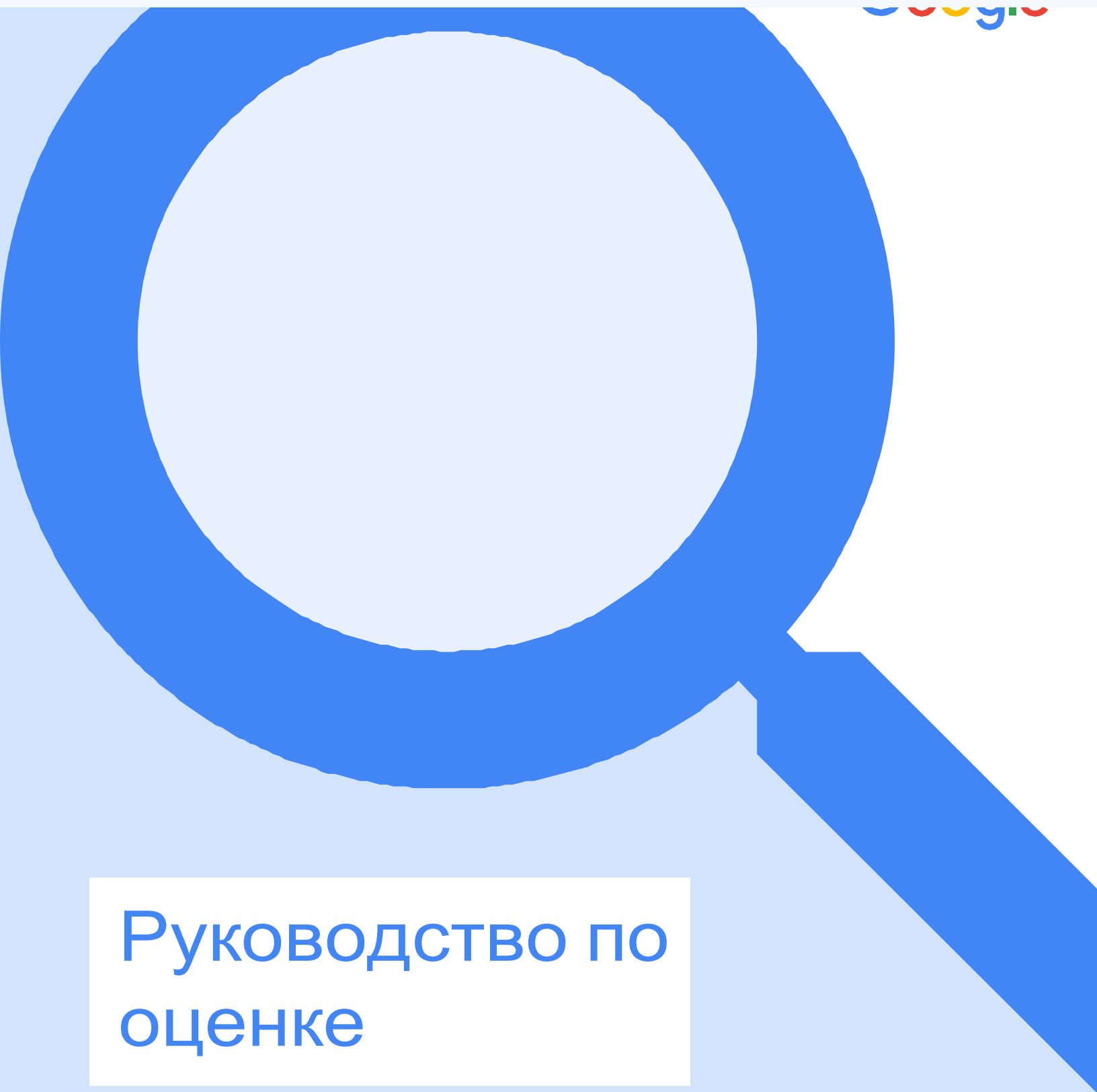




Руководство по оценке качества поиска:

