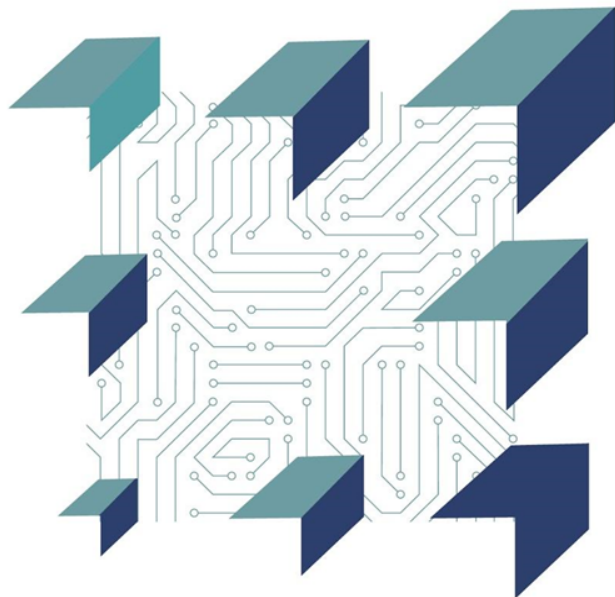


Латиница в госзакупках как пример “слепых закупок”

Способы искажения информации о государственных и муниципальных закупках в целях ограничения конкуренции и находимости информации. Как выявлять такие искажения

Иван Бегтин



Закупки в латинице — существенное продолжение



ivbeg | 4th Июн 2009 | E-Government | 1 Comment

Вокруг темы возник существенный ажиотаж у меня только сегодня уже несколько изданий запросили комментарии.

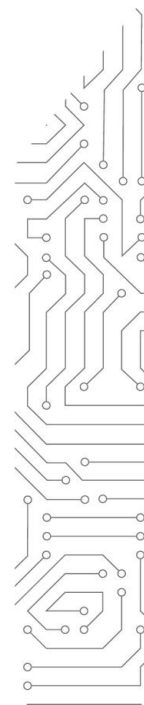
Должен сказать что подобного внимания к теме я не ожидал, тем более что это был просто самый заметный пример — причём изначально найденный не мной, а я лично увидел его на одном из форумов посвящённым госзакупкам. Его я привёл только для наглядности, самое же интересное можете посмотреть аттачем к этому посту.

Здесь приведены закупки где выявлены чередования русских и латинских букв в словах. Алгоритм не идеально точен, поэтому иногда попадают слова где английские буквы чередуются со спецсимволами, но их не так уж много.

Структура файлов:

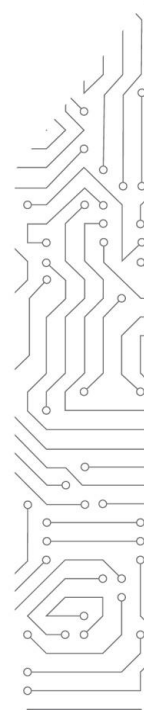
- слово
- название закупки
- ссылка на zakupki.gov.ru

В ZIP архиве HTML файл где русские и английские буквы подсвечены — русские зеленым,



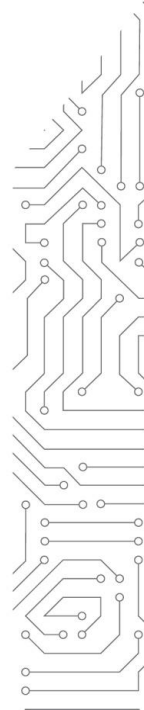
Откуда всё это возникло

1. Единый портал zakupki.gov.ru дал удобные инструменты поиска заказов для поставщиков
2. На сайте отсутствовал контроль качества информации (форматно-логический контроль)
3. Госзаказчики использовали трюки с описанием предметов закупки чтобы ограничивать возможности поставщиков найти информацию



Способы искажения информации

- **“латиница”**: замена кириллических букв на похожие по написанию латинские;
- **опечатки и ошибки**
- **отсутствие ключевых слов**
- **замена букв на похожие цифры**
- **использование тире в корне слова**
- **разделение букв пробелами**

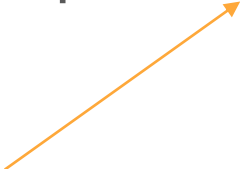


“Латиница”

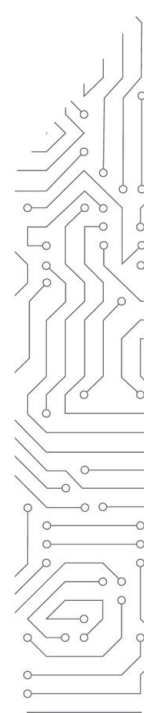

«латиница»—замена кириллических букв на аналогичные по написанию латинские.

