

Поиск в интернете: что и как ищут пользователи

По данным поиска Яндекса Осень 2008 года

Этот информационный бюллетень посвящен основному сервису Яндекса – поиску. Описать все, что связано с поиском в интернете, в одном исследовании невозможно. Поэтому наш отчет рассказывает только о том, что и как люди ищут в интернете – о формулировках поисковых запросов, ошибках и опечатках, появлении новых слов и новых запросов. В нем нет информации о релевантности выдачи, алгоритмах ранжирования и прочих вещах, интересных, например, вебмастерам, оптимизаторам и разработчикам поисковых систем. Исследование адресовано всем пользователям интернета, по крайней мере, тем, которые хотя бы иногда что-то ищут.

Все выводы основаны на данных поиска Яндекса. Информационный бюллетень включает в себя данные с сентября 1997 года по сентябрь 2008 года.

Содержание

1. Введение	1
2. «Внешний вид» поисковых запросов	2
3. Ошибки и опечатки	5
4. Популярные запросы	7
5. Региональные запросы	8
6. Всплески интереса	9
7. Появление новых слов	10
8. Сезонные колебания	11
Приложение 1. Основные факты и цифры	13

1. Введение

В 1997 году, когда поисковая система Яндекс только появилась, в Рунете было около 18 тысяч сайтов и около 3 миллионов страниц. Объем всех текстов Рунета составлял 7,5 Гб — то есть все существовавшие на тот момент сайты легко бы поместились сейчас на одну хорошую флешку. Количество пользователей Рунета не превышало 200 тысяч, и в день они задавали около тысячи запросов к Яндексу. Теперь в Рунете как минимум 13 миллионов сайтов, а полугодовая аудитория, по данным ФОМ, — более 32 миллионов человек.

По оценке Яндекса, сегодня **пользователи русскоязычного интернета просматривают страницы результатов поиска всех поисковых машин более 1,9 миллиарда раз в месяц**. Именно этот показатель — *число показов страниц результатов поиска* —

часто считают *числом поисковых запросов*. Реальных запросов меньше. К поисковым системам Рунета задают, по оценке Яндекса, около 48 миллионов запросов в день (при ежедневном просмотре страниц результатов поиска более 63 миллионов раз).

За поиском нужного ответа средний пользователь проводит меньше пяти минут. В день — тот день, когда среднему пользователю нужно что-то найти — он обращается к поиску один или два раза и делает в среднем пять запросов.

По данным TNS на август 2008, ежемесячно поиском на Яндексе пользуются 66,7% российских пользователей интернета, поиском на Mail.ru (по технологии Яндекса) — 27,2%, поиском на Рамблере — 28,3% (некоторые пользователи регулярно ищут с помощью нескольких поисковых систем). Поисковая система Google в исследованиях TNS не участвует.

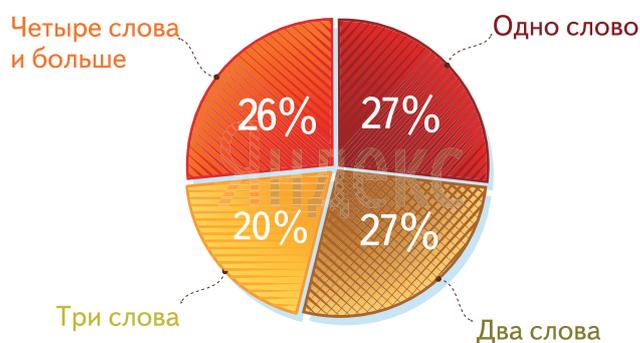
Аудитория поиска Яндекса на 53,6% состоит из мужчин и на 46,4% — из женщин, аудитория поиска на Mail.ru — больше женская (44,2% мужчин и 55,8% женщин), а поиском Рамблера мужчины и женщины пользуются практически поровну — 48,1% мужчин и 51,9% женщин.

Распределение пользователей поиска по возрастам похоже у всех трех порталов — больше всего ищут люди в возрасте от 25 до 34 лет (29-30%).

2. «Внешний вид» поисковых запросов

Типичный поисковый запрос состоит из двух или трех существительных, написанных кириллицей. Например, [реферат психология]¹.

Рис. 1. Количество слов в поисковых запросах



По данным поиска Яндекса, август 2008 г.

