ЧЕРНЫЕ ЯЩИКИ
ГОВОРЯТ И ПОКАЗЫВАЮТ

Михаил Волович, Сергей Вальковский,
«Ашманов и партнеры»

«Поисковый маркетинг и продвижение бизнеса в Интернете».
Москва, Digital October, 27.11.2015

Ашманов
и партнеры

КАК МЫ ЭТО АНАЛИЗИРУЕМ

Лаборатория поисковой аналитики «Ашманов и партнеры»

- Лингвисты, математики, программисты, асессоры
  ... в контакте с SEO-специалистами
- С 2008 года
- Анализаторы качества поиска AnalyzeThis.Ru
  (русские, английские, китайские, вьетнамские)
- Поисковый антиспам
- Участие в разработке нескольких поисковых машин
  Flexum.ru, *Wada.vn* и *WadaMarket.com*

ФАКТОРЫ РАНЖИРОВАНИЯ

По результатам поиска пытаемся разобраться,
какие факторы ранжирования работают

... и на основе полученных знаний выдаем рекомендации

- «Черные ящики»: Яндекс и Google
- 160 запросов регулярно + запросы клиентов
- 300+ параметров:
  - «коммерческие»
  - технические
  - текстовые
  - ссылочные
- Динамика: выкачки ежемесячно с сентября 2014 года

СТАТМЕТОДЫ

Зависимость статистически значима?

- Есть ли связь между фактором и позицией?

  Ранговый критерий Спирмена, коэффициент корреляции Пирсона, коэффициент классической линейной регрессии (МНК), U-тест Манна-Уитни

- Связано ли с фактором попадание в «десятку»?

  Точный тест Фишера, критерий χ^2

- Связано ли с фактором попадание в «тридцатку»?

  Точный тест Фишера, критерий χ^2

НЕОБХОДИМЫЕ ОГОВОРКИ

Статистические методы:

- есть ли *статистически достоверная связь* между фактором и ранжированием?

Если нет, *скорее всего*, фактор «не работает».

Если есть,

- *вероятно*, фактор учитывается в ранжировании;
- *возможно*, он скоррелирован с фактором или группой факторов, участвующими в ранжировании;
- *не исключено*, что так случайно получилось (но *статистически достоверная связь* ⇔ вероятность этого очень мала).

Так что же делать?

- здравый смысл;
- эксперименты;
- влияя на «сопутствующий» фактор, влияем и на реально работающие;
- *и в конце концов, какая альтернатива?*

ПОПАДАНИЕ В ТОП и ПОЗИЦИЯ

Фактор может быть связан:

- с попаданием в топ
  (топ-30 vs. фон)
- с позицией внутри топа
  (корреляция с позицией,
  топ-10 vs. 11–30)

Возможны любые комбинации!

Что важнее?

зависит от запроса

Рабочая гипотеза:

- топ <= грубый фильтр
- позиция <= тонкое ранжирование

	Яндекс Позиция	**Яндекс Топ-30**	**Google Позиция**	**Google Топ-30**
Сертификаты	**Сильное**	—	*Слабое*	**Сильное**
- на странице	**(Сильное)**	**Сильное**	*Слабое*	—
- авто	*Слабое*	—	—	**Сильное**
Скидки	Есть	Есть	**Сильное**	—
- на странице	Есть	—	*Слабое*	—
- авто	Есть	—	**Сильное**	—
Видео	Есть	—	**Сильное**	**Сильное**
- на странице	(—)	—	Есть	**Сильное**
Отзывы	Есть	—	**Сильное**	—
- на странице	*Слабое*	—	—	—
- авто	Есть	—	**Сильное**	Есть

КОММЕРЧЕСКИЕ ФАКТОРЫ

Коммерческие запросы

- большая конкуренция
- активная оптимизация сайтов
- *поэтому* много релевантных страниц

Как поисковику выбрать лучшее

- тексты? — переоптимизированы
- ссылки? — заспамлены, естественных мало
- поведенческие факторы? — хороши для корректировки, а не как основа

«Коммерческие» факторы

- качество сайта
- качество бизнеса
- + разные технические вещи

АСЕССОРСКИЕ ОЦЕНКИ

Почему асессоры,
а не автомат

- Многие факторы — про сайт, а не про страницу
- Не зависит от *наших* технологий
- Есть асессорский опыт

Около 80 «ручных» факторов

- для сайтов
- для страниц
- + автоматические версии (для страниц)

