружится. Тогда вам придется копать глубже. Первое, что нужно сделать в такой ситуации, — определить точную дату начала падения трафика. Например, предположим, что с некоторым сайтом это произошло 24 апреля 2012 года. Далее посетите веб-страницу, содержащую даты обновлений в ПО Google: это сайт Moz Google Algorithm Update Page (<http://moz.com/google-algorithm-change>). Посмотрите, не совпадает ли дата начала падения трафика с одной из дат, указанных на этой страничке.

В нашем примере именно 24 апреля 2012 года вышел новый алгоритм Google, получивший название “Penguin”, и очень много прежде успешных сайтов “попали под удар” и потеряли свой трафик. Возвращаясь к предыдущему примеру этого раздела (см. рис. 12.1), на сайте Moz Google Algorithm Update Page можно обнаружить, что 8 января 2016 года вышло “безымянное” обновление ядра поисковой машины Google. Хотя разработчики Google и не связали это событие с обновлением алгоритма Penguin, оно все же могло затронуть тот сайт, о котором идет речь в примере, хотя и с небольшим опозданием.



Для исследования провалов трафика сайта вам также может очень пригодиться утилита **Penguin Tool**: <http://www.barracudadigital.co.uk/penguin-tool/> от компании Barracuda Digital. Эта утилита позволяет наложить ваши данные от Google Analytics на график обновлений алгоритма Google. Данный инструмент значительно упрощает процедуру сравнения дат.

Но самой сложной ситуация будет тогда, когда вы не получали сообщений о штрафе в Google Search Console, а дата начала снижения трафика сайта не совпадает с объявленными изменениями в алгоритмах Google. В такой ситуации бывает очень сложно понять, что же в действительности стало причиной наблюдаемого падения трафика.



Следует учитывать, что разработчики Google *ежедневно* (!) вносят небольшие изменения в действующие алгоритмы. С их точки зрения, это очень небольшие изменения, которые, безусловно, не подпадают под категорию основных или важных обновлений. Однако даже такие небольшие изменения могут стать причиной катастрофического падения трафика отдельных сайтов.

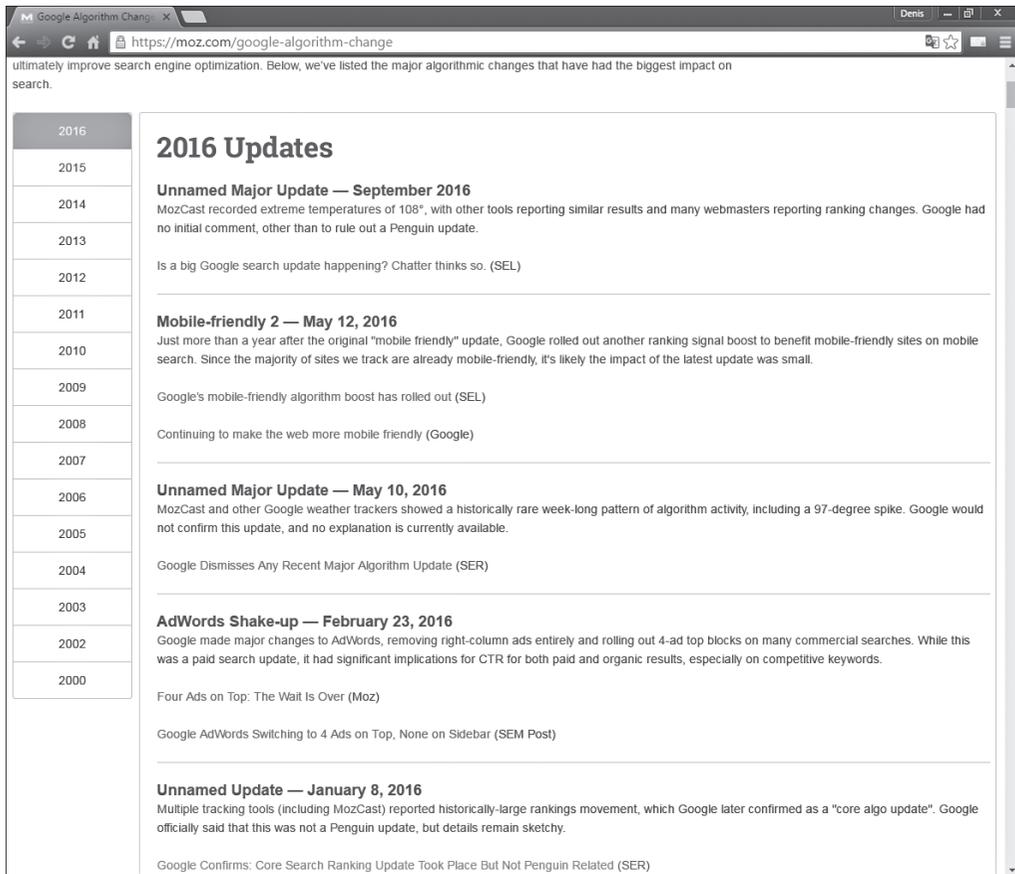


Рис. 12.3. Информация об основных обновлениях поисковых алгоритмов Google за 2016 год

Основные алгоритмы Google

В феврале 2011 года¹ были выпущены два основных алгоритма Google, нашумевшие на весь мир, — Панда (Panda) и Пингвин (Penguin). Алгоритмы Panda и Penguin подробно обсуждаются ниже в этой главе, а в табл. 12.1 приведены характеристики некоторых других алгоритмов, которые также существенно повлияли на органический поисковый трафик множества сайтов.

¹ Эта дата была выбрана мной специально, чтобы не вникать в старые алгоритмы, которые уже потеряли свою актуальность.