Средняя длина поискового запроса — **2,5 слова**. За одиннадцать лет работы поиска Яндекса пользователи стали многословнее — в 1997 году, когда поиск на www.yandex.ru только появился, средний запрос состоял всего из 1,2 слова. Теперь количество информации в интернете во много раз увеличилось, и для того чтобы найти ответ на свой вопрос, бывает нужно сформулировать запрос точнее.

¹ В тексте этого исследования поисковые запросы записываются в квадратных скобках. Части поисковых запросов или запросы, записанные не в точности так, как они были заданы, выделены курсивом. Количество запросов за период, о котором идет речь, может быть указано в тексте, а может — справа и сверху от запроса (как показатель степени).

Таблица 1. Рейтинг вопросительных слов

№	Вопрос	Количество за месяц	Примеры (количество запросов указано в среднем за месяц)
1.	как	4597176	[как правильно] ¹⁴¹⁹²⁶ [как целоваться] ⁹⁵⁷⁶ [как заработать в интернете] ⁸⁴⁷² [как завязать галстук] ⁸⁵⁴¹
2.	что	1504560	[что делать] ⁶¹⁷⁴⁸ [чем открыть] ²⁷¹⁷⁸ [что будет] ¹⁷⁰⁷⁷⁸² [чем закончится сериал] ⁴⁶⁴⁶
3.	где	949088	[где можно покупаться] ¹²²¹⁰ [где скачать песни] ²¹⁶¹ [где купить платье] ¹⁴⁹⁸
4.	какой	434165	[какие документы] ⁴⁵⁰⁰ [какие вопросы задать] ⁷³⁸ [какой антивирус лучше] ⁷²¹
5.	кто	420373	[кто с кем] ⁴¹⁸¹⁰⁹ [кто виноват] ¹⁷⁶⁷ [кто такие готы] ¹⁴⁸³
6.	сколько	406462	[сколько времени] ⁷⁷⁴⁸ [сколько стоит билет] ³⁰⁵³ [сколько варят кукурузу] ²⁵³
7.	когда	209951	[когда стричься] ¹⁴⁵¹ [когда запустят коллайдер] ¹¹⁶ [когда будет солнечное затмение] ¹³⁴
8.	куда	150074	[куда вложить деньги] ³⁹⁵¹ [куда поступать] ¹⁶⁵⁵ [куда обращаться] ⁹⁹⁰
9.	почему	130550	[почему нельзя] ²¹⁵⁴ [почему началась война] ³⁹ [почему желтеют листья] ⁵⁰
10.	зачем	55130	[зачем жить] ¹⁹¹ [зачем учиться] ³⁰⁸
11.	откуда	27314	[откуда берутся дети] ¹³⁸⁰ [откуда пошла фамилия] ³⁷³ [откуда происходит слово] ⁴¹⁶
12.	чей	15240	[чей номер] ¹⁸⁰² [чей туфля] ⁵⁰

По данным открытой статистики сервиса wordstat.yandex.ru, август 2007 — август 2008

Около 2,5% запросов сформулированы как вопрос. То есть это обычные вопросы, и пользователи, которые их задают, общаются с поисковой системой так, будто это живой человек. У Яндекса даже спрашивают [как пройти в библиотеку?] — в среднем 41 раз в месяц, [зачем герасим утопил муму?] — 53 раза и [кто убил лору палмер?] — 107 раз.

Самый популярный вопрос, с которым пользователи обращаются к Яндексу, — как. Вопросов, начинающихся с *как*, задают больше, чем вопросов, начинающихся со всех остальных вопросительных слов вместе взятых.

Интересно, что вопрос *что* значительно популярнее вопроса *кто*, и в частности **вопрос [что делать] интересует пользователей больше вопроса [кто виноват]**.

Самые короткие составляющие запросов к Яндексу — отдельные буквы и цифры. Пользователей интересуют все буквы русского алфавита, больше всего — предлоги и однобуквенные союзы (например, *и* в августе 2008 искали 4385966 раз), а меньше всего — *ъ* (меньше 9 тысяч запросов в месяц).

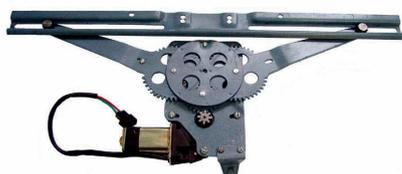
Наиболее длинные осмысленные слова — как правило, сложные существительные, состоящие из нескольких корней.