Обратный звонок	**В обвязке**	Есть	Трудно найти	Нет	
- Обратный звонок на странице	**Есть**	Нет			

Представление товара

Цена (на сайте)	**Есть**	Нет	Не актуально		
- Цена (на странице)	**Есть**	Нет	Не актуально		
Фотографии	Одна	**Несколько**	Нет	Не актуально	
Описание	**Есть**	Нет	Не актуально		
Характеристики / параметры	**Есть**	Нет	Не актуально		
Сравнение товаров	**Есть**	Нет	Не актуально		
Количество брендов	Один	Несколько	**Много**	Не актуально	
Ассортимент	Ограниченный	Средний	Широкий	**Огромный**	Не актуально

Полезные функции и опции

Кнопка заказа	**Есть**	Нет	Не актуально
Калькулятор	**Есть**	Нет	Не актуально
Подбор по параметрам	**Есть**	Нет	Не актуально
Гарантия	**Есть**	Нет	Не актуально
Постпродажное обслуживание	**Есть**	Нет	Не актуально

Оплата

Оплата	**Есть**	Нет	Не актуально
- Оплата на странице	Есть	**Нет**	
- Оплата картой	**Есть**	Нет	

ФАКТОРЫ И РЕЗУЛЬТАТЫ

Многие коммерческие факторы действительно работают!

			Связаны с попаданием в топ (Яндекс, Google)			
			Связаны с позицией (Яндекс, Google)			
		Значимых факторов				
	Всего факторов		**Я**	**G**	**Я**	**G**
Контакты и формы связи	25	20	10	15	12	9
«Товар» и его представление	32	29	5	28	16	4
Контент и дополнительные возможности	31	23	12	15	7	11
Социальные сети	18	17	4	13	0	10

ЧТО НЕ РАБОТАЕТ

Проще перечислить, что не влияет!

Не нашли достоверной связи с ранжированием:

- телефон на сайте (*потому что везде есть*)
- многоканальный телефон
- Skype
- реклама
- Flash
- скорость загрузки

Однако:

- есть коллекции запросов, где связь видна
- кое-где связь была раньше
- может появиться завтра

ЯНДЕКС vs. GOOGLE

Они похожи! (*Больше, чем мы ожидали.*)

Многие факторы влияют по-разному

- Яндекс — позиция, Google — топ
- *или наоборот*

Некоторые различия

телефон на странице:	Я+	G–
оплата картой:	Я–	G+
соцсети:	Я(–)	G+
кнопки *Поделиться*:	Я–	G+
счетчики:	Я+	G–

(Яндекс-метрика, LI, Google analytics)

ПРИМЕР:
ОДНА ГРУППА ФАКТОРОВ

	Яндекс		Google	
Корреляция / **Фактор**	С позицией	С попаданием в топ	С позицией	С попаданием в топ
Наличие информации об оплате на сайте	Нет	Сильная	Сильная	Нет
• На странице	Нет	Нет	Сильная	Нет
• Автоопределение	Нет	Сильная	Сильная	Нет
• Оплата картой	Нет	Нет	Сильная	*Слабая*
• Оплата Яндекс.Деньгами	Нет	Сильная	Сильная	Нет
• Оплата WebMoney	Нет	*Слабая*	Сильная	Нет