Обзор



Введение

Раздел 01 Как работает поиск

Раздел 02 Улучшение поиска

Раздел 03 Процесс оценки качества

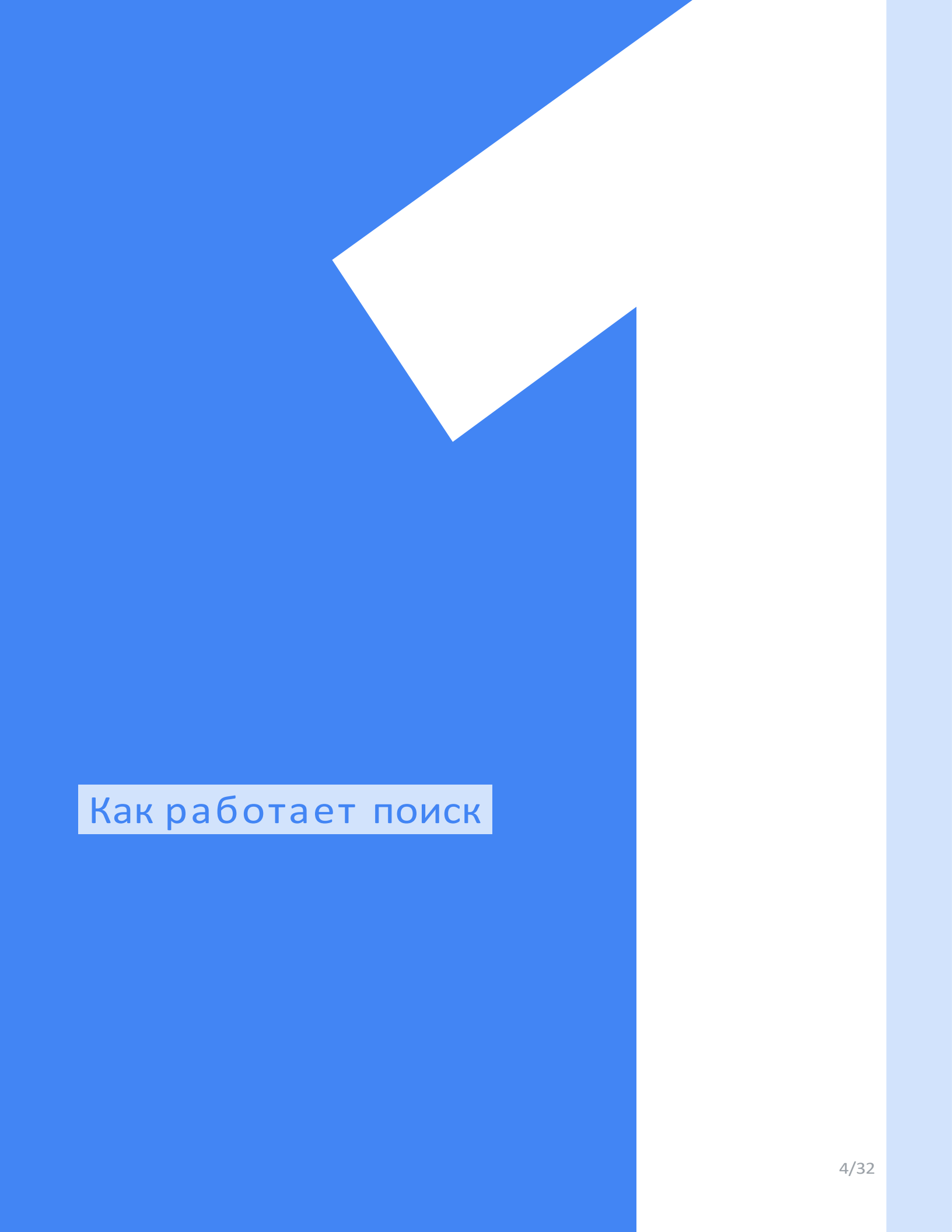
поиска Раздел 04 Резюме

Поисковые системы существуют для того, чтобы помогать людям находить полезную, актуальную и достоверную информацию. Для этого поисковые системы должны предоставлять разнообразный набор высококачественных результатов поиска, представленных в наиболее полезной форме.

В Google мы любим говорить, что Поиск - это не решенная проблема: мы постоянно вносим улучшения, чтобы Поиск работал лучше для наших пользователей. Мы подвергаем все предлагаемые усовершенствования продукта Search тщательному процессу оценки. Этот процесс включает в себя получение отзывов от "Оценщиков качества поиска", которые помогают нам определить, как люди могут воспринимать наши результаты.

В этом документе мы кратко объясняем, как работает Поиск и как мы улучшаем

Пройдитесь по этим оценкам качества поиска.



Как работает поиск

Для многих людей поиск Google - это место, куда они обращаются, когда хотят найти информацию по какому-либо вопросу, будь то получение более подробной информации по какому-либо вопросу или проверка фактов другом, цитирующим статистику футбольного матча в прошлые выходные.

Люди используют Поиск для миллиардов запросов каждый день, и одна из причин, по которой они продолжают обращаться к Google, заключается в том, что они знают, что здесь часто можно найти релевантную, надежную информацию, которой можно доверять. Чтобы помочь нам в этом, мы опираемся на три ключевых элемента, которые определяют наш подход к качеству информации:



Высококачественное автоматическое ранжирование: Мы разрабатываем наши автоматизированные системы ранжирования для определения информации, которую люди, скорее всего, сочтут полезной и надежной.



Полезные функции поиска: Мы разработали ряд функций поиска, которые не только помогают нашим пользователям разобраться во всей информации, которую они видят в Интернете, но и обеспечивают прямой доступ к информации из авторитетных источников, включая межправительственные организации или правительственные учреждения.



