

Новый поиск в Яндексе

*Стефанова Н.А., к.э.н., доцент, заместитель заведующего кафедрой
«Цифровая экономика»*

*Чупанов А.А., студент кафедры
«Цифровая экономика»*

*ФГБОУ ВО «Поволжский государственный университет телекоммуникаций
и информатики»*

e-mail: kilikopik@yandex.ru

Россия, Самара

Интернет уже давно стал неотъемлемой частью жизни практически каждого человека. В сети люди узнают много нового: читают книги, смотрят новости, ищут какую-либо информацию, общаются. Для удобства поиска информации, компания Яндекс разрабатывает новые поисковые алгоритмы. Никому не понравится если человек захочет найти в Интернете рецепты, а увидит абсолютно всё, что с ними не связано. Поэтому качество поисковых запросов и их выдачи очень важно для каждого пользователя сети Интернет.

Поисковые запросы являются одним из самых эффективных инструментов интернет-маркетинга по привлечению дополнительных пользователей, потенциальных клиентов, и как следствие получения прибыли. При этом на рынке поисковых систем обострилась конкурентная борьба между его лидерами Яндексом и Google. Последняя причем за последний год вышла по данным статистики на лидирующие позиции. Поэтому в стане Яндекса тщательным образом для возвращения доверия пользователей и завоевания их внимания ведутся работы по совершенствованию своих сервисов.

Так 22 августа 2017 года в Московском планетарии «Яндекс» представил новый поисковый алгоритм «Королёв». Крупнейшая поисковая система России запустила новую версию поиска, в основе которой сопоставление смысла запроса и веб-страницы, сообщает компания. Продукт презентовали руководитель поиска «Яндекса» Андрей Стыскин, руководитель службы релевантности и лингвистики Александр Сафронов и руководитель асессорской сети Ольга Мигорская.

«В прошлом году мы запустили новый алгоритм ранжирования “Палех”, в котором сделали первые шаги в сторону семантического поиска, а сегодня запускаем новый алгоритм ранжирования «Королёв». Почему мы выбрали такое название? Сергей Павлович Королев осуществил мечту человечества о полетах в космос. Для нас в «Яндексе» сегодняшний запуск является таким же важным технологическим прорывом к мечте о поиске, который понимает пользователей», – сказал на презентации поиска руководитель службы релевантности лингвистики «Яндекс» Александр Сафронов.

Яндекс уже давно начал идти в направлении машинного обучения поисковых алгоритмов, что даёт наиболее точное понятие поисковых запросов.

По данным LiveInternet, в июле 2017 года «Яндекс» занимал 51% рынка поиска в России, Google – 44%. Для бизнеса российской компании, большую

часть выручки которой приносит контекстная реклама, этот показатель критически важен. Но в течение нескольких лет «Яндекс» терял долю на рынке поиска: так, в августе 2014 года она составляла более 60%, тогда как у Google было 30%. На Android доля «Яндекса» гораздо скромнее: 28,1% против 71,1% у Google по итогам июля, по данным LiveInternet. Такому результату могли способствовать действия американского поисковика: в феврале 2014 года «Яндекс» пожаловался в Федеральную антимонопольную службу (ФАС), обвинив Google в давлении на вендоров, производящих мобильные устройства на базе Android. В частности, Google не предоставлял альтернативу своей поисковой системе и не позволял производителям предустанавливать на них приложения других компаний. Конфликт длился больше трех лет, но в апреле 2017 года ФАС и Google объявили о мировом соглашении.

Новый поисковой алгоритм полностью основан на нейросети. Нейросеть – это компьютерная модель, которая работает по тому же принципу, что и человеческий мозг. Она принимает решения исходя из множества факторов, а главное – самостоятельно устанавливает закономерности и учится на своих ошибках.

В новом алгоритме разработчики увеличили влияние нейронных сетей на выдачу. Так, предыдущая программа «Палех» обрабатывала только 150 страниц по запросу. «Королёв» воспринимает 200 тысяч страниц. Раньше нейросеть оценивала схожесть пары «запрос – заголовок страницы», а теперь также учитывает тело страницы. А ещё новый алгоритм определяет смысл запросов, по которым люди переходят на страницу.

На основании пользовательского поведения специалисты и создают обучающие примеры, или эталоны для нейронной сети. Например, если по какому-либо запросу пользователи очень часто кликают на одну и ту же страницу, то, вероятнее всего, запрос и текст страницы связаны по смыслу, то есть семантически.

«Умный» алгоритм понимает не только смысл выражений, но и визуальный контент. Это должно повысить эффективность поиска по картинкам, т.к. помимо заголовка / текста статьи, Яндекс сможет учитывать само изображение на картинке.