Товар и его представление / Оплата

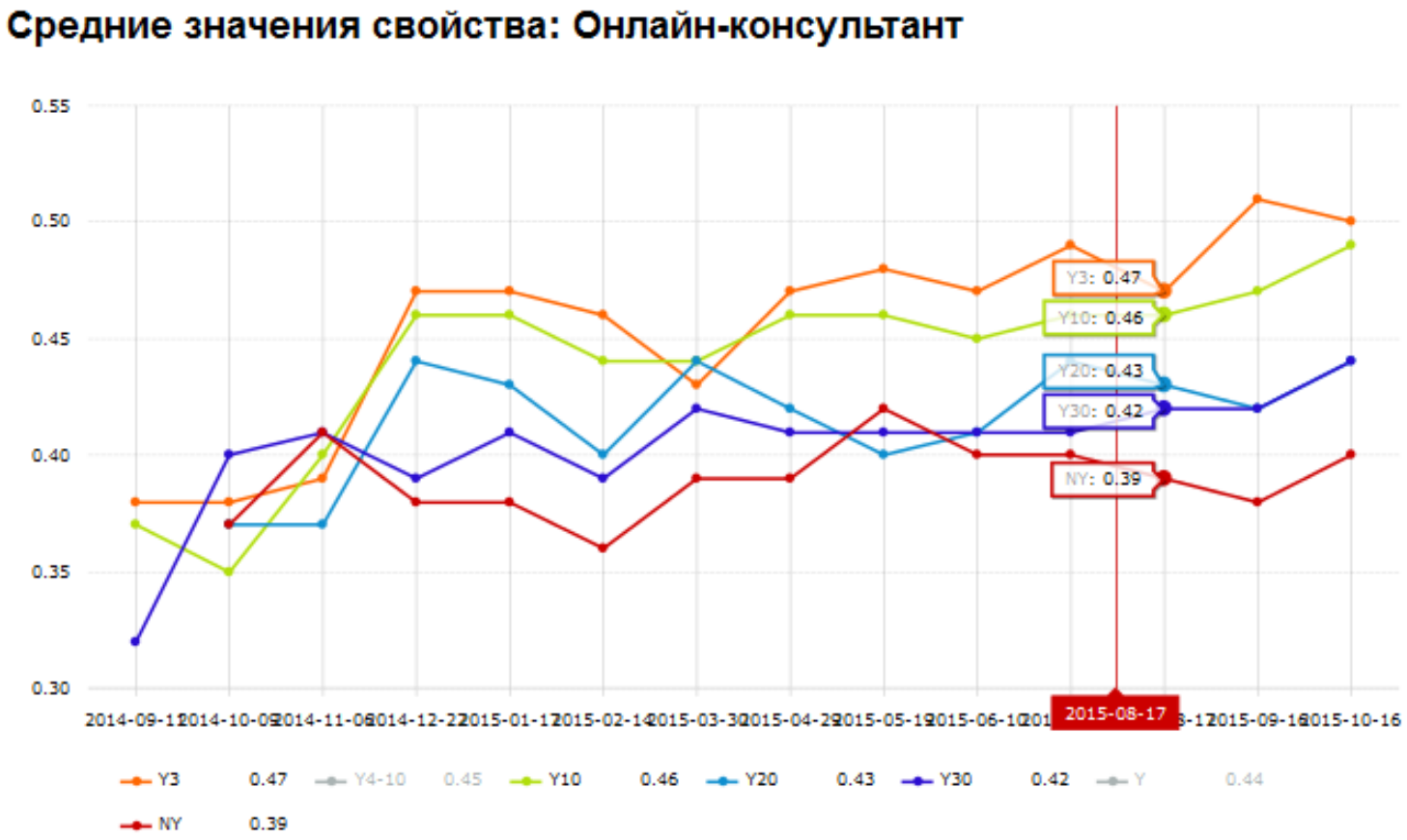
ДИНАМИКА И ТЕМАТИКИ

Уже год наблюдений

- Стабильный набор запросов
- Срезы раз в месяц
- Видно, какие факторы устойчивы
- «Сводки погоды»: рекомендации для SEO

Тематические выборки

- Данные для клиентов
- 100–500 запросов
- Автоматические параметры (ручные 3–10 на выборку)
- Там всё может отличаться



ТЕКСТОВЫЕ ФАКТОРЫ

Что

- точный запрос
- с точностью до форм
- «кворум»
- все слова запроса
- отдельные слова (с учетом веса)
- + синонимы
- + «дополнения»
- BM25

Где

- URL
- заголовок
- документ (текст)
- <h1>
- <h2>—<h4>
-
- <a>
- атрибут title

+ тексты внешних ссылок

ТЕКСТОВОЕ РАНЖИРОВАНИЕ: ЯНДЕКС vs. GOOGLE

Что	Где	Яндекс			Google		
		1–10	11–20	21–30	1–10	11–20	21–30
слова запроса	URL	0,48			0,44	0,38	0,37
точный запрос	<title>	0,46	0,43	0,41	0,47		
«кворум»		1,14	1,04	1,02	1,13	1,05	0,96
точный запрос	Текст	7,3	7	6,7	6,5	5,4	4,6
+ словоформы		12,1			11,4	9,3	7,7
«кворум»		18,3			18	14,5	11,7
слова запроса		32			32,4	27,5	22,6
Я-«дополнения»		4,4	4,1		4	3,9	2,7

BM25, ГОВОРИТЕ?

Прошли ли времена переоптимизированных текстов?

- Нужно ли нашпиговывать сайт ключевиками?
- Сколько запросов на килобайт?

Хорошие новости

«Плотность» роли не играет.
BM25 не работает!

Плохие новости

Хорошего запроса должно быть много.