Самое длинное слово, заданное в качестве запроса к Яндексу в августе 2008 года, состоит из 37 символов — *гиппотомомонстресеквициклопедофобия*. По этому запросу находится 4583 страницы (на сентябрь 2008 года). Среди самых длинных запросов, на которые существуют ответы в интернете, преобладают различные химические соединения (*этилоксиэтилпарафенилендиаминсульфат* — 35 символов), названия компаний (*средневожжсельэлектросетьстрой* — 32 символа) и разного рода фобии (например, *гексакосийгексеконтагексофобия* — 31 символ).

«Гиппотомомонстресеквициклопедофобия», как ясно из названия, — боязнь произнесения новых слов. «Гексакосийгексеконтагексофобия» — боязнь числа 666. В интернете еще встречается слово «гексакосийгексеконтагексапараскаведекатриафобия» (48 символов) — боязнь числа 666 в пятницу, 13 числа. Однако эту фобию в Яндексе пока не искали.

В запросах на поиск картинок самые длинные запросы — это *электростеклоподъемник* и *электроводонагреватель* — по 22 символа (см. рис. 2 и рис. 3).

Рис. 2. Электростеклоподъемник



Результат с сайта www.avtoalarms.ru

Рис. 3. Электроводонагреватель



Результат с сайта www.amtenzo.ru

По данным поиска Яндекса, 14 сентября 2008 года

Для общения с поисковой машиной чаще всего используют существительные — эту часть речи содержат 75% запросов к поиску. Вторая по распространенности часть

речи — прилагательные, они присутствуют в 16% запросов к веб-поиску. Глаголы используют только в 5% случаев, а наречия есть менее чем в 1% запросов.

Большинство пользователей не позволяют себе грубостей с поисковой системой — **матерную лексику содержит сравнительно небольшое количество запросов** — полтора процента. Тем не менее матерных слов больше, чем, например, наречий или числительных (количественных и порядковых вместе взятых). Числительные присутствуют всего в 0,2% запросов.

Свои запросы к поиску русскоязычные пользователи в основном пишут на кириллице (более 72% запросов). Латиницей написаны около 15% запросов, а около 11% содержат и кириллические, и латинские символы. Чуть более 1% — это запросы из одних чисел, пустые или бессмысленные.

Приблизительно в каждом десятом поисковом запросе присутствует название организации или сайта. В этом случае пользователь часто хочет найти вполне определенный сайт, и все остальные результаты поиска его не интересуют.

Иногда при формулировке поисковых запросов пользователи явно указывают желание получить (купить, скачать) или продать что-то. Запросов с такими уточняющими словами около 4%.

Интересно, что глагол *купить* в запросах встречается в три раза чаще, чем *продать*, а существительное *продажа* — в восемь раз чаще существительного *покупка*. Запросы [купить] и [продажа] синонимичны: и в том, и в другом случае пользователь хочет купить что-то и ищет предложение товара. Так что спрос в интернете существенно превышает предложение.

Таблица 2. Уточняющие слова

Уточнение	Доля запросов	Примеры распространенных запросов (количество на август 2008)
скачать	3,3%	[игры скачать] ⁸⁴⁸⁶⁰⁰
бесплатно	1%	[фильмы бесплатно] ²⁹⁸⁶⁷⁹
купить, куплю	0,6%	[купить москва] ⁸³³⁷⁷
цена, цены	0,4%	[цена билета] ⁴⁹¹⁷⁵
продать, продам, продаю	0,1%	[продаю квартиру] ³²⁴⁷⁴

По данным поиска Яндекса, 2008 г.

3. Ошибки и опечатки

14-15% всех запросов содержат различные ошибки и искажения. Ошибки в поисковых запросах возникают по разным причинам — не только из-за неграмотности, но и из-за стремления найти ответ как можно быстрее, неверной раскладки клавиатуры и т.п.

Две трети всех ошибок — орфографические (пропущенные буквы, неправильное написание слов и т.п.).

Самая распространенная ошибка сейчас — *однокласники* (с одной с). Запрос [однокласники] встречается в 8 раз реже грамотного запроса [однокласники], но входит в топ-10 запросов к Яндексу в августе 2008. Это единственный запрос с ошибкой среди ста самых частотных поисковых запросов.

Бывают слова, для которых нет устоявшейся нормы написания. Например, Яндекс знает 12 вариантов написания слова *риэлтор* — от распространенного [риэлтор] до экзотических [риэлтар] и [реелтр] (см. таблицу 3). Также бывают слова, которые чаще пишут неправильно. Одно из таких сложных слов — *трансъевропейский*. В поисковых запросах вариант [трансевропейский] встречается в полтора раза чаще словарного, и сайтов по нему находится почти в два раза больше.