Таблица 12.1. Некоторые основные алгоритмы Google

Название	Дата запуска	Описание
Тор Heavy Update (Большое обновление)	19 января 2012 года	Затронул сайты, на которых было слишком много рекламы, в том числе даже сайты с AdSense. Чтобы избежать штрафа, не следовало размещать слишком много рекламы в верхней части веб-страниц. Даже небольшой рекламный баннер в верхней части страницы мог подвести сайт под действие алгоритма. Последний раз обновлялся 9 октября 2012 года
Пенальти DMCA (или Пират)	10 августа 2012 года	Применял штрафы к сайтам, которые получали слишком много запросов о нарушении авторских прав. ("DMCA" означает "Digital Millennium Copyright Act".) Чтобы избежать штрафа от этого алгоритма, необходимо было публиковать на сайте только собственный оригинальный контент. Если вы не пишете контент сами, а нанимаете людей, убедитесь, что для вашего сайта они пишут действительно оригинальные статьи
EMD (Exact Match Domain)	27 сентября 2012 года	Изменял механизм точного совпадения домена. Некоторые владельцы сайтов регистрировали домены, содержащие основные ключевые слова. Например, если вы занимаетесь безопасностью данных, то могли просто зарегистрировать домен <code>data-security.com</code> и оказаться первым сайтом в выдаче по запросу data security . Цель этого алгоритма — устранить такое преимущество
Payday Loan (для сайтов спамеров)	11 июня 2013 года	Алгоритм был нацелен на сайты, которые по его мнению являются спамерскими. Также под удар этого алгоритма попали порнографические сайты и онлайн-казино. Дата последнего обновления — 16 мая 2014 года
Humminbird (Колибри)	26 сентября 2013 года	Это очень крупное обновление; по сути, в Google переписали всю платформу поиска. Основная цель — сделать поиск более гибким и научить его адаптироваться к будущему. Начиная с этого момента поисковая машина Google стала лучше обрабатывать запросы на естественном языке
HTTPS/SSL ²	6 августа 2014 года	С этого момента в Google преимущество в поиске отдается тем сайтам, которые поддерживают протокол HTTPS. Это означает, что если есть два сайта, А и Б, у которых были примерно одинаковые позиции по определенному ключевому слову, то теперь преимущество отдается сайту, адрес которого начинается с <code>https://</code> . Если вы только создаете собственный сайт, сразу добавьте к нему поддержку протокола SSL

Название	Дата запуска	Описание
RankBrain ³	26 октября 2015 года	Корпорация Google объявила, что отныне ее поисковая машина будет самообучаться. По сути, RankBrain — это система искусственного интеллекта от Google. Она помогает поисковой машине Google отслеживать релевантные страницы для определенных запросов, не содержащихся на страницах сайта в прямом вхождении
AdWords Shake-up	23 февраля 2016 года	Внесены значительные изменения в сервис AdWords, имевшие серьезные последствия для CTR как платных, так и органических результатов, особенно для конкурентоспособных ключевых слов

Как уже было отмечено, каждый день Google может делать множество изменений в своих алгоритмах. В табл. 12.1 приведены только некоторые наиболее крупные изменения. За 2016 год было реализовано пять наиболее крупных обновлений Google. Ознакомиться с другими изменениями за 2016 год (и не только) можно по приведенной ранее ссылке⁴.

Алгоритм Panda

Хронология релизов алгоритма Panda

Алгоритм Панда (Google Panda) был выпущен 23 февраля 2011 года. Вообще говоря, 2011 и 2012 годы были очень урожайными в плане всяческих изменений, внесенных в поисковые алгоритмы Google. Так, в 2011 году было внесено более 20, а в 2012 году — более 30 только лишь *крупных* изменений. Для сравнения в 2015 году таких крупных обновлений было внесено всего 5, а за 2016 год — тоже 5 (правда, на момент написания этих строк год еще не закончился и окончательная картина была еще не ясна).

Алгоритм Panda — это не просто обновление. Это внесение радикальных изменений в само ядро поисковой машины Google. Алгоритм Panda был нацелен на снижение рейтинга некачественных сайтов, которые мало что могут сообщить нового их посетителям, поскольку просто копируют контент с других сайтов, а также на исключение из результатов поиска всех сайтов, которые просто бесполезны. Всего алгоритм затронул 12% (11,8% если соблюдать точность) поисковых запросов. В первую очередь, под удар попали сайты тонкого контента, фермы контента, сайты с большим количеством рекламы (особенно те, на которых рекламы было больше, чем контента) и другие сайты с некачественным контентом.

² <https://webmasters.googleblog.com/2014/08/https-as-ranking-signal.html>

³ Дополнительная информация: <http://searchengineland.com/faq-all-about-the-new-google-rankbrain-algorithm-234440>.

⁴ <https://moz.com/google-algorithm-change#2016>



Нужно отметить, что алгоритм Panda — это очень важное обновление. Есть очень немного обновлений, которые затрагивали бы более 10% всех поисковых запросов. Основная цель алгоритма Panda — найти сайты с некачественным контентом и оштрафовать их. Также этот алгоритм штрафует сайты, не имеющие существенной уникальной информации, например фермы контента.

Уже в апреле 2011 года⁵ вышла вторая версия алгоритма Panda — Panda 2.0. Если первая версия алгоритма затрагивала только англоязычные запросы, то вторая версия стала “международной” и теперь данный алгоритм был распространен на запросы, введенные на других языках мира. Также в алгоритм были интегрированы и другие сигналы ранжирования, например “черный список” сайтов, формируемый пользователями с помощью расширения **Blocklist** браузера Google Chrome. Расширение **Blocklist** предназначено для указания пользователями тех веб-страниц, которые они хотят полностью исключить из возвращаемых им результатов поиска (рис. 12.4).

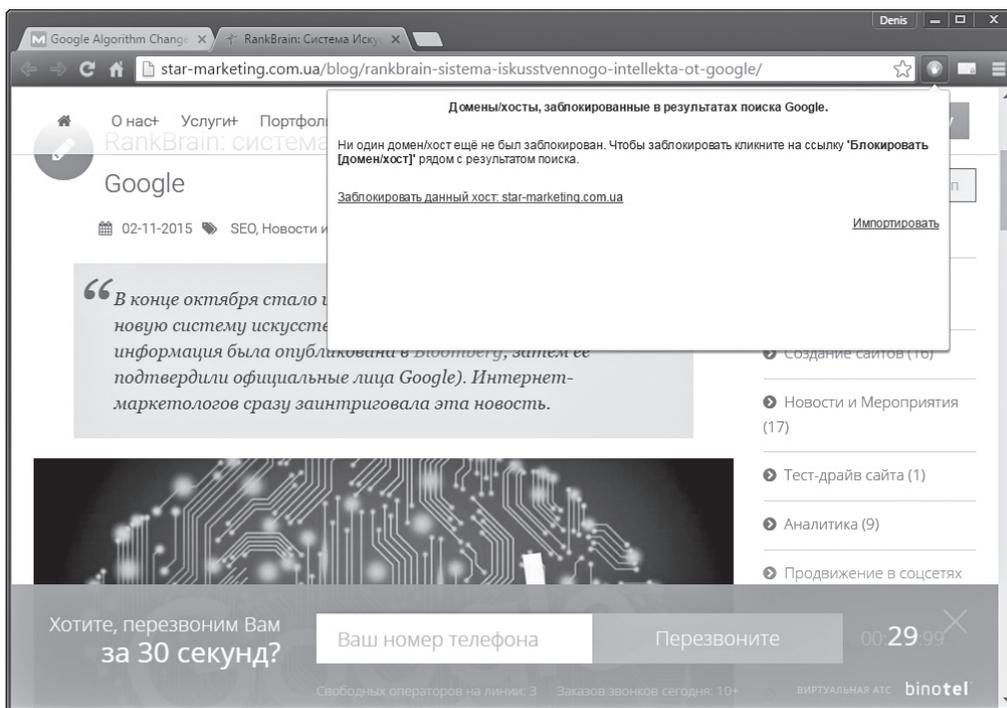


Рис. 12.4. Расширение **Blocklist** в браузере Google Chrome



Этим изменением корпорация Google впервые публично подтвердила, что будет использовать форму прямого ввода данных со стороны пользователей как фактор ранжирования сайтов в одном из своих алгоритмов.