Например, «окон» и «окон»

Кириллическая “о”, код **U+043E**



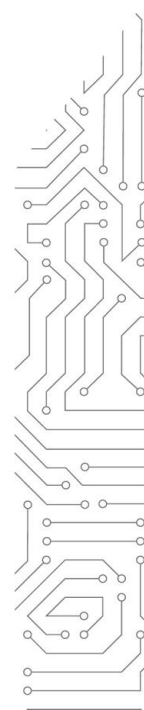
Латинская “o”, код **U+006F**



Опечатки и грамматические ошибки

Опечатки и грамматические ошибки – использование опечаток, а также грамматических ошибок, в том числе, в ключевых словах, определяющих предмет закупки.

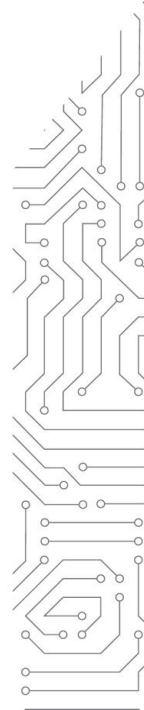
Например, «**поставка мяса (гавядина)**» или «**поставку автомобиля**»;



Отсутствие ключевых слов

Отсутствие ключевых слов – в наименовании закупки используются только общие слова, не позволяющие однозначно определить предмет закупки.

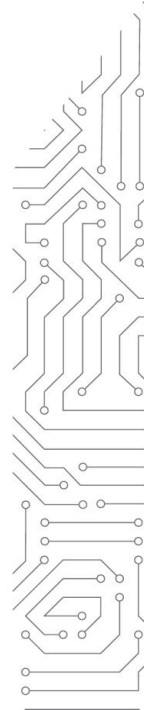
Например, «**Лот N15**» или «**закупка товаров**»;



Замена букв на похожие по написанию цифры

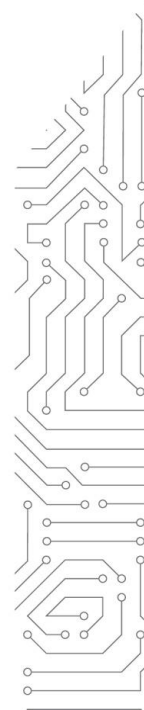
Замена букв на похожие по написанию цифры – в этом случае буквы «о» и «б» могут заменяться на цифры «0»(ноль) и «6»(шесть).

Например, «**поставка мебели**»



Использование тире в корне слова

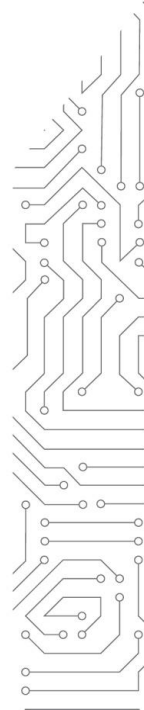
Использование тире в корне слова – в этом случае знак тире размещается в корне слова, как при переносе текста по правилам русского языка. Например, «**поставка мо-лочных про-дуктов**»;



Разделение букв в составе слова пробелами или иными знаками

Разделение букв в составе слова пробелами или иными знаками –в этом случае каждая буква или часть слова отделена от другой пробелом.