По поисковым запросам можно понять, какой способ написания слов, недавно появившихся в русском языке, выбирают пользователи.

Варианты [тег]²⁶⁴²⁵, [бренд]⁵¹³³³ и [тренд]⁷⁹⁸⁶ победили варианты [тэг]³⁶⁴⁰, [брэнд]⁷⁷⁰⁴ и [трэнд]³⁶¹.

Большинство пишет [браузер]⁴⁹¹¹⁴, а не [броузер]³⁷⁰⁰ и [фитнес]²¹⁰³⁶¹, а не [фит-несс]³²⁹⁸⁹.

Таблица 3. Варианты написания слова «риэлтор»

Запрос	Среднее число запросов в месяц
риэлтор	22125
риелтор	4331
риэлтер	3046
риелтер	643
реэлтор	527
реелтор	309
риэлтр	132
реэлтер	71
реелтер	56
риелтр	40
реелтр	14
риэлтар	13

По данным открытой статистики сервиса wordstat.yandex.ru, август 2007 — август 2008

Четверть ошибок в поисковых запросах пользователей связана с лишними и пропущенными пробелами или со смысловыми ошибками — когда ошибка или опечатка в одном слове меняет смысл запроса. Например, [скачатьбесплатно] вместо [скачать бесплатно] или [дипломные роботы] вместо [дипломные работы].

И ещё около 5% ошибок возникают из-за неверной раскладки клавиатуры.

Для некоторых пользователей не существует адресной строки браузера, поэтому они вводят адреса сайтов в строку поиска. Таких запросов достаточно — около 4%, из них 1,3%

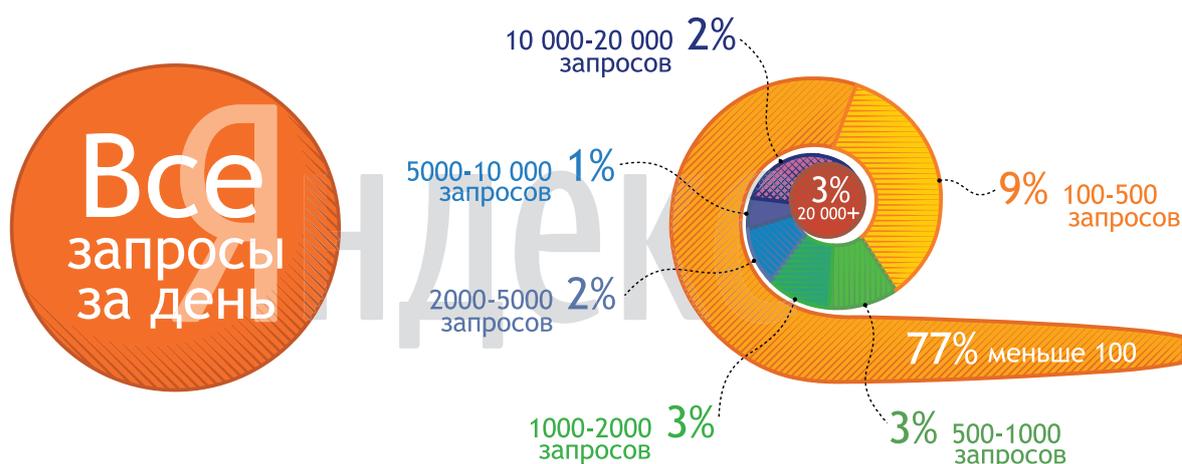
содержат `www`, например, [`www.mail.ru`] и 2,6% — не содержат, например, [`yandex.ru`]. Адреса электронной почты есть в 0,2% запросов.

4. Популярные запросы

Самые популярные запросы — только очень небольшая часть всех запросов к поиску. По данным на август 2008, общее количество запросов из первой сотни составляет около семи процентов от всех поисковых запросов (около 1,8 миллиона запросов к Яндексу в день). Большинство запросов достаточно редки — каждый из них задается реже ста раз в день. Именно такие, низкочастотные, запросы составляют три четверти поисковых запросов (более 20 миллионов в день).

На рисунке 4 показано, сколько процентов составляют запросы различной частотности от общего ежедневного количества поисковых запросов.