⁵ Точнее, 11 апреля 2011 года.

Ранее уже было отмечено, что с апреля 2011 года алгоритм Panda стал поддерживать обработку запросов на других языках, помимо английского. Но полностью он стал действительно международным только с 12 августа 2011 года — это была версия Panda 2.4. С этого дня работа алгоритма Panda была развернута по всему миру, за исключением таких стран, как Япония, Китай и Корея. И вообще, с момента запуска первой версии в алгоритм Panda было внесено много изменений. В табл. 12.2 приводится полный список всех известных релизов алгоритма Panda.

Таблица 12.2. Сведения о выпуске новых релизов алгоритма Panda⁶

Релиз	Дата запуска	Описание
Panda 1.0	23 февраля 2011 года	Затронуто 11,8% результатов поиска, первая версия алгоритма
Panda 2.0	11 апреля 2011 года	Используется расширение Chrome Blocklist Extension, поддерживаются запросы на разных языках
Panda 2.1	9 мая 2011 года	Изначально предполагалось назвать релиз Panda 3.0, но потом сочли, что изменения не такие уж глобальные, поэтому релизу присвоили версию 2.1. Глобальных изменений не было, дополнительная информация доступна по адресу http://searchengineland.com/its-panda-update-2-not-3-google-says-76508
Panda 2.2	21 июня 2011 года	Некритические изменения в Panda
Panda 2.3	23 июля 2011 года	Еще одно изменение в самом Panda, изменились только данные самого алгоритма
Panda 2.4	12 августа 2011 года	Задействован на действительно всем международном уровне, затронул 6–9% запросов во всем мире, кроме Японии, Китая и Кореи
Panda 2.5	28 сентября 2011 года	Каких-либо сведений об этом алгоритме нет, и точно непонятно, что он затрагивал, но некоторые сайты сообщали о значительных потерях ⁶
Panda “Flux”	5 октября 2011 года	Называется «Flux». Затронул 2% запросов, незначительные изменения были также 11.10.11, 13.10.11 и 12.11.11
Panda 3.1	18 ноября 2011 года	Несмотря на то, что не было официального выпуска 3.0, эта версия почему-то получила номер 3.1. Основана на Panda Flux
Panda 3.2	18 января 2012 года	Хотя Google и подтвердила выпуск нового обновления, оно не затронуло большого числа запросов
Panda 3.3	27 февраля 2012 года	Не затронуло большого числа запросов
Panda 3.4	23 марта 2012 года	Затронуло 1,6% запросов
Panda 3.5	19 апреля 2012 года	Не затронуло большого числа запросов
Panda 3.6	27 апреля 2012 года	Не затронуло большого числа запросов
Panda 3.7	8 июня 2012 года	Затронуло менее 1% запросов
Panda 3.8	25 июня 2012 года	Не затронуло большого числа запросов

⁶ <http://searchengineland.com/google-panda-losers-today-show-winners-youtube-95257>

Релиз	Дата запуска	Описание
Panda 3.9	24 июля 2012 года	Затронуло около 1% запросов
Panda #18	20 августа 2012 года	С этого момента нумерация выпусков Panda изменилась
Panda #19	18 сентября 2012 года	Не затронуло большого числа запросов
Panda #20	27 сентября 2012 года	Затронуло 2,4% запросов
Panda #21	5 ноября 2012 года	Затронуло 1,1% запросов
Panda #22	21 ноября 2012 года	Не затронуло большого числа запросов
Panda #23	21 декабря 2012 года	Затронуто 1,3% запросов
Panda #24	22 января 2013 года	Затронуто 1,2% запросов
Panda #25	14 марта 2013 года	Дата определена Moz. Разработчики Google сделали алгоритм Panda частью основного алгоритма
Panda #26	11 июня 2013 года	С этого момента мелкие обновления больше не анонсировались, но они будут происходить каждый месяц с точностью плюс минус 10 дней
Panda #27	18 июля 2013 года	Смягчает влияние алгоритма Panda (по слухам)
Panda 4.0	20 мая 2014 года	Большое обновление, подтвержденное Google. Теперь Panda влияет на большинство сайтов, но не на новые сайты
Panda 4.1	23 сентября 2014 года	Значительное обновление, включающее изменения в алгоритмическом компоненте. Затронуло 3–5% запросов
Panda 4.2	17 июля 2015 года	Точно неизвестно, на что влияет это обновление, но абсолютно ясно, что оно не является значительным

На что влияет алгоритм Panda

За время, прошедшее с момента первого выпуска, корпорация Google так и не предоставила полную информацию относительно того, как алгоритм Panda определяет качество сайта. Однако сейчас точно известно, что Panda не любит следующее.

- ✓ **Тонкий контент.** К категории “тонкий контент” относят страницы с небольшим количеством контента. Такими страницами могут оказаться страницы профилей пользователей на форумах, страницы некоторых интернет-магазинов, на которых продается, например, 1000 продуктов, а описаний этих продуктов нет или оно совсем маленькое. (То есть получается, что на сайте есть 1000 страниц, по одной на каждый из предлагаемых продуктов, и на каждой из этих страниц присутствует одно-два предложения.) Такие сайты называются сайтами тонкого контента, и они не очень нравятся алгоритму Panda. Так что же делать в таких ситуациях? Если у вас форум, то проще всего закрыть от индексации все профили пользователей. Если же у вас интернет-магазин, добавьте на страницы продуктов их фотографии, приведите более подробные описания. Очень желательно также добавить

возможность комментирования каждого продукта и предложить их видеообзоры. Если такой подход вас не устраивает, просто удалите все страницы с описанием отдельных товаров и создайте единственную страницу, на которой будут перечислены сразу все товары. (Полезно будет также добавить на нее инструменты фильтрации, например, по категориям, чтобы облегчить поиск товара пользователям.) Добавьте кнопки покупки прямо на эту страницу — напротив каждого товара. Безусловно, страниц у сайта станет меньше, но зато они будут качественными.