Политика в отношении контента: У нас есть политика в отношении того, что может появляться в Поиске, чтобы убедиться, что мы показываем высококачественный и полезный контент.

Менее чем за полсекунды наши системы сортируют сотни миллиардов веб-страниц, пытаясь найти наиболее релевантные и надежные результаты. Поэтому наши системы ранжирования разработаны с целью возвращения релевантных и надежных результатов. Для обеспечения релевантности результатов поиска у нас есть целый ряд систем, которые стремятся сопоставить слова и понятия в запросах пользователей с соответствующей информацией в нашем индексе. Эти системы варьируются от систем, которые понимают такие вещи, как опечатки или синонимы, до более продвинутых систем на основе искусственного интеллекта, которые могут понимать более сложные запросы на естественном языке.

Однако, когда речь идет о надежности - то есть о поиске высококачественной, заслуживающей доверия информации - даже с нашими передовыми возможностями понимания информации, поисковые системы, такие как Google, не понимают содержание так, как это делает человек.

Наши системы часто не могут определить по одним только словам или изображениям, является ли что-то преувеличенным, неправильным, некачественным или иным бесполезным.



Поисковые системы в значительной степени понимают качество контента с помощью так называемых "сигналов". Их можно рассматривать как подсказки о характеристиках страницы, которые соответствуют тому, что люди могут интерпретировать как высокое качество или надежность. Например, количество качественных страниц, ссылающихся на определенную страницу, является сигналом того, что эта страница может быть надежным источником информации по данной теме.

Как работает

Узнайте больше о том, как работает поиск, [здесь](#)



Улучшение поиска

Мы постоянно совершенствуем Поиск - только в 2021 году их было более 5 150. Эти изменения могут быть как крупными, так и небольшими, но все они направлены на то, чтобы Поиск работал лучше для наших пользователей и чтобы пользователи могли находить релевантную, высококачественную информацию, когда она им нужна.

Мы подвергаем все возможные изменения в Search тщательному процессу оценки, чтобы проанализировать показатели и решить, стоит ли внедрять предложенное изменение. Данные этих оценок и экспериментов проходят тщательную проверку опытными специалистами по анализу данных, менеджерами по продуктам и инженерами, которые затем определяют, будет ли изменение одобрено для запуска. Запуски также проходят дополнительные оценки, такие как проверка конфиденциальности и юридическая экспертиза.

1

У нас есть
представление о
том, как
улучшить Поиск

2

Мы развиваем
идею до
возможного
изменения

3

Мы оцениваем,
возвращает ли
эта идея более
полезные
результаты
поиска через...

- Исследования
- пользователей
- Живые
- эксперименты
- Тесты качества
поиска

4

В случае успеха
мы запускаем
изменение в
Search

Мы оцениваем поиск различными способами, включая тесты качества поиска, исследования пользователей и живые эксперименты:

Тесты качества поиска

Чтобы оценить, действительно ли наши системы предоставляют информацию, которую люди, осуществляющие поиск, считают релевантной и надежной, как и предполагалось, мы запрашиваем отзывы о предлагаемых улучшениях у специалистов по оценке качества поиска, которые используют наши Руководства для анализа изменений. Эти оценщики качества поиска представляют различные регионы мира, которые совместно выполняют миллионы пробных поисковых запросов и оценивают качество результатов в соответствии с сигналами, которые мы ранее установили.

В нашем [руководстве по оценке качества поиска](#) описано, на что обращают внимание наши оценщики. В разделе 3 мы объясним процесс оценки качества поиска более подробно.

Исследование пользователей

У нас есть команда по исследованию пользователей, в обязанности которой входит общение с людьми по всему миру, чтобы понять, как Поиск может быть более полезным и как люди в разных сообществах получают доступ к информации в Интернете.

Мы также приглашаем людей дать нам обратную связь по различным итерациям наших проектов.

Живые эксперименты

Мы проводим эксперименты с живым трафиком, чтобы увидеть, как реальные люди взаимодействуют с предложенным изменением, прежде чем запустить его для всех. Мы включаем функцию в отношении лишь небольшого процента людей, обычно от 0,1% или меньше. Затем мы сравниваем экспериментальную группу с контрольной группой, у которой эта функция не была включена.