«Королёв» обучался на big data массивах пользовательской информации по принципу DSSM. Создатели демонстрировали нейронной сети огромные количества правильных и ложных примеров. После тренировки программа научилась сравнивать смысл текста по семантическому вектору в 300 чисел. Чем ближе по смыслу тексты запроса и потенциального ответа, тем ближе их семантические векторы.

Основная задача: улучшение качества выдачи по многословным низкочастотным запросам, по которым качество выдачи было низким (заведомо хуже, чем у Google – основного конкурента в рунете). В данный сегмент фраз часто попадают и голосовые запросы, заданные с переносных устройств на естественном языке (растущий спрос).

Что отличает данный алгоритм «Королёв» от предыдущего «Палеха»? Добавление в общий набор факторов ранжирования, которые учитывают:

- схожесть того «смысла», который скрыт в поисковой фразе и «смысла» всего документа, а не только заголовка окна браузера Title;
- качество ответа документа на схожие по «смыслу» запросы пользователей;
- новая техническая реализация с расчётом ряда факторов на этапе индексирования и внедрением дополнительного индекса.

От семантического поиска следует ожидать изменений в некоторых факторах ранжирования.

При этом отмечается и понижающее действие алгоритма:

- SEO-тексты не смогут дать прежнего эффекта. Потребуется новые подходы к оптимизации;
- некоторые вебмастера навязчиво вплетают не относящиеся к тематике статьи / ресурса популярные инфоповоды. Такое будет исключаться из выдачи.

Алгоритм воздействует, в первую очередь, на низкочастотные длинные запросы. А такие фразы задают обычно пользователи голосового поиска. Для понятных и привычных индентов «Королев», вероятнее всего не приведет дополнительного трафика. Это тематики такси, одежды, промышленных товаров с техническими характеристиками и артикулами.

Для пользователей Яндекса, в первую очередь, стоит ожидать прироста трафика в сложных и непонятных для обычного пользователя тематиках – юриспруденция, медицина, искусство и так далее. Когда посетитель сайта ищет контент естественным языком, без применения терминологии. И не имеет значения – коммерческий сайт или нет. Реакция рынка на представления данного алгоритма была однозначна, акции Яндекса резко возросли и пробили максимальную отметку.

С точки зрения пользователей, бывает странно, что одинаковые по смыслу, но отличающиеся по написанию запросы выдают разные результаты. Многие пользователи задают запросы в поисковую систему так, как будто спрашивают у друга. Новому алгоритму отвечать на эти запросы будет проще.

Вебмастерам же придётся делать более хорошие продукты, создавать качественный контент и не думать о том, чтобы специально продвигать свой сайт в поисковых системах. В реальности им часто приходится подстраивать тексты и сам сайт под поисковые машины.

Часть методов, которые помогали продвигать сайты раньше (например, SEO-копирайтинг) уже не дадут такого эффекта. Конечно, у «черных» SEO-оптимизаторов будут попытки обхитрить и новый алгоритм, но часть сил все же будет направлена на создание качественного контента.

Пока судить о качестве нового алгоритма рано, но чем больше ответов он даст, тем лучше он станет. Поэтому в перспективе пользователи должны почувствовать разницу.

Конечно, предстоит серьезно апробировать механизм «умных» запросов и точность выдачи нейронной сети и смыслового анализа на основе поисковой статистики в собственных проектах. Главный вопрос: а можно ли научить поисковую систему оперировать смыслами на уровне человека, как заявляет «Яндекс»? На взгляд мировых нейрофизиологов и ученых – это невозможно в ближайшее время, так как не родился ещё человек на Земле, который бы раскрыл все механизмы сознания и человеческого мышления, оперирующего смыслами.

Но всё же, технологии развиваются с огромной скоростью. Интернет всё больше «вливается» в жизнь каждого человека. И очень важно чтобы Интернет понимал каждого пользователя и давал ему всё что требуется.

Список литературы

1. Блог о технологиях Яндекса – [Электронный ресурс] / Режим доступа: <https://yandex.ru/blog/yatechnologies/yandeks-predstavil-novyy-algoritm-korolev>

2. Анна Ященко, «Королев» - все про новый алгоритм яндекса в одном месте – [Электронный ресурс] / Режим доступа: <https://seoandme.ru/korolev-algoritm-yandeksa.html>

3. Тимур Гаскаров, Как яндекс научил искусственный интеллект понимать смысл документов – [Электронный ресурс] / Режим доступа: <https://habr.com/company/yandex/blog/336094/>

4. Виктория Кравченко, В «Яндексе» запустили новый поисковый алгоритм «Королёв» – мнения экспертов – [Электронный ресурс] / Режим доступа: <https://rb.ru/opinion/s-yandeksom-v-kosmos/>