Рис. 4. Длинный хвост поисковых запросов



По данным поиска Яндекса, август 2008

Список самых популярных поисковых запросов в целом изменяется крайне медленно. На топ-10 не влияют ни новости, ни сезонность — речь идет о десятках тысяч одинаковых запросов в день. Однако за последние полтора года самые часто задаваемые запросы к Яндексу изменились достаточно сильно (см. таблицу 4). Это первое серьезное изменение первой десятки за несколько лет. Основная причина этих изменений — появление и огромная популярность социальных сетей. Первый раз за несколько лет запрос [порно] перестал быть самым популярным, уступив запросам [одноклассники] и [в контакте]. В топ-50 самых частых запросов в августе 2008 года входит двенадцать запросов, связанных с социальными сетями.

Из десяти запросов, наиболее распространенных в январе 2007, сильно опустились только два — [скачать] и [ваз], на 108 и 533 место соответственно. Остальные находятся в первой или второй десятке рейтинга.

Таблица 4. Топ поисковых запросов

№	Январь 2007	Август 2008
1.	порно	одноклассники
2.	погода	в контакте
3.	знакомства	порно
4.	mail.ru	погода
5.	скачать	mail.ru
6.	работа	вконтакте
7.	сонник	контакт
8.	ваз	одноклассники
9.	из рук в руки	mail
10.	гороскоп	зайцев нет

} 2,7% запросов

По данным поиска Яндекса

Рейтинг запросов, которые пользователи задают к поиску с мобильных телефонов, выглядит несколько иначе — там социальная тематика пока не победила сексуальную.

5. Региональные запросы

Обычно пользователи из разных регионов ищут в интернете похожие вещи. Но, конечно, в каждом регионе существуют специфические, «местные» запросы — например, названия аэропортов или магазинов. Чтобы выделить такие запросы, можно использовать специальный показатель — *региональность*.

Региональность запроса — это отношение частоты, с которой задавали этот запрос пользователи определенного региона, к количеству всех запросов из этого региона. Чем больше это отношение (то есть доля того или иного запроса среди всех запросов), тем более этот запрос «местный», то есть наиболее популярный в определенном регионе. Сравнение показателей региональности в разных регионах позволяет увидеть, какие запросы характерны для того или иного региона.

Как правило, доли самых распространенных запросов оказываются больше там, где не так много пользователей — в небольших городах или в других странах. Так, например, запросы [в контакте] и [вконтакте] наиболее свойственны Гатчине — конечно, здесь не самое большое количество пользователей этой социальной сети, но самая большая концентрация интереса к ней. Запрос [зайцев нет] свойственен Люберцам, [порно] — Нидерландам, а [работа] и [секс] — Индии.

Для крупных городов локальные запросы часто содержат указание города или региона ([новосибирский зоопарк], [знакомства в ростове]) или местных брендов. Например, один из кировских запросов — [квадрат] — означает не геометрическую фигуру, а сеть магазинов техники. Среди локальных часто оказываются запросы со спортивными клубами ([фк урал], [рубин]), университетами ([двгу], [пгту]), названиями аэропортов ([толмачево], [аэропорт курумоч]) и т.п.

Уточнить поисковый запрос можно не только непосредственно его формулировкой, но и при помощи дополнительных интерфейсных возможностей поисковой системы. Например, в Яндексе можно ограничить поиск по тому или иному месту, поставив галочку «искать только в определенном регионе». Этой возможностью пользуются в среднем около 5% посетителей Яндекса. В большинстве крупных городов этой опцией пользуются чаще, чем в среднем. Интересно также, что пользователи из Санкт-Петербурга достаточно редко указывают свой город непосредственно в запросе, но при этом ежедневно около 15% петербургских запросов ограничены по этому региону. Вообще самый популярный регион при ограничении поиска — Северо-Запад.

6. Всплески интереса

При помощи поисковых запросов можно увидеть, как меняется интерес пользователей к различным понятиям и событиям. Для этого в январе 1999 года был придуман НИНИ-индекс — индекс Непостоянства Интересов Населения Интернета. Он показывает, к каким запросам интерес пользователей внезапно вырос, а к каким — упал.

Резкое появление интереса со стороны пользователей связано с различными новостями — общественно-политическими или светскими.

В июле 2008 года никаких политических событий не происходило, и наибольший всплеск интереса вызвала радиопередача с участием Кати Гордон и Ксении Собчак (запросы [ксения собчак], [собчак], [собчак гордон], [гордон] и т.п.) Другие июльские всплески интереса были вызваны смертями Нонны Мордюковой и Михаила Пуговкина.

В августе 2008 года пользователей интересовали уже совсем другие вещи — российско-грузинский конфликт ([война в осетии], [война в южной осетии], [война грузия]) и олимпиада ([дневник олимпиады], [медали олимпиады], [итоги олимпиады]). На третьем месте — свадьба Ксении Бородиной.