Достаточно сделать текст большим

- Яндекс и Google любят огромные простыни-витрины
- Чем больше размер HTML и размер текста, тем лучше
  Яндекс: позиция, топ; Google: позиция
- Главные страницы сайтов тоже должны быть побольше
  чтобы больше ключевиков поместилось

ВРЕДНОЕ ПРОИЗВОДСТВО

№	Запросы	Вхождение слов запроса в URL	Вхождения в текст: точный запрос	+ словоформы	кворум	отдельные слова ↑	Вхождения в <title>: точный запрос	+ словоформы	кворум	отдельные слова
1	кресло руководителя	0.54	31.6	35.3	51.6	74.9	0.30	0.70	1.37	1.57
2	женские сумки	0.55	7.96	21.6	42.9	73.7	0.81	1.11	1.59	1.94
3	торшер	0.75	46.2	68.8	68.8	68.8	0.21	1.46	1.46	1.46
4	шкафы купе	0.67	19	62.4	62.4	68.7	1	1.60	1.60	1.63
5	люстры	0.62	28.0	66.3	66.3	66.3	1.10	2	2	2
6	бра	0.81	65.1	65.1	65.1	65.1	1.44	1.44	1.44	1.44
7	диваны и кресла	0.66	6.57	8.50	12.8	60.2	0.90	1.03	1.13	1.41
8	светодиодные лампы	0.74	13.7	28	44.1	59.2	1.10	1.15	1.40	1.33
9	купить люстру	0.56	2.30	3.50	12.8	57.9	0.60	0.70	0.93	1.70
10	универсальный фрезерный станок	0.68	4.74	9.52	15.1	57.7	0.22	0.67	0.85	1.31
11	мезотерапия	0.80	29.1	56.4	56.4	56.4	1.13	1.47	1.47	1.47
12	септик	0.83	25.8	53.8	53.8	53.8	0.96	2.21	2.21	2.21
13	токарно винторезный станок	0.65	13.5	24.1	31.6	52.5	0.45	0.95	1.09	1.21
14	токарный станок по металлу	0.76	3.54	7.38	10	51.3	0.35	0.88	1	1.46

Запросы — лидеры по вхождениям слов запроса («отдельные слова»)

Невеселая мораль: У кого *шкафов купе* меньше 60, в топ-30 не пустят!

ССЫЛОЧНЫЕ ФАКТОРЫ

Ссылочные факторы

- очень сильные
- в т.ч. у Яндекса, несмотря на ...
- скорее всего, *некоторые* ссылочные факторы были отключены

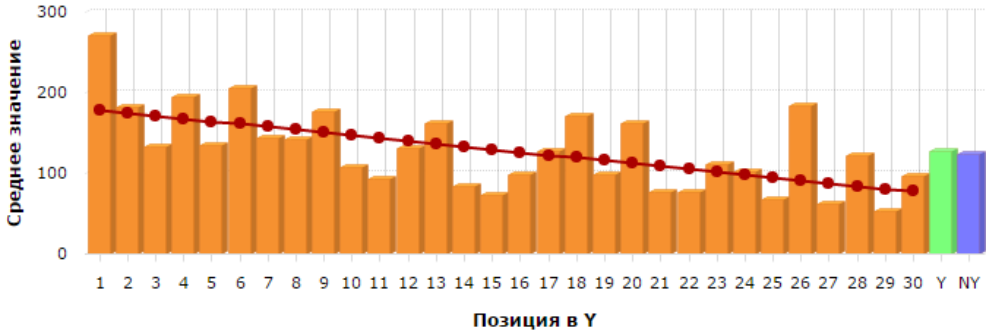
«Минусинск»

- сильно перемешал выдачу
- все факторы стали слабее, ссылочные особенно
- к октябрю всё выправилось на новом уровне