- ✓ **Неоригинальный контент.** К такому контенту Google относит как полностью скопированный контент, так и скопированный контент с небольшими изменениями. Поисковая машина довольно легко обнаруживает такие страницы. Если на вашем сайте много страниц с неоригинальным контентом, он может быть оштрафован этим алгоритмом. И вообще, поисковой машине Google очень не нравится неоригинальный или некачественный контент. Качественный оригинальный контент — вот залог “благополучия” вашего сайта и гарантия высоких позиций на страницах с результатами поиска.
- ✓ **Недифференцированный контент.** Представим, что вы действительно пишете оригинальный контент для своего сайта. Но и этого может быть недостаточно. Если вы на своем сайте затрагиваете темы, которые уже обсуждались на других сайтах многие тысячи раз, особой пользы от вашего сайта не будет. Вот пример: посмотрите на рис. 12.5. На нем показаны результаты поиска в Google по запросу **торт наполеон**. Обратите внимание: результат поиска предоставляет собой 616 000 страниц! Какой толк от новой статьи, даже качественно написанной, если это будет 616 001 рецепт торта “Наполеон”? Весьма вероятно, что алгоритм Panda сочтет ваш сайт практически бесполезным, и правильно сделает.
- ✓ **Плохое качество контента.** Если в текстах страниц вашего сайта много орфографических и грамматических ошибок, то Google будет расценивать ваш сайт как содержащий контент плохого качества. Следовательно, о высоких позициях в выдачах вам придется, скорее всего, забыть. Но есть одно исключение — если контент вашего сайта будет совершенно уникальным, например в нем описывается изобретенный вами действующий способ телепортации на Марс. Вот тогда ваш сайт будет первым (и единственным) в списке результатов, даже если он будет содержать множество орфографических ошибок. Хотя именно в таком контенте их, вероятнее всего, не будет — едва ли изобретатель действительно работающего телепорта окажется настолько безграмотным.

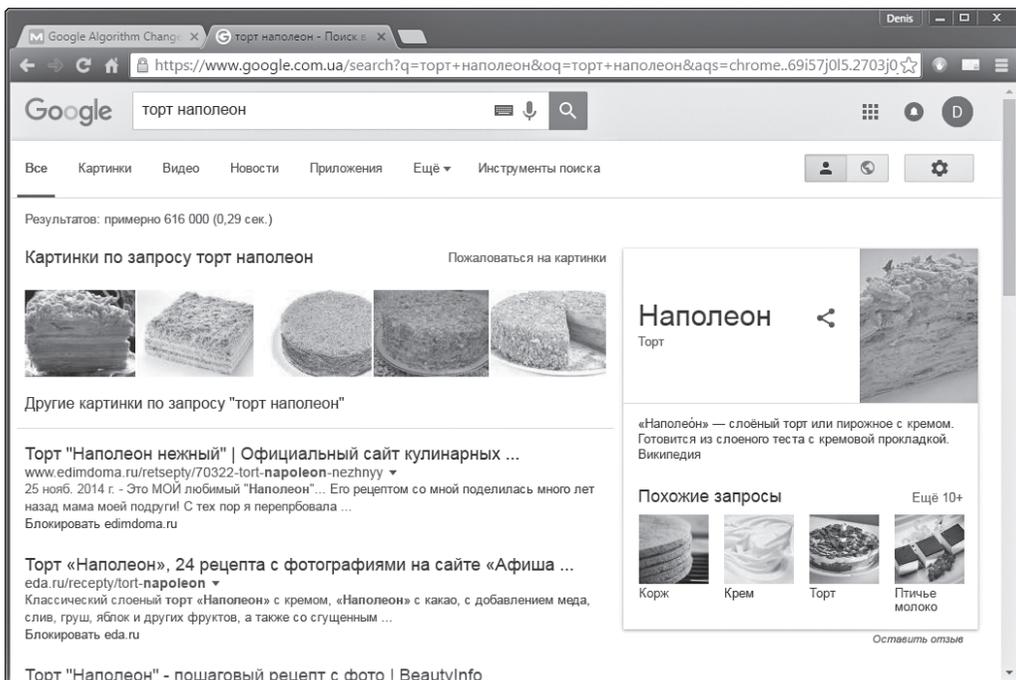


Рис. 12.5. Количество результатов по запросу **торт наполеон** огромно

✓ **Фермы контента.** Одной из основных целей поиска и отбраковки для алгоритма Panda являются так называемые “фермы контента”. Сайты ферм контента публикуют множество статей по одной теме со схожими заголовками. Это сайты компаний, производящих огромное количество текстового, фото- и видеоматериала любыми доступными способами — копированием, переписыванием, привлечением авторов-фрилансеров, работающих на заказ. Материалы необходимы для привлечения большого количества посетителей сайта, которые будут их просматривать, но цель создания ферм контента — вовсе не просвещение посетителей или утоление их интеллектуального голода. Фермы создаются для получения дохода от рекламы, показываемой их посетителям. Сам контент и его качество мало волнуют владельцев ферм — это лишь “мед” для привлечения простачков. Гораздо больше владельцев ферм волнуют позиции, которые их сайты занимают в результатах поиска, ведь это напрямую влияет на поток посетителей. В результате фермы контента по своей природе настолько сильно обесценивают результаты поиска, отодвигая действительно полезные сайты в нижнюю часть выдачи, что компания Google давно и очень активно занимается поиском и удалением их из своих результатов.

Подведем итог. Исходя из всего вышесказанного вот что вы можете сделать для своего сайта прямо сейчас, так сказать, в лечебно-профилактических целях.

- 1. Убедитесь, что на вашем сайте есть оригинальный, качественный контент.** Это особенно важно, если контент сайта писали наемные сотрудники.
- 2. Безо всякого сожаления удалите с сайта любой бесполезный контент.** В частности, это может быть пресловутый 616 001-й рецепт торта “Наполеон”.
- 3. Тщательно проверьте орфографию текстового контента.** Обнаруженные ошибки следует немедленно исправить.
- 4. При необходимости выполните полную реорганизацию сайта, например с целью устранить тонкий контент.**

Конечно, если ваш сайт — типичная ферма контента, ничего вы уже не сделаете. Скорее всего, придется искать другие способы заработка.