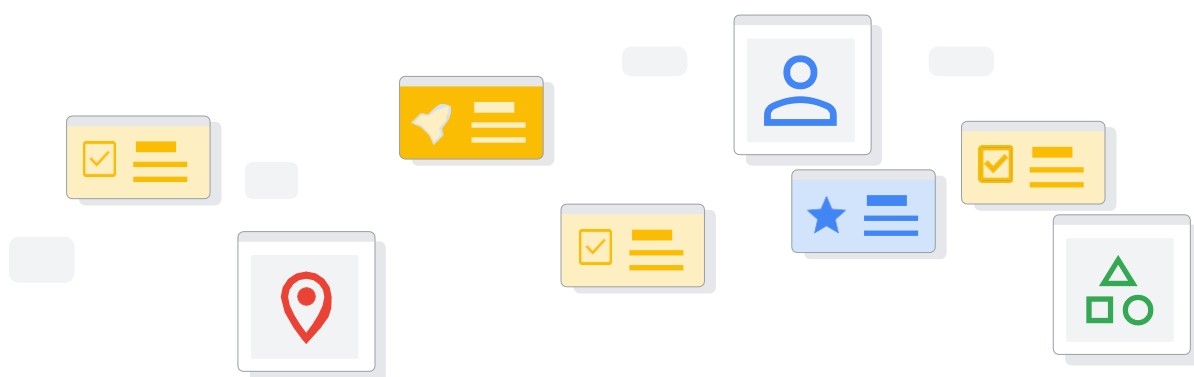
Мы смотрим на очень длинный список показателей, например, на что люди кликают, сколько запросов было сделано, были ли запросы брошены, сколько времени прошло, прежде чем люди кликнули на результат, и так далее. Мы используем эти результаты для оценки того, насколько положительным является использование новой функции, чтобы убедиться, что вносимые нами изменения повышают релевантность и полезность наших результатов для всех.



Процесс оценки качества поиска

Наш процесс оценки качества поиска оценивает качество результатов поиска на постоянной основе. Мы работаем с более чем 14 000 внешними оценщиками качества поиска, которые выставляют оценки на основе наших рекомендаций и представляют реальных пользователей и их вероятные информационные потребности, используя свои лучшие суждения для представления своей местности.

Чтобы оценить качество результатов поиска, оценщики работают с общим набором рекомендаций - и получают конкретные задания для выполнения.



Как работает задание по оценке

Сначала мы создаем выборку поисковых запросов (скажем, несколько сотен) для анализа определенного вида поиска или потенциального изменения рейтинга. Группа проверяющих получает этот набор поисковых запросов, и их просят выполнить определенные задания. Например, одним из видов заданий является эксперимент "бок о бок",

в котором эксперту показывают два разных набора результатов поиска: один с уже реализованным предлагаемым изменением, а другой - без него. Мы спрашиваем их, какие результаты они предпочитают и почему. Оценщики также дают ценные отзывы о других системах, например, о правописании. Чтобы оценить предлагаемые улучшения в системе правописания, мы спрашиваем у проверяющих, неправильно ли написан исходный запрос, и является ли исправление, созданное улучшенной системой правописания, точным для отдельных запросов.

Оценщики также просматривают каждую страницу, указанную в наборе результатов, и оценивают эту страницу по шкале **Рейтинга качества поиска в оптимизированном руководстве по оценке** частей:



Качество страницы

Насколько хорошо страница достигает своей цели



Оцените, является ли страница вредной
Оцените рейтинг



Удовлетворение потребностей

Насколько полезен результат для данного поиска



Оцените намерение пользователя
Оцените рейтинг

Важно отметить, что ни один отдельный рейтинг - ни один рейтинг - не влияет непосредственно на то, как данная страница или сайт ранжируется в Поиске. При наличии триллионов страниц, которые постоянно меняются, люди никак не могут оценить каждую страницу на регулярной основе. Простое использование рейтингов качества

Процесс оценки качества поиска

поиска для целей ранжирования не представляется возможным, поскольку люди никогда не смогут индивидуально оценить каждую страницу в открытом Интернете. Собственно, именно для этого и были разработаны поисковые системы.

Но помимо того, что это невыполнимая задача, использование только рейтингов качества поиска для определения ранжирования не даст достаточно сигналов для определения того, как должно работать поисковое ранжирование. Существует очень много аспектов качества и релевантности, которые имеют решающее значение, например, такие сигналы, как:

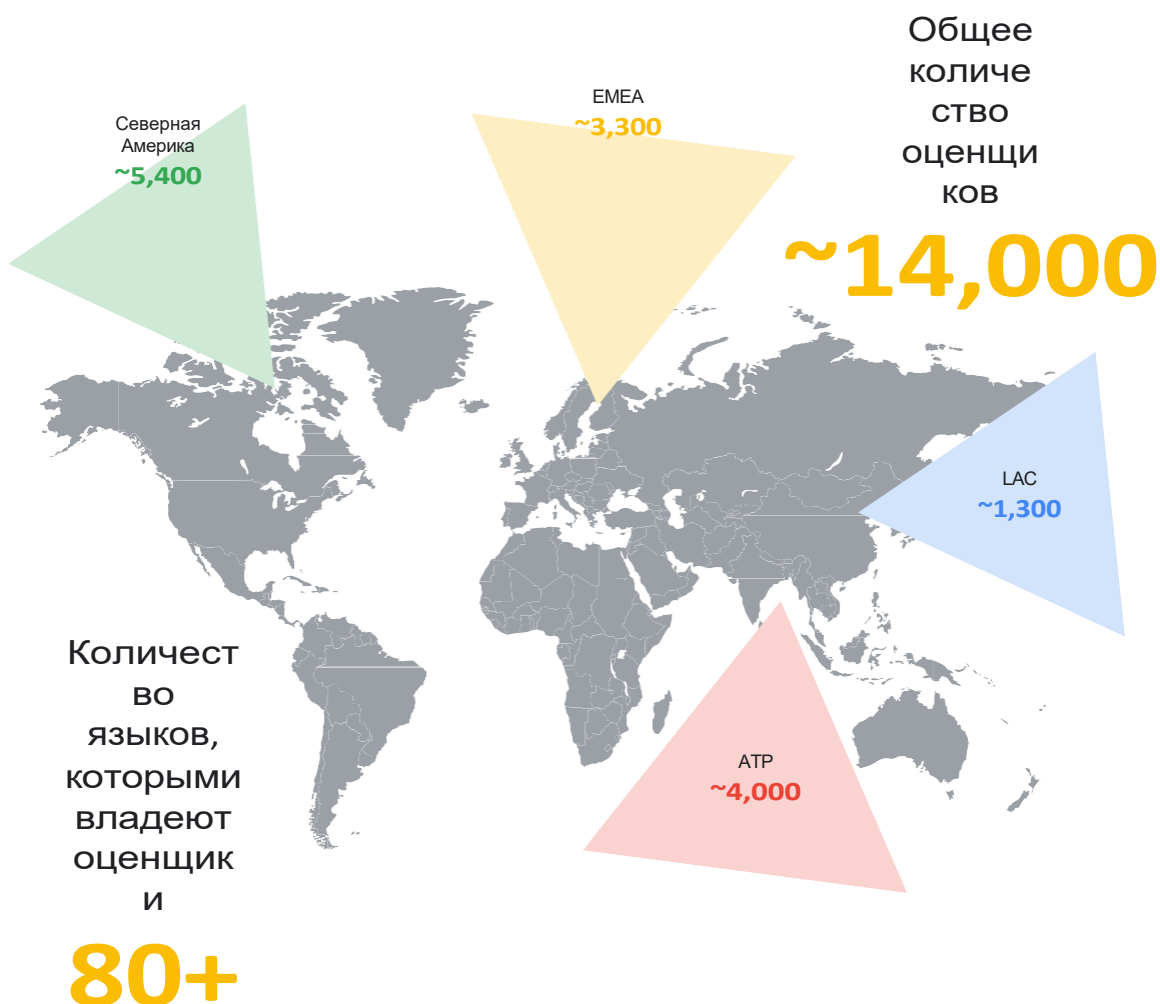
- сообщают нам, может ли что-то быть
- спамом; отмечают, что сайт может быть
- небезопасным или небезопасным; или указывают, что страница может быть устаревшей.

Таким образом, ни один источник информации - например, рейтинг качества поиска - никогда не сможет охватить все аспекты, важные для такой сложной задачи, как ранжирование.

Мы используем агрегированные рейтинги для оценки того, насколько хорошо работают наши системы для предоставления полезного контента по всему миру. Кроме того, рейтинги используются для улучшения наших систем, предоставляя им положительные и отрицательные примеры результатов поиска.

Рейтеры

Оценщики представляют пользователей поисковой системы в своей местности. Они должны быть хорошо знакомы с языком задания и местом его проведения, чтобы представлять опыт людей в своей местности, а также быть знакомы и чувствовать себя комфортно при использовании поисковой системы.



Оценщики нанимаются внешними поставщиками, и их число может меняться в зависимости от оперативных потребностей. В связи с важностью

масштабирования процесса оценки качества поиска по регионам, разнообразие местностей остается приоритетом.

Оценки не должны основываться на личных мнениях, предпочтениях, религиозных убеждениях или политических взглядах оценивающих. Вместо этого, оценщикам предписано всегда использовать свои лучшие суждения и представлять культурные стандарты своей местности. Оценщики также должны пройти тест на соответствие рекомендациям, чтобы гарантировать соблюдение наших собственных стандартов качества.

Рейтинг качества страницы

Цель рейтинга качества страницы (PQ) - оценить, насколько хорошо страница достигает своей цели.