Например, «**работ по из г о т о в л е н и ю п р о д у к ц и и (п о л и г р а ф и ч е с к а я)**»

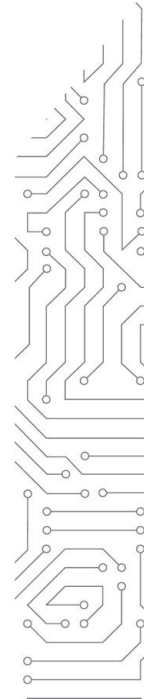


Как это было в 2012 году

Опублик: Опубликовано 31.07.2012 08:56 (MSK (UTC+4) Москва, Волгоград) (по местному времени заказчика)

Общая	Общая информация	Документы заказа	Изменения и разъяснения	Протоколы	Журнал событий
-------	------------------	------------------	-------------------------	-----------	----------------

Общая	Общая информация о заказе	
Способ	Способ размещения заказа	Открытый конкурс
Размещ	Размещение заказа осуществляет	Уполномоченный орган ИНН 5321026264 КПП 532101001 комитет по управлению государственным имуществом области
Заказчи	Заказчик	Сведения приведены ниже
Наимен	Наименование заказа	оказание услуг по распространению информации о деятельности Новгородской областной Думы, в т.ч. о законодательной работе и взаимодействии с населением в печатных изданиях, распространяемых на территориях Батецкого, Волотовского, Демянского, Любытинского, Маловишерского, Маревского, Мошенского, Новгородского, Окуловского, Парфинского, Пестовского, Поддорского, Солецкого, Хвойнинского, Холмского, Чудовского, Шимского районов Новгородской области для государственных нужд Новгородской области в 2012 году.
Наимен	Наименование лота	оказание услуг по распространению информации о деятельности Новгородской областной Думы, в т.ч. о законодательной работе и взаимодействии с населением в печатных изданиях, распространяемых на территориях Батецкого, Волотовского, Демянского, Любытинского, Маловишерского, Маревского, Мошенского, Новгородского, Окуловского, Парфинского, Пестовского, Поддорского, Солецкого, Хвойнинского, Холмского, Чудовского, Шимского районов Новгородской области для государственных нужд Новгородской области в 2012 году.
Началь контра	Начальная (Максимальная) цена контракта	269 100,00 Российский рубль
Этап ра	Этап размещения заказа	Размещение заказа завершено



Каждое слово красным шрифтом с латиницей

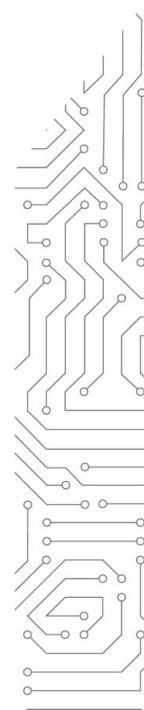
 Печатная форма  RSS-подписка на события заказа

Опубликовано 30.07.2012 12:19 (MSK (UTC+4) Москва, Волгоград) (по местному времени заказчика)

Общая информация | Документы заказа | Изменения и разъяснения | Протоколы | Жалобы/Результаты контроля | Информация о контракте | Журнал событий

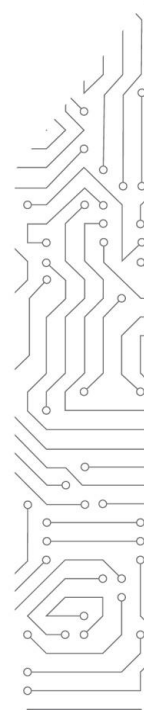
Общая информация о заказе

Способ размещения заказа	Открытый аукцион в электронной форме
Аукцион в электронной форме проводится на электронной площадке по следующему адресу	http://roseltorg.ru
Размещение заказа осуществляет	Заказчик
Заказчик	ИНН 3718000813 КПП 371801001 Администрация Пестяковского муниципального района Ивановской области
Краткое наименование аукциона	Право на заключение муниципального контракта на поставку материальных ресурсов для подключения многоквартирного дома № 29 по ул. Гагарина и здания Дома детского творчества к котельной МУЗ «Пестяковская ЦРБ» и частичной замены действующего участка теплотрассы на больший диаметр
Полное наименование аукциона (предмет контракта)	Право на заключение муниципального контракта на поставку материальных ресурсов для подключения многоквартирного дома № 29 по ул. Гагарина и здания Дома детского творчества к котельной МУЗ «Пестяковская ЦРБ» и частичной замены действующего участка теплотрассы на больший диаметр
Начальная (Максимальная) цена контракта	1 286 900,00 Российский рубль
Классификация товаров, работ и услуг	- 4560252 Наружные тепловые сети
Количество поставляемого товара, объема выполняемых работ, оказываемых услуг	В соответствии со спецификацией (пункт 1 раздела III аукционной документации, приложение 1 к проекту муниципального контракта)