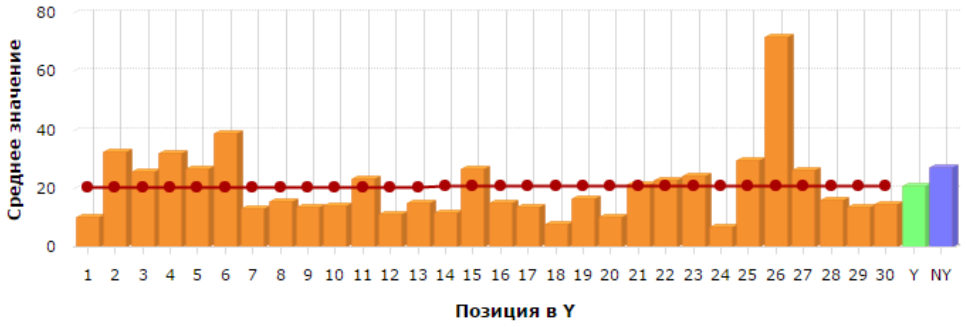
«МИНУСИНСК» ПОКАЗЫВАЕТ ЗУБЫ

Вхождения запроса («кворум») в ссылки на сайт: Яндекс

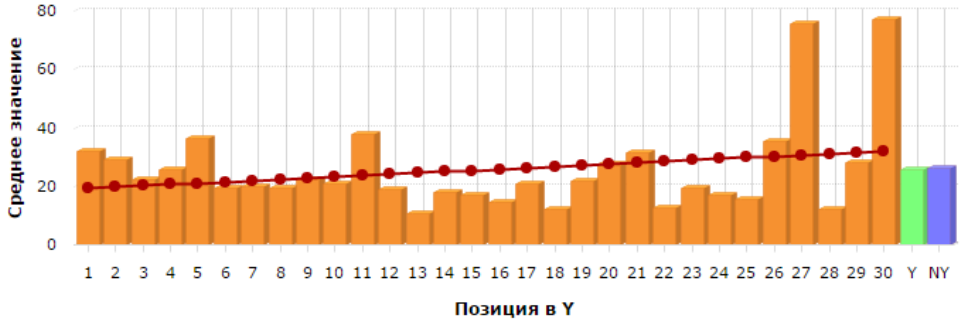
1. март



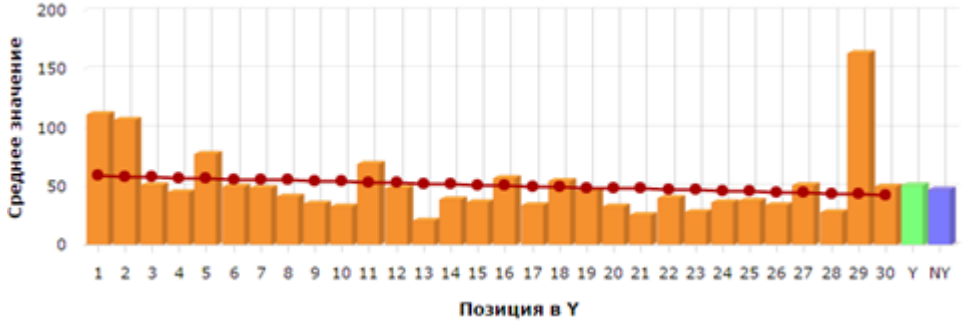
2. август



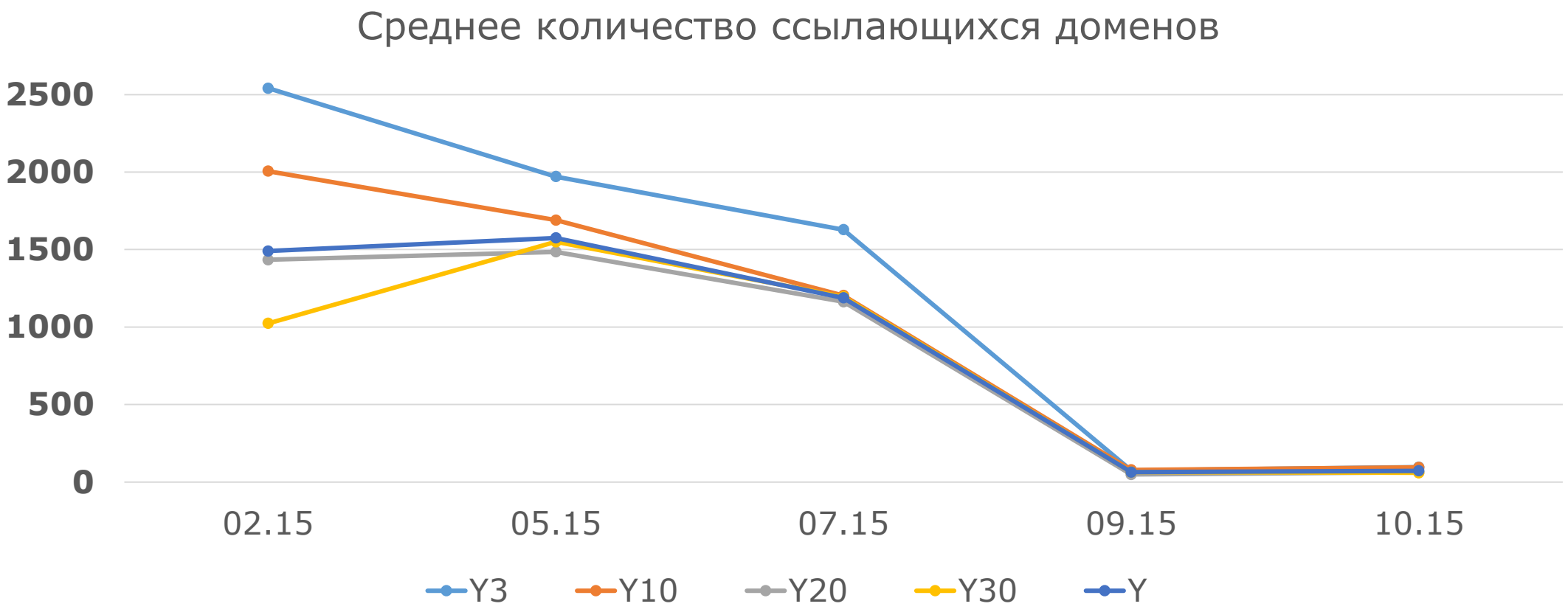
3. сентябрь



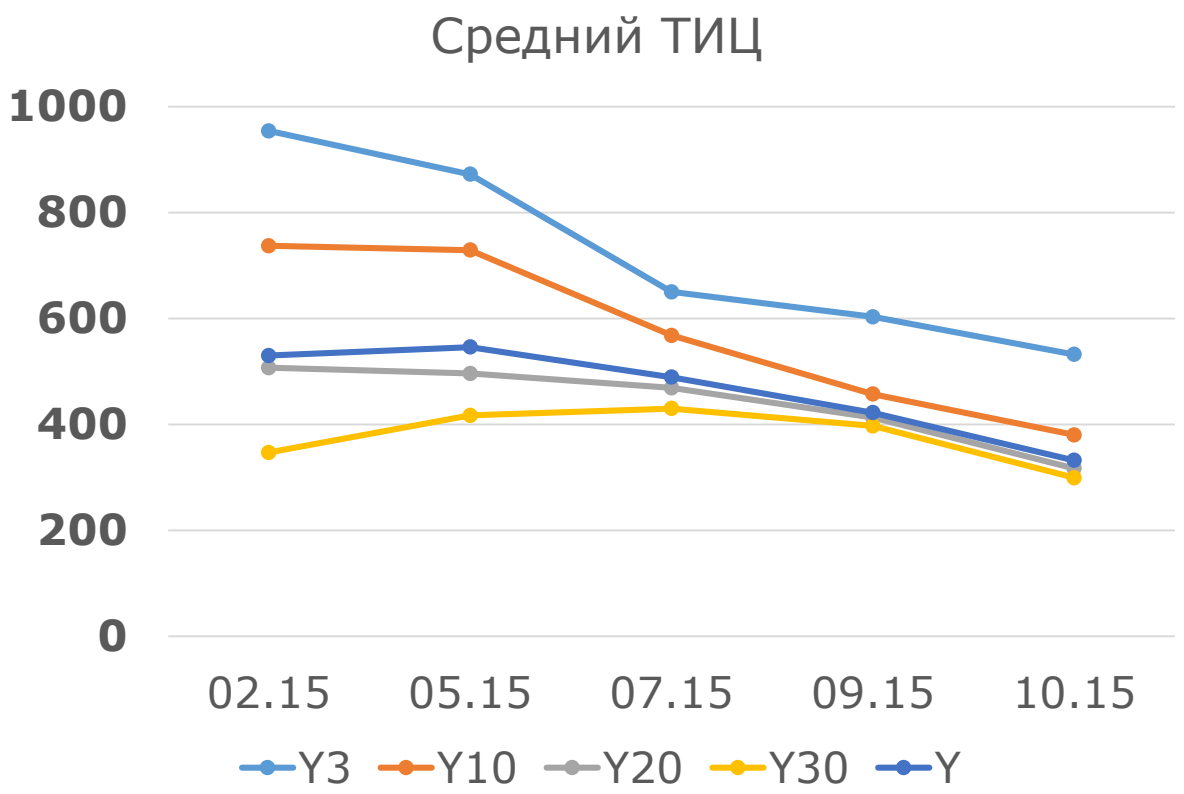
4. октябрь

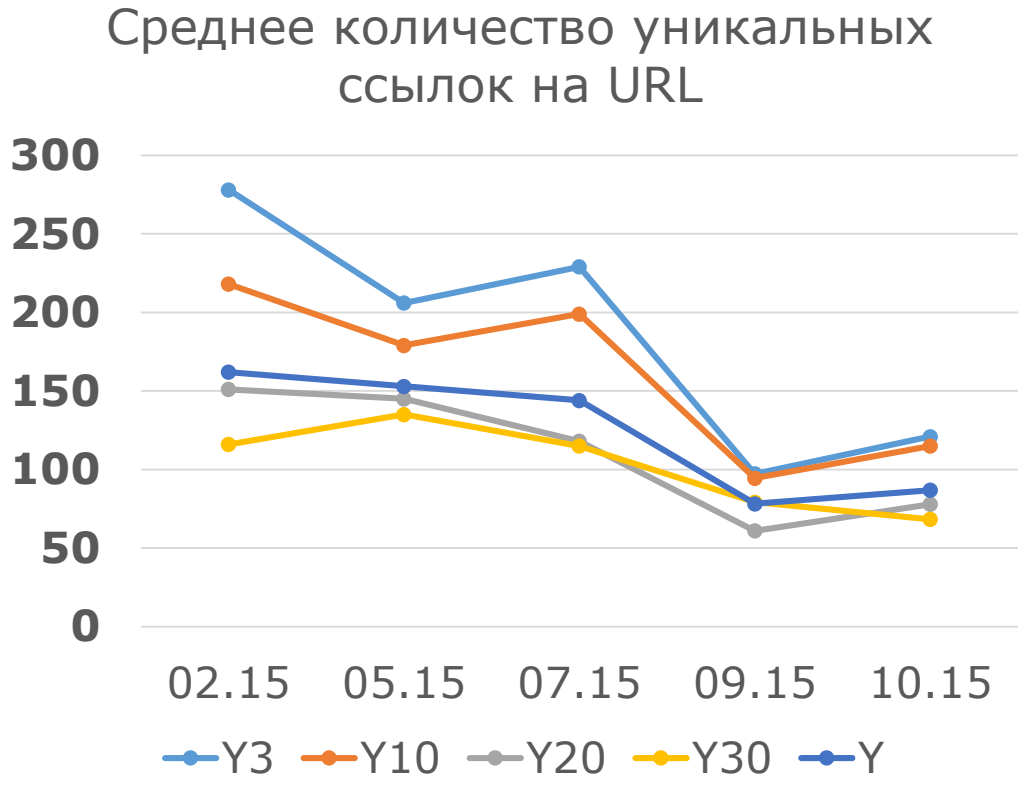


«МИНУСИНСК» РАБОТАЕТ!



НО И ССЫЛКИ ТОЖЕ...





ПРАКТИЧЕСКАЯ АНАЛИТИКА

Новый фокус проекта

- практические рекомендации