Процесс оценки PQ состоит из трех этапов:



Качество страницы

Насколько хорошо страница достигает своей цели



Определите цель Оцените,
является ли страница
вредной Определите
рейтинг



Удовлетворение потребностей

Насколько полезен результат для данного поиска



Определите намерение
пользователя Определите
рейтинг

Шаг 1: Определение цели страницы



Качество страницы

Насколько хорошо страница достигает своей цели



Определите цель Оцените,
является ли страница
вредной Определите
рейтинг

Для того чтобы присвоить рейтинг, оценщики должны сначала понять цель страницы. Например, цель главной страницы новостного сайта - информировать пользователей о последних или важных событиях.

Это позволяет эксперту лучше понять, какие критерии важны при оценке качества данной конкретной страницы на этапе 3. Поскольку различные типы веб-сайтов и веб-страниц могут иметь совершенно разные цели, наши ожидания и стандарты для различных типов страниц также различны.

Вот несколько примеров, по которым легко понять назначение страницы:

Тип страницы	Назначение страницы
 Главная страница новостного сайта	Информировать пользователей о последних или важных событиях.
 Страница покупок	Продавать или предоставлять информацию о продукте.
 Видеостраница	Чтобы поделиться видео.
 Страница конвертера валют	Рассчитывать эквивалентные суммы в разных валютах.

Веб-сайты и страницы должны создаваться для того, чтобы помогать людям. Если это не так, то рейтинг "Низкий" может быть оправдан. Если страница создана для того, чтобы помогать людям, мы не будем считать, что какая-то конкретная цель или тип страницы выше по качеству, чем другая. Например, страницы энциклопедий не обязательно будут более качественными, чем страницы юмора.

Существуют веб-страницы самого высокого и самого низкого качества самых разных типов и назначения: страницы покупок, страницы новостей, страницы форумов, страницы видео, страницы с сообщениями об ошибках, PDF, изображения, страницы сплетен, страницы юмора, домашние страницы и все другие типы страниц. Тип страницы не определяет рейтинг PQ - для определения рейтинга необходимо понять назначение страницы.

Шаг 2: Оцените, соответствует ли цель страница вредна, или страница потенциально может причинить вред



Качество страницы

Насколько хорошо страница достигает своей цели



Определите цель

Оцените, является ли страница вредной

Определите рейтинг

Если сайт или страница имеют вредную цель или призваны обмануть людей относительно своей истинной цели, они немедленно получают оценку "Низшее качество" по шкале оценки PQ. Сюда входят сайты или страницы, которые наносят вред людям или обществу, не заслуживают доверия или являются спамом, как указано в руководстве.

В Интернете есть много материалов, которые некоторые сочтут спорными, односторонними, неприятными или отвратительными, но не сочтут их вредными.

указанные в руководстве. Оценщикам предписано следовать стандартам, изложенным в [разделе 7.0 руководства](#), где определено, что считается вредным.

Шаг 3: Определение рейтинга



Качество страницы

Насколько хорошо страница достигает своей цели

Определите цель Оцените, является ли страница вредной Определите рейтинг

Оценка PQ основывается на том, насколько хорошо страница соответствует своему назначению по шкале от самого низкого до самого высокого качества.



качества, но также имеет слабые характеристики Низкого качества. Сильные высокие характеристики качества затрудняют присвоение странице низкой оценки.

Высочайшее: Страница наивысшего качества служит полезной цели и достигает ее очень хорошо.



Основным критерием качества страницы является экспертность, авторитетность и достоверность (Е-А-Т) страницы. При оценке качества страницы оценщики учитывают:

- Опыт создателя.
- Авторитетность создателя, самого основного контента и сайта.
- Достоверность создателя, основного содержания и веб-сайта.

Оценщики определяют рейтинг качества страницы по:

- Анализ количества и качества основного содержания (ОС): Как правило, МК является высококачественным, если для его создания требуется значительное количество времени, усилий, опыта или таланта/навыков.
- Изучение доступной информации о сайте и его создателе: В зависимости от типа веб-сайта, для того чтобы пользователи могли доверять его содержанию или интерфейсу, требуется достаточное количество информации.
- Изучение репутации сайта и его создателя: Репутация сайта основана на опыте реальных пользователей, а также на мнении людей, которые являются экспертами в теме сайта.

Некоторые типы страниц требуют иных стандартов качества, чем другие. Например, некоторые типы страниц или темы могут потенциально повлиять на будущее счастье, здоровье, финансовую стабильность или безопасность человека. Мы называем такие страницы "Ваши деньги или ваша жизнь", или YMYL. Темы YMYL включают новости, финансы, здоровье и безопасность, а также правительство.

Рейтинги применяют очень высокие стандарты PQ для страниц YMYL, поскольку некачественные страницы YMYL могут потенциально негативно повлиять на счастье, здоровье, финансовую стабильность или безопасность человека. Примеры страниц YMYL можно найти в разделе 2.3 руководства.

Аналогичным образом, другие сайты или страницы, которые наносят вред людям или обществу, не заслуживают доверия или являются спамом, как указано в настоящем руководстве, должны получить самый низкий рейтинг. Примеры вредных страниц подробно описаны в разделе 7.0 руководства.