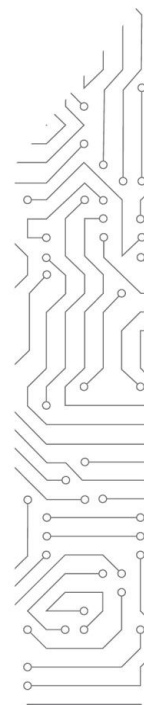
Как проводилось исследование

- Разработана методика выявления закупок с искажениями названий
- Собрана база всех искажений за 2012 год
- Вручную проведена проверка на намеренность искажений в названиях
- Разработана методика оценки ограничения доступа к участию в государственных закупках
- Оценен ущерб от ограничений к участию в государственных закупках



В цифрах за 2012 год

- **несколько миллионов** - извещений о закупках
- **около 13 тысяч** закупок выявлено с искажениями
 - **662** закупки с признаками умышленного сокрытия информации
 - **около 2500** закупок требующих внимания
 - **около 9000** закупок с единичными нарушениями





РУКОВОДИТЕЛЬ
ЦИФРОВОЙ
ТРАНСФОРМАЦИИ

А что сейчас?



Опечатки

Закупка у единственного поставщика
(исполнителя, подрядчика)

 Процедура завершена / 223-ФЗ

Начальная цена

906 876 ,25

Российский рубль

№ 31300563078

Заказчик:

МОСКОВСКОЕ ОТКРЫТОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО ТРУБОПРОВОДНОГО ТРАНСПОРТА
НЕФТЕПРОДУКТОВ.

Купля-продажа серверного **оборудования**

Размещено: 20.09.2013

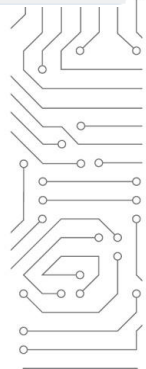
Обновлено: 20.09.2013



[Сведения](#) / [Документы](#) / [Жалобы](#)

Пример

<http://zakupki.gov.ru/223/purchase/public/purchase/info/common-info.html?regNumber=31300563078>



Отсутствие смысловых слов

Закупка у единственного поставщика
(подрядчика, исполнителя) (до
01.07.18)

✔ Процедура завершена / 223-ФЗ

Начальная цена

237 120 ,00

Российский рубль

№ 31908212779

Заказчик:

МУНИЦИПАЛЬНОЕ АВТОНОМНОЕ ОБЩЕОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ СРЕДНЯЯ
ОБЩЕОБРАЗОВАТЕЛЬНАЯ ШКОЛА № 13 ИМЕНИ П. А. ЛЕОНОВА ГОРОДА ЮЖНО-САХАЛИНСКА

Поставка товаров



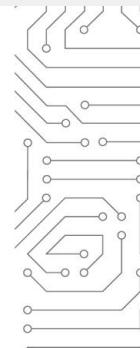
Размещено: 19.08.2019

Обновлено: 19.08.2019

[Сведения](#) / [Документы](#) / [Жалобы](#) / [Протоколы](#)

Пример

<http://zakupki.gov.ru/223/purchase/public/purchase/info/common-info.html?regNumber=31908212779>



Технические аспекты

Открытый код

<https://code.ach.gov.ru/labs/blindprocurement>

Экспериментальные лабораторные проекты > Слепые закупки > Details

 **Слепые закупки** 
Project ID: 11




☆ Star 0 Clone ▾

 MIT License  1 Commit  1 Branch  0 Tags  0 Bytes Files

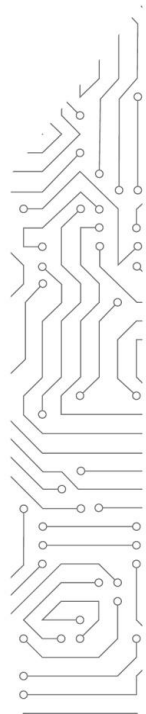
Код исследования "Слепые закупки" по выявлению закупках с подменой символов (латиница, лишние пробелы и тд) для снижения находимости для поставщиков.