Обеспечение разнообразия в результатах поиска

Вернемся к нашему примеру с тортом “Наполеон”. Давайте изменим условия задачи и посмотрим, что будет, если ввести в поисковик укороченный запрос: **наполеон**. Но что именно требуется найти в данном случае? Можно предположить разное: биографию великого французского императора, рецепт торта, информацию о сериале “Наполеон” или сайты, на которых можно купить одноименной коньяк.

Получается, что в данном случае Google должен угадать, что *именно* вас интересует, и предоставить наиболее релевантные результаты. Посмотрите на рис. 12.6: чтобы решить эту проблему, в выдаче Google предоставлен максимальный спектр результатов: здесь есть информация и о Наполеоне Бонапарте, и рецепты торта “Наполеон”, и ссылка на страницу сериала “Наполеон” (я, например, даже не знал, что такой существует).

Обратите внимание, что результаты поиска в верхней части представляют сведения о французском императоре. Предполагается, что пользователь ищет информацию именно о нем. Это и понятно — торт, коньяк и сериал названы в его честь. Не было бы того самого Наполеона, не было бы и всех этих продуктов. Тем не менее в верхней части поисковой выдачи Google предоставляет информацию и о других страницах, которые могли бы заинтересовать пользователя (рецепты, фильмы и т.д.).

Как видите, Google вносит корректировки в содержимое страниц результатов поиска, используя концепцию QDD (Query Deserves Diversity — Запрос заслуживает разнообразия). В результате, даже если обычные сигналы ранжирования ставят след за предыдущей еще одну страницу об императоре, данная концепция требует заменить ее страницей, связанной с тортом, кино, коньяком или чем-то еще.

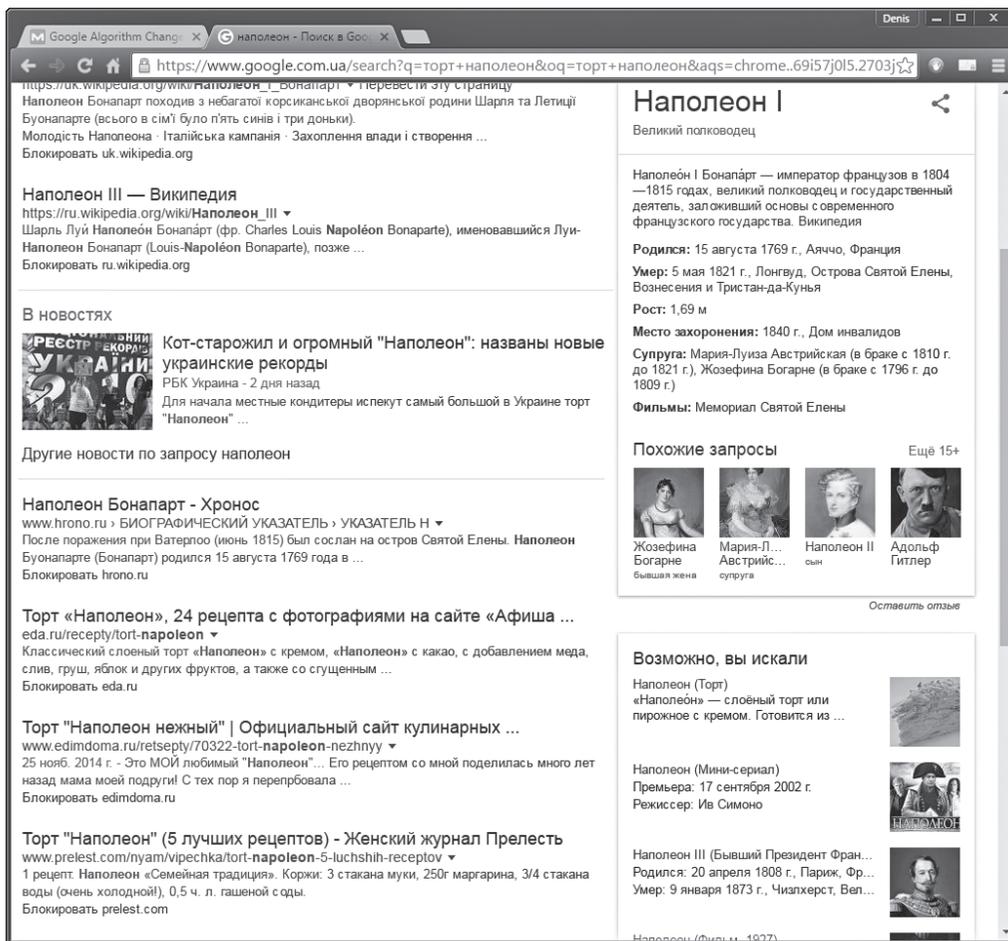


Рис. 12.6. Результаты поиска по запросу **наполеон**

Авторитет сайта влияет на результаты

Предположим, что вас интересуют выборы в Белоруссии, поэтому вы вводите в поисковик запрос **выборы Беларуси**. Давайте проанализируем полученные результаты — они показаны на рис. 12.7. На первом месте сайт Википедии, затем идет сайт *belta.by*, далее — *vybory2016.by*, а официальный сайт с календарем выборов *rec.gov.by* оказался лишь на пятом месте. Почему так получилось? Да потому, что авторитет сайта Википедии выше прочих, именно поэтому результаты с этого сайта появляются в выдачах раньше других.