Пример

Веб-страница/Тип содержимого

Страница о сезонном гриппе
(YMYL)

Качественные характеристики

Очень высокий уровень E-A-T
для целей страницы

Очень положительная
репутация (веб-сайт)

Рейтинг PQ и объяснение



Это справочная страница по гриппу на заслуживающем доверия и авторитетном медицинском сайте. Этот сайт имеет репутацию одного из лучших веб-ресурсов для получения медицинской информации такого рода.

Оценка удовлетворения потребностей

Задачи рейтинга Needs Met сосредоточены на потребностях пользователей и на том, насколько полезен результат для людей, использующих Google Search.

Полезность результата поиска зависит от намерений пользователя, как они интерпретируются в запросе, и от того, насколько полно результат удовлетворяет этим намерениям. Поэтому оценка удовлетворения потребностей состоит из двух этапов:



Качество страницы

Насколько хорошо страница достигает своей цели



Определите цель Оцените, является ли страница вредной Определите рейтинг



Удовлетворение потребностей

Насколько полезен результат для данного поиска



Определите намерение пользователя Определите рейтинг

Шаг 1: Определение намерения пользователя на основе запроса



Удовлетворение потребностей

Насколько полезен результат для данного поиска



Определите намерение пользователя

Определите рейтинг

Запрос - это то, что пользователь набирает или произносит в нашей поисковой системе. Мы используем содержание запроса и, если уместно, местоположение пользователя, чтобы определить его намерения.

Например, если пользователь ищет "кофейни" и он находится в Лондоне, мы можем определить, что он хочет найти кофейни в Лондоне. Однако многие запросы также имеют более одного значения. Например, если пользователь ищет "ртуть", его намерением может быть узнать больше о планете ртуть или о химическом элементе.

Мы предполагаем, что пользователи ищут актуальную информацию по теме, а оценщиков инструктируют думать о текущем значении запроса во время оценки.

Шаг 2: Определение рейтинга



Удовлетворение потребностей

Насколько полезен результат для данного поиска



Определите намерение пользователя

Определите рейтинг

Оценка удовлетворения потребностей



О
ч
е
н
ь

с
о
о
т
в
е
т
с
т
в
у
е
т

ольшинст
ва
пользо
телей
мобильн
ых
устройст
в.
Некоторые
пользоате
ли могут
пожелать
увидеть
дополнител
ьные
результаты.

(
Н
М
)

О
ч
е
н
ь

п
о
л
е
з
н
о

д
л
я

м
н
о
г
и
х

и
л
и

Рейтинг NM основан на том, насколько результат поиска соответствует намерениям пользователя.

При определении оценки эксперт учитывает, в какой степени результат:

- "соответствует" запросу
- является
- всеобъемлющи
- м является
- актуальным
- исходит из авторитетного источника
- удовлетворяет пользователя, или хочет ли пользователь увидеть дополнительные результаты

Пример

Запрос и намерение пользователя

Запрос:
[Бэтмен]

Местоположение пользователя:

Анахайм,
Калифорния
Намерение пользователя:

Найдите информацию о вымышленном супергерое, который появляется в американских комиксах, фильмах и телевизионных шоу.

Результат



Рейтинг



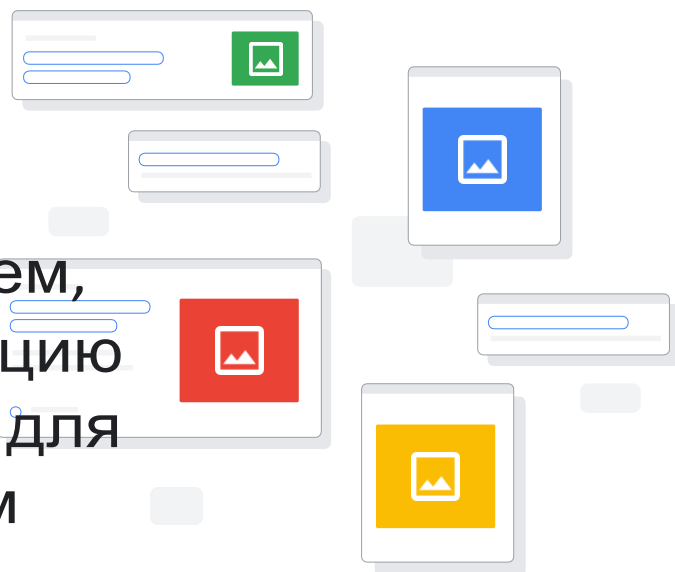
Пояснение

Крайне маловероятно, что этот запрос ищет информацию о городе в Турции под названием Батман, учитывая, что пользователь находится в США. Ни один или почти ни один пользователь не будет удовлетворены этим результатом.