Оригинальный репозиторий: <https://github.com/infoculture/blindprocurement> Требуется ревизия

master ▾ blindprocurement History 🔍 Find file 📄 ▾

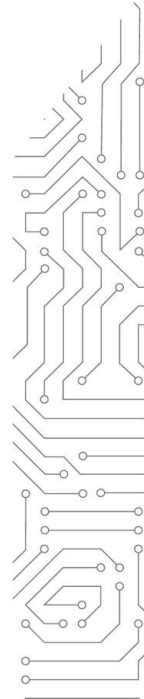
 Первоначальный импорт
Иван Бегтин authored 2 months ago  7dc41cfd 

 README



Поиски аномалий в наименованиях

- нет никаких смысловых слов
- есть слишком длинные слова, более 20 символов
- слова разных кодовых таблиц вперемешку: кириллица и латиница
- спецсимволы и разделители внутри слов (тире, пробелы и тд)
- опечатки, их частота и локализация



Частотный словарь

Что

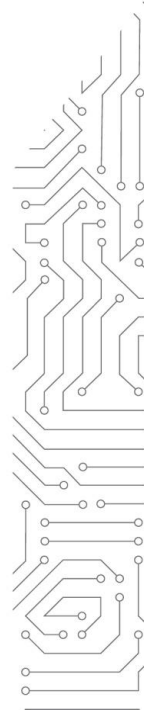
Словарь из всех слов использованных когда-либо в наименовании закупок. На выходе словарь из 749544 слов

Как

Обработка наименований более 2 миллионов закупок

Зачем

Для последующего выявления аномалий



Словарь опечаток

Что

Словарь всех встречающихся опечаток в базе

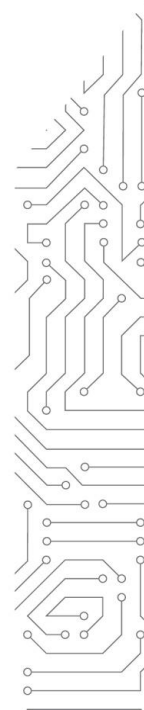
Как

Автоматический расчёт расстояния Левенштейна для всех слов и отсева редко встречающихся слов как опечаток

https://ru.wikipedia.org/wiki/Расстояние_Левенштейна

Зачем

Для идентификации опечаток



Список вспомогательных слов

Что

Словарь из 150 вспомогательных слов

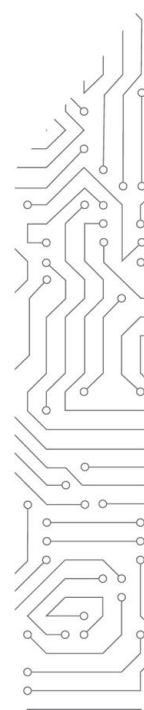
Пример: поставка, закупка, аукцион, ГБУЗ, лот и тд.

Как

Ручная выборка из 5000 наиболее часто употребляемых слов

Зачем

Для выявления закупок где нет ни одного смыслового слова





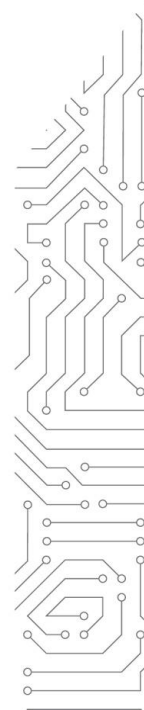
РУКОВОДИТЕЛЬ
ЦИФРОВОЙ
ТРАНСФОРМАЦИИ

Незавершённое



Планировалось, но незавершено

- Ежегодная проверка данных ЕИС на искажение
- Выявление склеенных слов
- Идентификация отсутствия содержания в описании предмета закупки (текст предмета закупки идентичен классификационному коду)
- сопоставление текста предмета закупки и классификационного кода: выявление несоответствий





Вопросы?

Иван Бегтин, ibegtin@infoculture.ru

Председатель Ассоциации участников рынка данных
Директор АНО “Информационная культура”