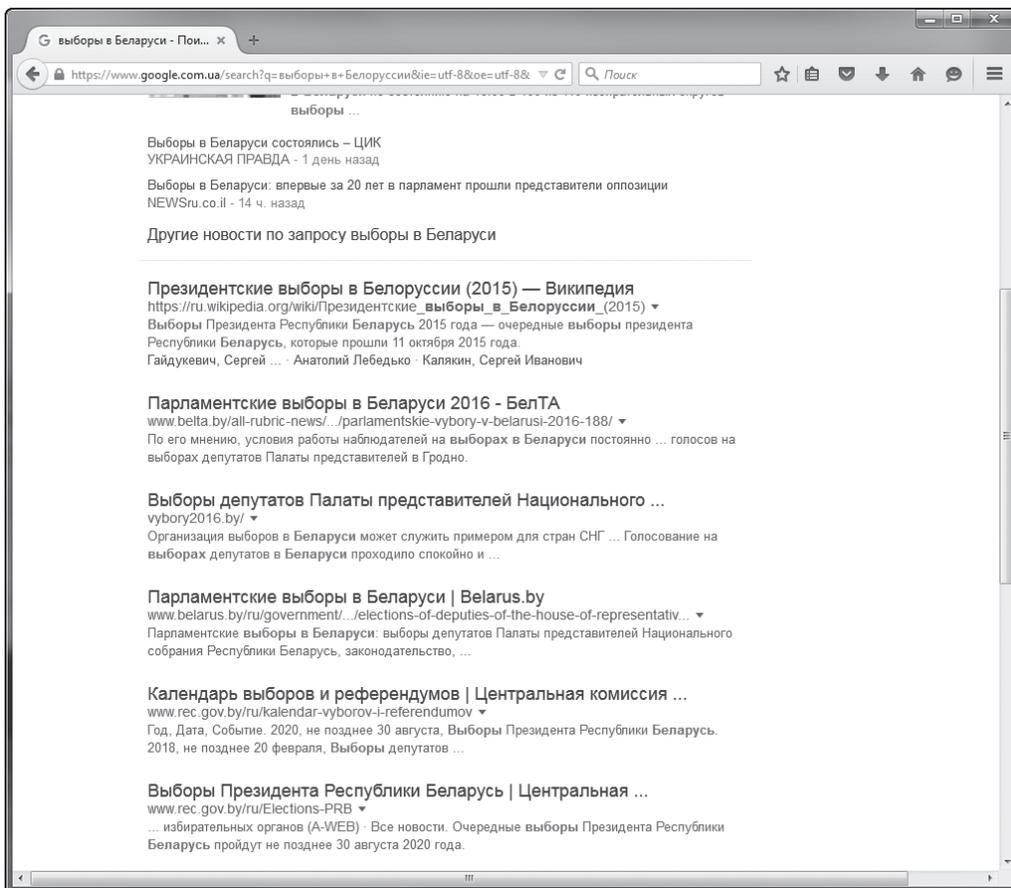


Рис. 12.7. Влияние авторитета сайтов на результаты поиска

Но почему все обстоит именно так, как именно принимается решение о высоком или низком авторитете? Прежде всего потому, что высокоавторитетным сайтам не свойственны характеристики или поведение, которое так не нравится Google (см. предыдущий раздел). На таких сайтах вы не найдете “копипаста”, тонкого контента, орфографических ошибок (если они и есть, то быстро исправляются) и других объектов внимания (или триггеров) алгоритма Panda.



Некоторые специалисты считают, что Google может применять более свободные критерии контроля для высокоавторитетных сайтов. Однако это предположение является всего лишь мнением и ничем конкретным не подтверждено.

Влияние слабого контента

Слабым считается любой контент, не несущий никакой ценности для посетителей сайта. Он не увеличивает конверсию, не содержит полезной информации, не оправдывает затрат на оптимизацию сайта, а также создает помехи для поисковых систем.



Если на сайте есть слабый контент, даже относительно немного, например 10–20%, то алгоритм Panda может понизить рейтинг вашего сайта. Важно регулярно исследовать сайт на наличие слабого контента. Обнаружив такой контент, немедленно удаляйте его безо всякого сожаления. Избавляйтесь от него, как от ненужного хлама. Запомните, пока на сайте есть такой контент, вы не сможете восстановиться после штрафа, наложенного алгоритмом Panda.

При исследовании своего сайта уделите больше внимания поиску слабого контента. Это могут быть как отдельные страницы, так и целые разделы, содержащие очевидно слабый материал. Если вы все же не желаете удалять такие страницы, то хотя бы воспользуйтесь следующими рекомендациями.

- ✓ **Улучшите контент.** Переработайте контент слабых страниц так, чтобы он содержал больше полезной для пользователей информации. Сделайте его более привлекательным для посетителей сайта. Надеюсь, об уникальности контента напоминать не стоит.
- ✓ **На страницы со слабым контентом добавьте метатег noindex.** Он закроет такой контент от индексации, и алгоритм Panda не найдет причин для наложения штрафов.
- ✓ **Если просто удалить такой контент невозможно, воспользуйтесь методом перенаправления.** Удалите слабый контент и создайте на этой странице 301-е перенаправление, уводящее пользователей на другую страницу вашего сайта, содержащую контент, похожий на удаленный или близко связанный с ним.
- ✓ **Если похожего контента нет и перенаправление невозможно, то все же удалите контент со страницы, а при попытке ее посещения возвращайте пользователю статус HTTP 410 — попытка посещения удаленной страницы.** Такой подход четко укажет поисковой системе, что страницы были удалены с вашего сайта.
- ✓ **Для удаления из индекса Google сведений о странице вашего сайта с плохим контентом воспользуйтесь следующим инструментом Google: <https://www.google.com/webmasters/tools/removals?pli=1>.**



Если вы — начинающий веб-мастер, то, возможно, еще не знаете, как можно организовать на сайте 301-е перенаправление и выдачу кода ответа 410. Так вот, оба эти действия можно реализовать с помощью файла .htaccess.

Создайте новый текстовый файл, присвойте ему имя `.htaccess` и поместите его в корневой каталог вашего веб-сайта. Чтобы организовать для некоторой страницы 301-е перенаправление, добавьте в него следующую строку:
`Redirect 301 /weak-content/ /related-content/`

В данном случае при посещении страницы `/weak-content/` произойдет автоматическое перенаправление на страницу `/related-content/`. Для отправки пользователю кода ответа 410 строка должна быть следующей:
`Redirect 410 /weak_content.html`

В этом случае при посещении пользователем страницы `/weak_content.html` ему вместо данных этой страницы будет отправлен код ответа 410 (рис. 12.8). В языке HTML это означает, что запрошенный пользователем материал был удален с сервера.

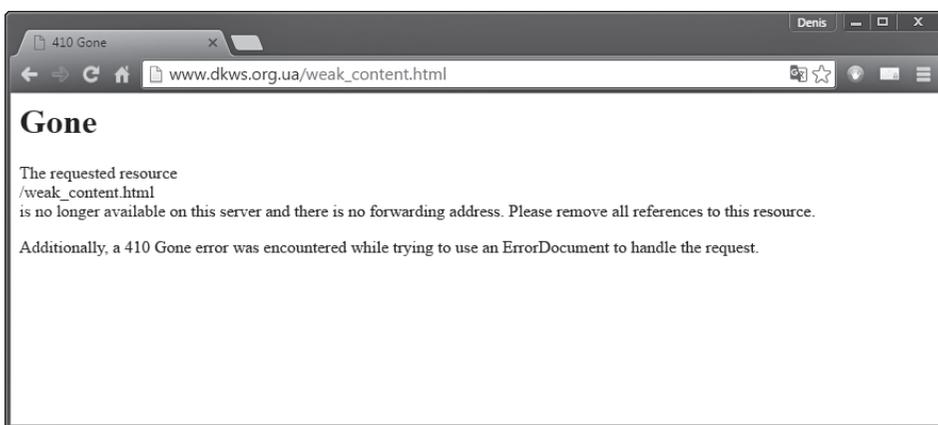


Рис. 12.8. Вот что увидит пользователь при обращении к удаленной странице



Запомните: если просто удалить страницу, не используя перенаправление 404, то сервер будет сам автоматически отправлять пользователю код ответа 404 и при этом на панели **Search Console** появится сообщение о “битых” ссылках на вашем сайте.