Резюме

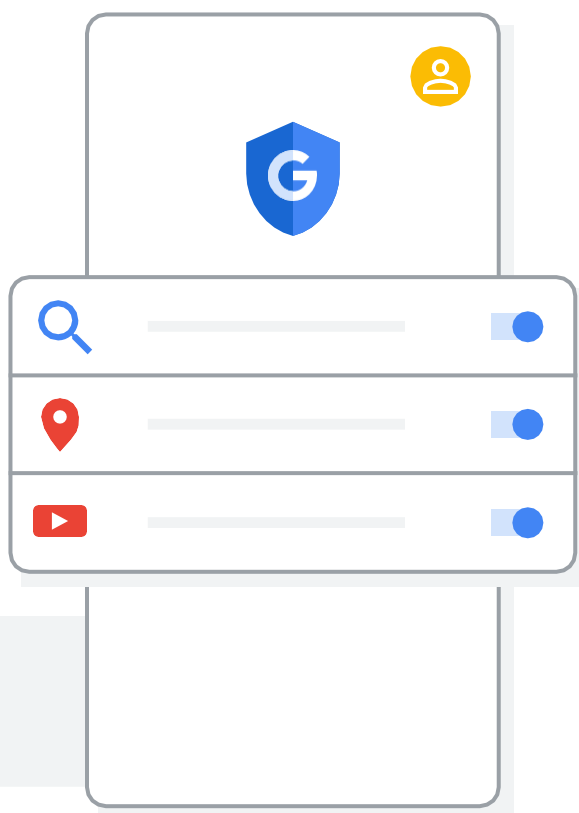
Доступ к информации лежит в основе миссии Google, и каждый день мы работаем над тем, чтобы сделать информацию из Интернета доступной для всех. Мы разрабатываем наши системы ранжирования таким образом, чтобы выявлять информацию, которую люди, скорее всего, сочтут полезной и надежной.

Поскольку Интернет и информационные потребности людей постоянно меняются, мы постоянно совершенствуем алгоритмы поиска, чтобы не отставать от них. Улучшения, которые мы вносим, проходят процесс оценки, чтобы люди во всем мире продолжали находить в Google полезную информацию, которую они ищут.



Ключевой частью нашего процесса оценки является получение отзывов от обычных пользователей о том, насколько эффективно работают наши системы ранжирования и предлагаемые усовершенствования. У нас работает более 10 000 специалистов по оценке качества поиска из разных регионов мира, которые, следуя инструкциям, которые каждый может прочитать в нашем руководстве по оценке качества поиска, оценивают результаты по примерным запросам и определяют, насколько страницы, попавшие в список, демонстрируют характеристики надежности и полезности.

Специалисты по оценке качества поиска должны выполнять задание по оценке качества страниц, которое позволяет оценить надежность результатов поиска. Для этого задания оценщики рассматривают качество веб-страницы, используя критерии, которые мы называем E-A-T: Экспертность, Авторитетность и Достоверность. У нас очень высокие стандарты оценки качества страниц для страниц "Ваши деньги или ваша жизнь" (YMYL), поскольку некачественные страницы YMYL могут негативно повлиять на счастье, здоровье, финансовую стабильность или безопасность человека. Страницы, которые считаются вредными для пользователей, поскольку, например, вводят пользователя в заблуждение или содержат опасный контент, получают самую низкую оценку качества страницы.



Специалист по оценке качества поиска также выполняет задание по оценке удовлетворения потребностей, в котором оценивается полезность результатов для пользователя путем анализа того, удовлетворяют ли они намерениям пользователя, как они интерпретируются на основе поискового запроса и местоположения пользователя. Оценщики должны учитывать степень полноты, актуальности или авторитетности результатов поиска. Например, результат, оцененный как "полностью удовлетворяющий", означает, что пользователь будет немедленно и полностью удовлетворен результатом и ему не потребуется просматривать другие результаты, чтобы удовлетворить свою потребность.



Оценки качества поиска сами по себе не влияют непосредственно на отображение конкретной веб-страницы, веб-сайта или результата в Поиске Google, а также не приводят к перемещению конкретных веб-страниц, веб-сайтов или результатов вверх или вниз на странице результатов поиска. Вместо этого рейтинги будут учитываться в совокупности и использоваться для оценки эффективности работы алгоритмов поисковых систем для широкого круга поисковых запросов.

Дополнительную информацию о рейтинге качества поиска можно найти на следующих ресурсах:



[Руководство по оценке качества](#)

[поиска Как работает поиск](#)

[Как Google предоставляет достоверную информацию в Поиске](#)

[Как информация, полученная от людей со всего мира, делает Google Поиск лучше](#)



[Как реальные люди делают поиск Google лучше](#)

[Могли бы вы ходить в обуви Google? Принятие сложных решений при работе с поисковыми объявлениями](#)