- с учетом конкретных «раскладов»
  - по выборке
  - по запросу
  - у конкретных конкурентов

Автоматическое построение отчетов

- критические и просто важные проблемы
- коммерческие факторы (по группам)
- текстовые факторы
- ссылочные факторы:
  общие, текстовые, риск «Минусинска»
- подробно по запросам

ОТЧЕТ: САМОЕ ГЛАВНОЕ

Самое главное

	Сайт Клиента	Топ 10	Топ 30	Конкурент 1	Конкурент 2	Конкурент 3
Вхождение слов запроса в URL	0	0.52	0.53	0.72	0.62	0
Вхождения в <title>: кворум	0	1.44	1.30	1.46	0.82	1.86
Вхождения в текст: кворум	2.13	25.5	23.5	28	23.4	8.29
Вхождения в текст: отдельные слова	4.85	32.2	31.2	28.9	25.9	10.4
Основная локализация в Москве	Нет	99%	99%	Есть	Есть	Есть
Онлайн-консультант	Нет	61%	55%	Нет	Есть	Есть

А вот другой сайт с другими главными проблемами:

Вхождение слов запроса в URL	0.33	0.58	0.61	0.32	0.92	0.32
Риск «Минусинска»	3.06	0.30	0.38	0.15	0.13	0.15
Наличие счетчика LI	Нет	76%	78%	Есть	Есть	Нет
Цена (на странице)	33%	94%	92%	Есть	Есть	Есть