Как восстановиться после штрафа Panda

Первое, что нужно запомнить, — процесс восстановления после штрафа может занять много времени. Часто для восстановления может потребоваться полная переоценка существующей бизнес-модели сайта. Вам придется набраться смелости и критически посмотреть на свой сайт. Понятно, что часто это будет очень сложно сделать. То, что вы создавали годами, больше не работает, и для восстановления потребуется капитальный ремонт.



В некоторых случаях эффективна практика привлечения стороннего специалиста для оценки вашего сайта. Именно он сможет непредвзято посмотреть на состояние дел и увидеть все недостатки вашего сайта.

Любой капитальный ремонт начинается с разрушения того, что было сделано ранее. Поэтому, начав процесс восстановления после штрафа, на каком-то этапе вы обнаружите, что весь сайт практически разрушен. Но это не причина для отчаяния, просто вам необходимо существенно перенастроить сайт, и вы должны быть к этому готовы.

Взгляните на сайты конкурентов, которые не были оштрафованы. Посмотрите, как все устроено у них. Это очень хорошая практика. Правда, здесь есть одна опасность — вы можете обнаружить у них плохой контент (тонкий контент, слабый контент, контент плохого качества) и по этой причине отказаться от его удаления на своем сайте. Вот этого нельзя допустить. Возможно, алгоритм Panda просто еще не добрался до сайта вашего конкурента, и в этом случае падение рейтинга вашего сайта может сыграть вам на руку. Когда Panda доберется до конкурентов (а это обязательно случится, будьте в этом уверены), вы будете уже на полпути к восстановлению.

Как уже не раз было отмечено, залог здоровья сайта — качественный контент. Помимо рейтингов в поисковых машинах, в качестве бонуса вы получите еще и много сигналов социального взаимодействия — много репостов в социальных сетях. Хотя официально Google и не использует социальные сигналы как фактор ранжирования, все может измениться в будущем. К тому же существует сильная корреляция между большим количеством социальных сигналов и тем, что Google считает хорошим контентом, — в социальных сетях едва ли будут активно “репостить” плохой контент.



Существует целая наука о том, как создать хороший контент, цепляющий пользователя. Дополнительную информацию об этом можно получить по адресу <https://searchenginewatch.com/sew/how-to/2278413/after-google-panda-standout-content-is-essential>.

И все же как долго ждать восстановления? Хотя Google больше и не сообщает об обновлениях алгоритма Panda, в среднем они происходят раз в месяц. Следовательно, после внесения в сайт необходимых изменений вам останется только ждать. Рано или поздно Google заново просканирует ваш сайт и увидит проведенные изменения. Приготовьтесь, что может пройти несколько месяцев, прежде чем начнется заметное повышение трафика на ваш сайт. Конечно, улучшение будет только в том случае, если вы все сделаете правильно. Вполне вероятно, что с первого раза у вас может и не получиться.



Если показатели сайта долго не меняются, значит, вы что-то сделали не так. Возможно, вы удалили не весь плохой контент с вашего сайта или же Google просто не получает достаточного количества сигналов о том, что людям нравится ваш контент. В любом случае на сайте есть что-то, что не нравится алгоритму Panda.

Вы должны продолжать работать, пока не сделаете свой сайт интересным для пользователей, и как следствие ваши рейтинги вырастут, а Google снимет все санкции. Не воспринимайте процесс работы над сайтом как процесс борьбы с алгоритмом Panda, просто продолжайте совершенствовать свой сайт, пока это не принесет результаты. Конечно, все это потребует времени и определенного творческого подхода.

Не надейтесь, что у вас что-то получится без необходимых инвестиций и затрат рабочего времени. Но если вы вложили много денег (например, в написание контента) и внесли много улучшений, а результата пока нет, это может означать только одно: работа была выполнена не очень хорошо. И что же теперь? Да все то же, просто на восстановление вам понадобится больше времени и сил.

Всегда помните, что алгоритм Panda постоянно развивается. Даже если на сегодня вы еще не были оштрафованы этим алгоритмом, он вполне может оштрафовать вас завтра. В действительности Panda во всеуслышание передает владельцам сайтов основное послание Google: поисковик всегда будет “награждать” только сайты с полезным контентом. Поэтому работать над сайтом вам нужно только в одном направлении — улучшать его контент. Этим вы не только повысите шансы на восстановление после штрафных санкций алгоритма Panda, но и застрахуете себя от разрушительного воздействия будущих выпусков этого алгоритма.

Алгоритм Penguin

Алгоритм Penguin (Пингвин) появился 24 апреля 2012 года. С момента его выпуска он запомнился многим оптимизаторам сайтов как первый алгоритм Google, нацеленный исключительно на борьбу с плохими ссылками.

Релизы алгоритма

Как и в случае с алгоритмом Panda, с момента выпуска первой версии алгоритма Penguin неоднократно выходили его новые релизы. В табл. 12.3 приведена информация об основных выпусках алгоритма Penguin. Данные, как обычно, взяты с сайта <https://moz.com/google-algorithm-change>.

Таблица 12.3. Выпуски алгоритма Penguin

Релиз	Дата	Примечания
Penguin 1.0	24 апреля 2012 года	Затронул 3,1% запросов в США
Penguin 1.1	25 мая 2012 года	
Penguin 1.2	5 октября 2012 года	Затронул 0,3% запросов
Penguin 2.0	22 мая 2013 года	Что именно изменилось, точно не известно, но можно с уверенностью сказать, что алгоритм стал работать на уровне страницы, а не сайта
Penguin 2.1	4 октября 2013 г.	Общее влияние было незначительным, но некоторые веб-мастера сообщали, что он существенно повлиял на их поисковый трафик в худшую сторону
Penguin 3.0	17 октября 2014 года	Затронул 1% запросов
Penguin 4.0	23 сентября 2016 года	После двухлетнего перерыва вышла следующая версия алгоритма. На данный момент влияние алгоритма оказалось на удивление мало. Но, может быть, прошло недостаточно времени — с момента запуска алгоритма до момента написания этих строк прошло всего две недели

Редкие релизы — это характерная особенность этого алгоритма. Поэтому, если вы пострадали от него, вам придется ждать восстановления довольно долго. Вот почему очень важно разобраться, что именно не нравится этому алгоритму, чтобы он не наложил санкции на ваш сайт.