Кроме новостей и важных событий резкие всплески интереса вызывают мемы — практически любая информация, которая неожиданно становится популярной и мгновенно распространяется среди пользователей интернета. Например, слово «превед», вопрос «как вы относитесь к пробуждению Ктулху?» и т.п. Одно из последних таких явлений — запуск большого адронного коллайдера в Швейцарии, который готовился летом 2008 года. Еще год назад адронными коллайдерами никто, кроме нескольких специалистов, не интересовался — ежемесячное количество запросов не превышало нескольких десятков. Этим летом пользователи задавали десятки тысяч запросов [адронный коллайдер] и [большой адронный коллайдер].

Кроме того, пользователи искали [адронный коллайдер]²⁰¹³² и [андроидный коллайдер]¹⁸²³ (количество запросов за август 2008).

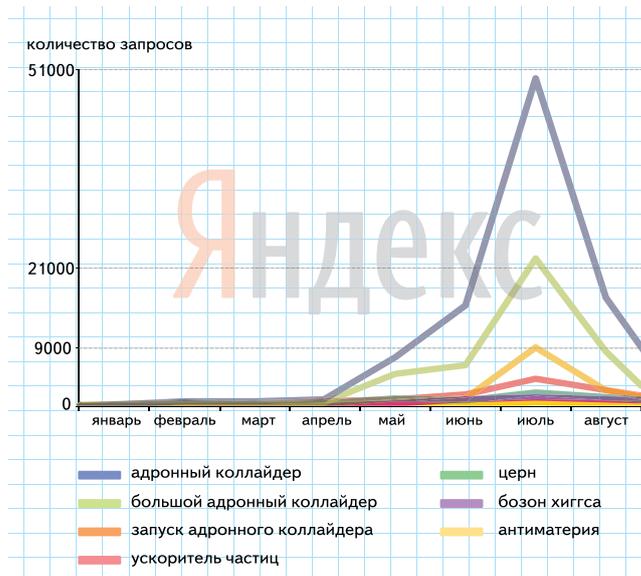
Среди *связанных* запросов — то есть таких, которые пользователи искали в рамках одной поисковой сессии

Большой адронный коллайдер (англ. *Large Hadron Collider, LHC*) — ускоритель, предназначенный для разгона протонов. Коллайдер построен в исследовательском центре Европейского совета ядерных исследований CERN (Conseil Européen pour la Recherche Nucléaire). Первоначально его запуск был назначен на лето 2008 года, а потом перенесен на сентябрь. Некоторые люди считали, что после запуска коллайдера образуется черная дыра и наступит конец света.

с запросом [коллайдер], — были запросы [черная дыра] и [конец света].

Кроме того, многие из тех, кто искал [большой адронный коллайдер], искали также различные водонагреватели — видимо, потому что слово «бак», кроме всего прочего, — сокращение от большого адронного коллайдера. Увидев это словосочетание в результатах поиска, некоторые пользователи решили узнать, что оно значит. Люди, искавшие просто [адронный коллайдер], нагревателями не интересовались.

Рис. 5. Большой адронный коллайдер

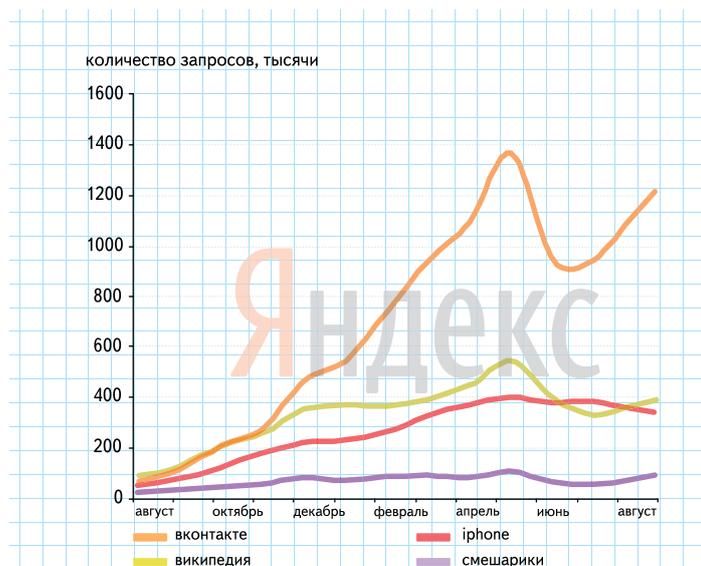


По данным открытой статистики сервиса wordstat.yandex.ru, январь 2007 — август 2008

7. Появление новых слов

Поисковые запросы также показывают, как появляются новые предметы и явления, а вместе с ними — новые слова в языке. До 2006 года, пока не появилась социальная сеть Вконтакте.ру, запроса [вконтакте] без пробела не существовало, а теперь он — один из самых часто задаваемых.