ОТЧЕТ: ГРУППА ФАКТОРОВ

Продукт и его представление

Важно	Сайт Клиента	Топ 10	Топ 30	Конкурент 1	Конкурент 2	Конкурент 3	Конкурент 4
Цена (на странице)	7%	86%	80%	Есть	82%	Есть	Есть
Подбор по параметрам	Нет	85%	76%	Есть	Есть	Есть	Есть
Информация о доставке	Нет	92%	87%	Есть	Есть	Есть	Есть
Почта или транспортные компании	Нет	81%	76%	Есть	Есть	Есть	Есть

Полезно	Сайт Клиента	Топ 10	Топ 30	Конкурент 1	Конкурент 2	Конкурент 3	Конкурент 4
Несколько фотографий	Нет	48%	46%	Есть	Нет	Есть	Нет
Оплата	Нет	75%	67%	Есть	Есть	Есть	Есть
Оплата Яндекс.Деньгами	Нет	19%	14%	Есть	Нет	Нет	Нет
Курьер	Нет	73%	66%	Есть	Есть	Есть	Есть
Самовывоз	Нет	58%	52%	Есть	Есть	Есть	Нет

Сильные стороны	Сайт Клиента	Топ 10	Топ 30	Конкурент 1	Конкурент 2	Конкурент 3	Конкурент 4
Цена (на сайте)	Есть	97%	95%	Есть	Есть	Есть	Есть
Фотографии	Есть	97%	97%	Есть	Есть	Есть	Есть
Описание	Есть	87%	88%	Есть	Есть	Есть	Есть
Сравнение товаров	Есть	54%	47%	Есть	Есть	Есть	Есть

ОТЧЕТ: ПРОБЛЕМНЫЙ ЗАПРОС

Особенно проблемные запросы

аренда генератора

[Кратко|Подробно]

Ожидаемая страница: Страница клиента

Сайт ***не попал*** в результаты

Характеристики запроса

Веса: *аренда:0.38 генератора:0.62* (кворум: *0.99*)

Добавленные слова: *arenda, generatora, аренда, генератора, москве*

Тип сайта	Топ 10	Топ 30
Продавец / Поставщик услуг	9	25
Агрегатор	1	3
Доска объявлений	—	1
Прочее (игры, спам, пустые, ...)	—	1

Тип страницы	Топ 10	Топ 30
Главная	7	9
Главная в разделе	2	17
Внутренняя	1	2

	Топ 10	Топ 30
Количество элементов в URL	1.50	2.10

Текстовые факторы

Критично	Сайт Клиента	Топ 10	Топ 30	Конкурент 1	Конкурент 2	Конкурент 3	Конкурент 4
Вхождение слов запроса в URL	0.00	0.30	0.62	1.00	0.00	1.00	0.00
Вхождения в текст: кворум	0	18.3	17.9	6	8	44	14
Вхождения в текст: отдельные слова	7.62	30.7	28.2	11.9	23.5	70.9	17.9
Вхождения в <title>: кворум	0	1.30	1.50	2	1	1	1

Полезно	Сайт Клиента	Топ 10	Топ 30	Конкурент 1	Конкурент 2	Конкурент 3	Конкурент 4
Вхождение слов запроса и синонимов в URL	0.00	0.37	0.70	1.00	0.00	1.00	0.00
Вхождения в текст: точный запрос	0	3.90	3.54	2	0	8	5

ОТЧЕТ: СЛОВА ЗАПРОСА

Количество вхождений слов запроса в текст страницы							
- такси	6	19.6	17.6	20	14	21	6
- программу	6	12.3	12.3	28	8	11	8
- купить	4	2.1	1.7	7	0	1	4
Количество вхождений синонимов в текст страницы							
- программное	0	1.7	2.2	1	2	2	0
- таксопарков	0	0.8	1.1	0	0	1	0
- софт	0	0.2	0.2	0	0	2	0
- software	0	0.0	0.1	0	0	0	0
- soft	0	0.0	0.0	0	0	0	0
Количество вхождений дополнений запроса в текст страницы							
- таксистов	0	0.6	0.7	1	0	0	0
- taxi	0	0.4	1.6	3	0	0	1
- москва	1	0.2	1.2	0	1	0	0
- покупка	0	0.2	0.4	1	0	0	1
- покупать	0	0.0	0.1	0	0	0	0
- приобрести	0	0.0	0.2	0	0	0	0
Основные термины							
- заказа	5	4.00	10.48	1	4	14	1
- водителями	2	3.67	6.08	10	7	11	0

СПАСИБО ЗА ВНИМАНИЕ!

Михаил Волович

руководитель лингвистического отдела

mv@ashmanov.com

Сергей Вальковский

ведущий аналитик