На что нацелен алгоритм Penguin

В главе 9 были рассмотрены ссылки и прочие вещи, которые не нравятся Google, но алгоритм Penguin во многом более привередлив. Ему точно не нравятся следующие типы ссылок.

- ✓ **Дешевые каталоги.** Именно поэтому оптимизация сайта путем его “прогона” по многим каталогам уже не эффективна. А раньше это был очень популярный способ оптимизации, а главное — дешевый. Многие SEO-организации предлагали автоматический прогон сайта по каталогам всего за несколько десятков долларов. Конкретная цена зависела от количества каталогов, в которых будет регистрироваться заявленный сайт. Также всего за несколько долларов можно было купить базу адресов таких каталогов и выполнить регистрацию в них самостоятельно. Логика была проста: за небольшие деньги вы быстро получаете несколько тысяч ссылок на ваш сайт. Однако сейчас такие ссылки уже ничего не значат. Более того, алгоритм Penguin может оштрафовать ваш сайт за такую практику. Регистрировать сайт нужно только в основных каталогах: Yahoo! Directory, DMOZ, Business.com, Best of the Web и некоторых отраслевых каталогах.
- ✓ **Каталоги статей.** Первый выпуск Penguin был нацелен как раз на каталоги статей, точнее на сайты, получающие ссылки из каталогов статей. Не всегда эти ссылки означают плохие намерения со стороны издателя, но Google заметил, что обычно это сайты плохого качества.
- ✓ **Спам в комментариях.** Следите за комментариями к статьям, записям, блогам, на форуме. Спам в комментариях — это проблема не только для сайта, на который ссылаются “спамные” комментарии, но и для сайта, на котором такие комментарии содержатся. Раньше существовал еще один популярный метод оптимизации — “прогон” сайта по форумам и блогам, когда на многочисленных блогах (где было разрешено добавление комментариев) и форумах размещалось объявление рекламного характера и ссылка на “раскручиваемый” сайт. Сейчас такое поведение считается абсолютно недопустимым, а все сайты, на которых практиковался такой способ оптимизации, были оштрафованы.
- ✓ **Неактуальные международные ссылки.** По мнению Google ссылки с сайтов других стран, в которых вы *не продаете* свой товар или услуги, следует рассматривать как проблемные. В любом случае от таких ссылок лучше отказаться. В главе 9 было показано, как это можно сделать.

- ✓ **Чрезмерное использование богатого якорного текста.** Эта ситуация также считается проблемой с точки зрения алгоритма Penguin. Этот алгоритм штрафует сайты, на которых имеется много экземпляров одного и того же якорного текста, расставленного в ссылках на любой из URL-адресов вашего сайта.
- ✓ **Страницы, состоящие только из ключевых слов.** Такой контент алгоритм Penguin считает очевидным поисковым спамом; следовательно, вам необходимо полностью исключить возможность его употребления.

Как восстановиться после штрафа

Помните, что новые релизы алгоритма Penguin выходят нечасто, поэтому вам нужно прямо сейчас, срочно избавиться от всех сомнительных ссылок на своем сайте. Позднее будет показано, как это сделать.

Самое страшное здесь то, что вы не можете подать запрос на повторную проверку⁷. Если алгоритм Penguin оштрафовал вас, то для восстановления нормальных условий работы вам придется ждать его повторного релиза. Напоминаю, что предыдущий был *два года* назад. В общем, штраф от алгоритма Penguin — это всерьез и надолго.



Как только вы избавитесь от всех плохих ссылок на свой сайт, вам останется только ждать. И только когда выйдет новый выпуск алгоритма Penguin, вы сможете узнать, все ли плохие ссылки были удалены.

Алгоритм RankBrain

Год назад стало известно, что при формировании поисковой выдачи Google использует систему искусственного интеллекта RankBrain (Ранжирующий интеллект), созданную на базе алгоритмов машинного обучения. Поэтому вполне можно сказать, что RankBrain — это не просто алгоритм, а система искусственного интеллекта.



Что такое *машинное обучение*? Под этим определением понимается процесс, когда машина обучается выполнению определенных задач, чтобы в дальнейшем следовать инструкциям и принимать решения чисто автоматически, без вмешательства человека. Алгоритм RankBrain является частью общего поискового алгоритма Google; по сути, это компьютерная программа, которая из миллиардов страниц, о которых ей известно, отбирает наиболее релевантные конкретным поступившим запросам.

Алгоритм RankBrain — это часть общего поискового алгоритма Hummingbird (Колибри). В течение многих лет у основного поискового алгоритма Google не было собственного названия, но в 2013 году в корпорации Google все же решили присвоить собственное имя и ему. Так появилось название “Колибри”.

⁷ В поисковой машине Google вы можете сделать запрос только в том случае, если штраф был сделан вручную, а не алгоритмически.

Так вот, алгоритм RankBrain является частью Колибри. Нужно отметить, что RankBrain обрабатывает не все запросы, а лишь часть из них. В состав Колибри входят и другие алгоритмы, такие как алгоритмы-фильтры Panda (Панда) и Penguin (Пингвин) (они были рассмотрены ранее), алгоритм локального поиска Pigeon (Голубь), алгоритм для борьбы с “переспамленными” запросами Payday (День полочки), алгоритм Mobile-friendly (Для мобильных устройств) и алгоритм Pirate (Пират), который призван бороться с нарушением авторских прав.



Задача алгоритма RankBrain заключается не в том, чтобы штрафовать сайты, а в том, чтобы выводить пользователям наиболее релевантные их запросу результаты. Алгоритм RankBrain является важным сигналом ранжирования. *Сигналы ранжирования* — это все то, что поисковая машина Google использует для определения, какую именно позицию нужно присвоить данной веб-странице в поисковой выдаче. Так вот, алгоритм RankBrain — новый и очень важный фактор в ранжировании страниц. Он используется преимущественно для интерпретации поисковых запросов. Главная цель этого сигнала — подобрать страницы, релевантные запросу, но такие, на которых точные вхождения ключевых слов этого запроса могут и не присутствовать.

Алгоритм RankBrain помогает Google лучше классифицировать веб-страницы. Данная цель достигается путем анализа контента, который содержат страницы. Возможно, RankBrain способен лучше, чем другие математические модели Google, понять, о чем идет речь на странице. Но пока это только предположения.