Рис. 6. Появление новых поисковых запросов



По данным открытой статистики сервиса wordstat.yandex.ru, август 2007 — август 2008

8. Сезонные колебания

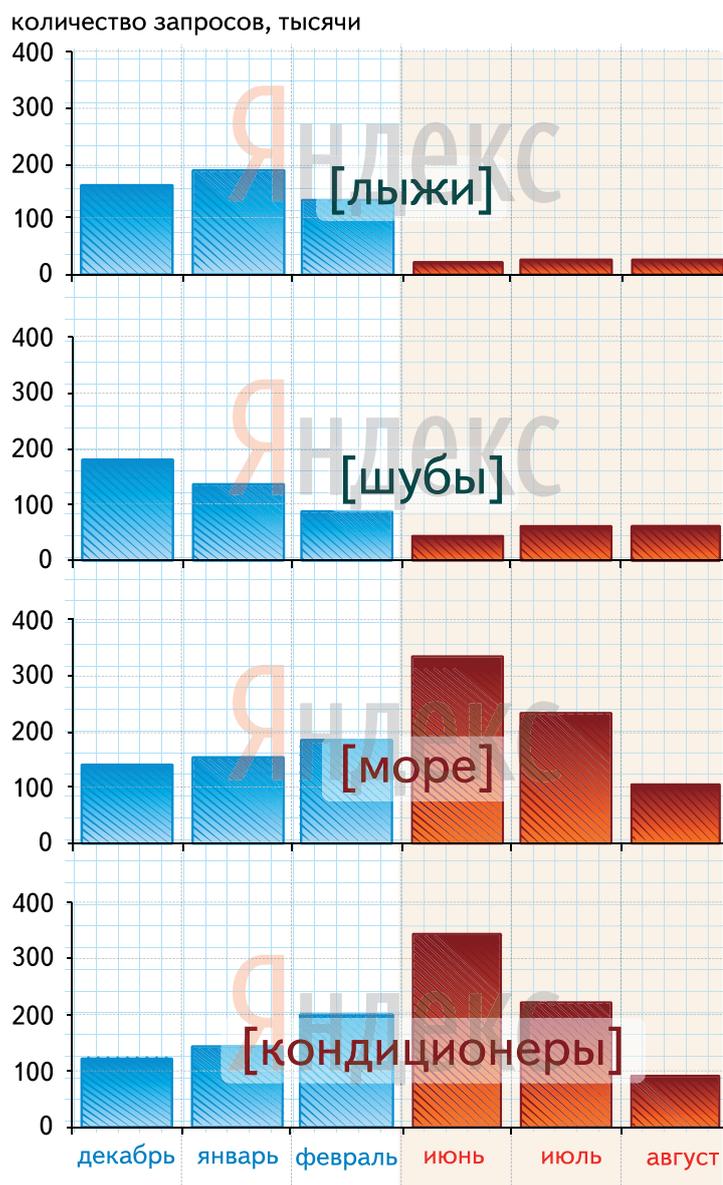
Очевидно, что интересы пользователей (и как следствие — поисковые запросы) зависят от времени года — так, например, зимой больше ищут катки, а летом — водоемы.

Для наблюдения за сезонными колебаниями подходит, например, школьная программа по литературе. По поисковым запросам, содержащим названия произведений, хорошо видно, когда люди начинают искать эти произведения в интернете.

Оказывается, почти никто, кроме школьников или их родителей, книгами из школьной программы не интересуется. Запросы к поиску на тему различных книг в течение всего года распределяются равномерно, и их количество относительно невелико — кроме одного месяца, когда наблюдается значительный всплеск интереса. Видимо, как раз в этом месяце книгу проходят в школе. Разница между максимальным и средним числом запросов достаточно большая.

В таблице 5 для всех произведений указано количество запросов за тот месяц, когда интерес к ним максимальный, и разница по сравнению со среднемесячным количеством запросов.

Рис. 7. Летние и зимние запросы



По данным открытой статистики сервиса wordstat.yandex.ru, август 2007 — август 2008

Таблица 5. Поиск литературных произведений
из школьной программы для 7-8 классов

Месяц	Произведение	Запрос	Количество запросов	По сравнению со средним
Сентябрь	Гомер, «Одиссея»	[одиссея гомер] другие запросы: [поэма гомера одиссея], [гомер илиада одиссея], [гомер одиссея краткое]	2408	+126%
Октябрь	Д. Фонвизин, «Недоросль»	[недоросль] другие запросы: [фонвизин недоросль], [комедия недоросль], [сочинения недоросль]	44555	+408%
Ноябрь	А. Пушкин, «Капитанская дочка»	[капитанская дочка] другие запросы: [сочинение капитанская дочка], [пушкин капитанская дочка]	28436	+227%
Декабрь	М. Лермонтов, «Мцыри»	[мцыри] другие запросы: [лермонтов мцыри], [поэма мцыри]	23855	+315%
Декабрь	Н. Гоголь, «Ревизор»	[ревизор] другие запросы: [гоголь ревизор], [ревизор хлестаков], [ревизор характеристики]	32500	+170%
Январь	М. Салтыков-Щедрин, сказки	[салтыков щедрин сказки] другие запросы: [анализ сказок салтыкова щедрина], [салтыков щедрин сказка премудрый], [салтыков щедрин сказка пескарь]	5754	+255%
Февраль	А. Толстой, «Князь Серебряный»	[князь серебряный] другие запросы: [толстой князь серебряный]	2227	+138%
Март	М. Горький, «Детство»	[горький детство] другие запросы: [м горький детство], [горький детство скачать]	5646	+233%
Апрель	А. Твардовский, «Василий Теркин»	[василий теркин] другие запросы: [твардовский василий теркин], [василий теркин кратко]	10878	+316%

По данным открытой статистики сервиса wordstat.yandex.ru, сентябрь 2007 — май 2008

Приложение 1. Основные факты и цифры

По оценке Яндекса, сегодня пользователи русскоязычного интернета просматривают страницы результатов поиска всех поисковых машин более 1,9 миллиарда раз в месяц.

За поиском нужного ответа средний пользователь проводит меньше пяти минут за один раз. В день — тот день, когда среднему пользователю нужно что-то найти — он обращается к поиску один или два раза и делает пять запросов.

Типичный поисковый запрос состоит из двух или трех существительных, написанных кириллицей. Например, [реферат психология].

Средняя длина поискового запроса — 2,5 слова. По сравнению с 1997 годом запросы стали вдвое длиннее — тогда средний запрос состоял всего из 1,2 слова.

Около 2,5% поисковых запросов сформулированы как вопрос. Самый популярный вопрос, с которым пользователи обращаются к Яндексу — *как*. Вопросов, начинающихся с *как*, задают больше, чем вопросов, начинающихся со всех остальных вопросительных слов вместе взятых. Вопрос *что* значительно популярнее вопроса *кто*, и в частности **вопрос [что делать] интересует пользователей больше вопроса [кто виноват]**.

Самое длинное слово, заданное в качестве запроса к Яндексу в августе 2008 года, состоит из 37 символов — *гиппотомомонстресесквиппедалиофобия*.

Для общения с поисковой машиной чаще всего используют существительные — эту часть речи содержат 75% запросов к «большому» поиску и 96% запросов к поиску по картинкам.

Большинство пользователей не позволяют себе грубостей с поисковой системой — **матерную лексику содержит сравнительно небольшое количество запросов** — полтора процента.

14-15% всех запросов содержат различные ошибки и искажения. Две трети всех ошибок — орфографические (пропущенные буквы, неправильное написание слов и т.п.).

Самая распространенная ошибка сейчас — *однокласники* (с одной *с*). Этот запрос входит в топ-10 запросов к Яндексу в августе 2008.

Для некоторых пользователей не существует адресной строки браузера, поэтому они вводят адреса сайтов в строку поиска. Таких запросов достаточно — около 4%.

По поисковым запросам можно понять, какой способ написания слов, недавно появившихся в русском языке, выбирают пользователи.

Варианты [тег]²⁶⁴²⁵, [бренд]⁵¹³³³ и [тренд]⁷⁹⁸⁶ победили варианты [тэг]³⁶⁴⁰, [брэнд]⁷⁷⁰⁴ и [трэнд]³⁶¹.

Аналитическая группа департамента маркетинга
компании «Яндекс»
analyst@yandex-team.ru