

Яндекс

Яндекс

В окрестностях Палеха

Александр Садовский

OPTIMIZATION 2016, Москва

От слов — к смыслам

- › TF*IDF модели
- › Морфология
- › Расширения запроса
- › ...

синхрофазотрон купить
москва

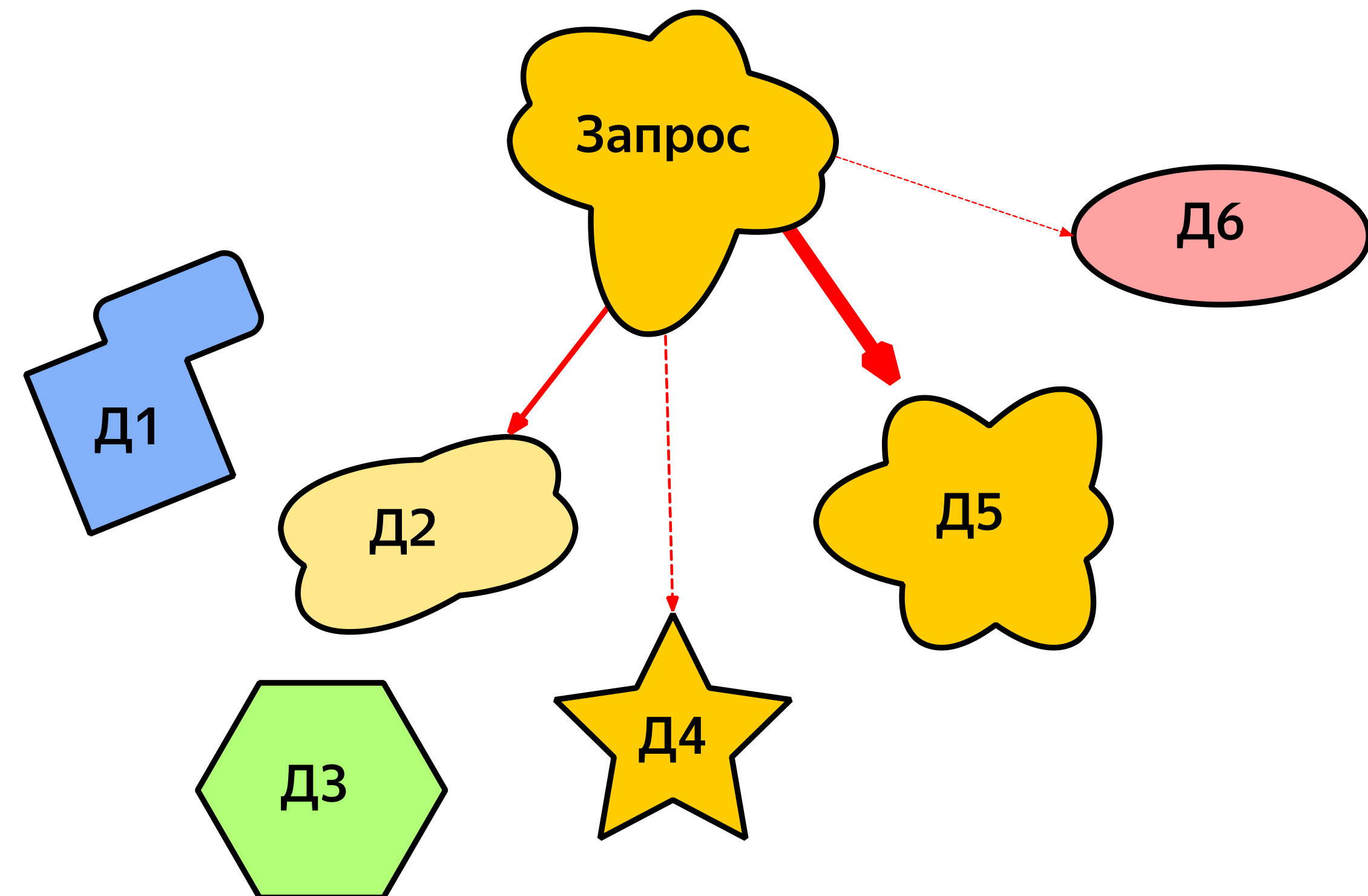
платье | платья | ...

насос | nasos | pump | ...

Но хочется искать по смыслу.

Как искать по смыслу?

1. Представим запрос и тексты векторами
(слов, n-грамм и словосочетаний)
2. Уменьшим размерность
3. Измерим сходство

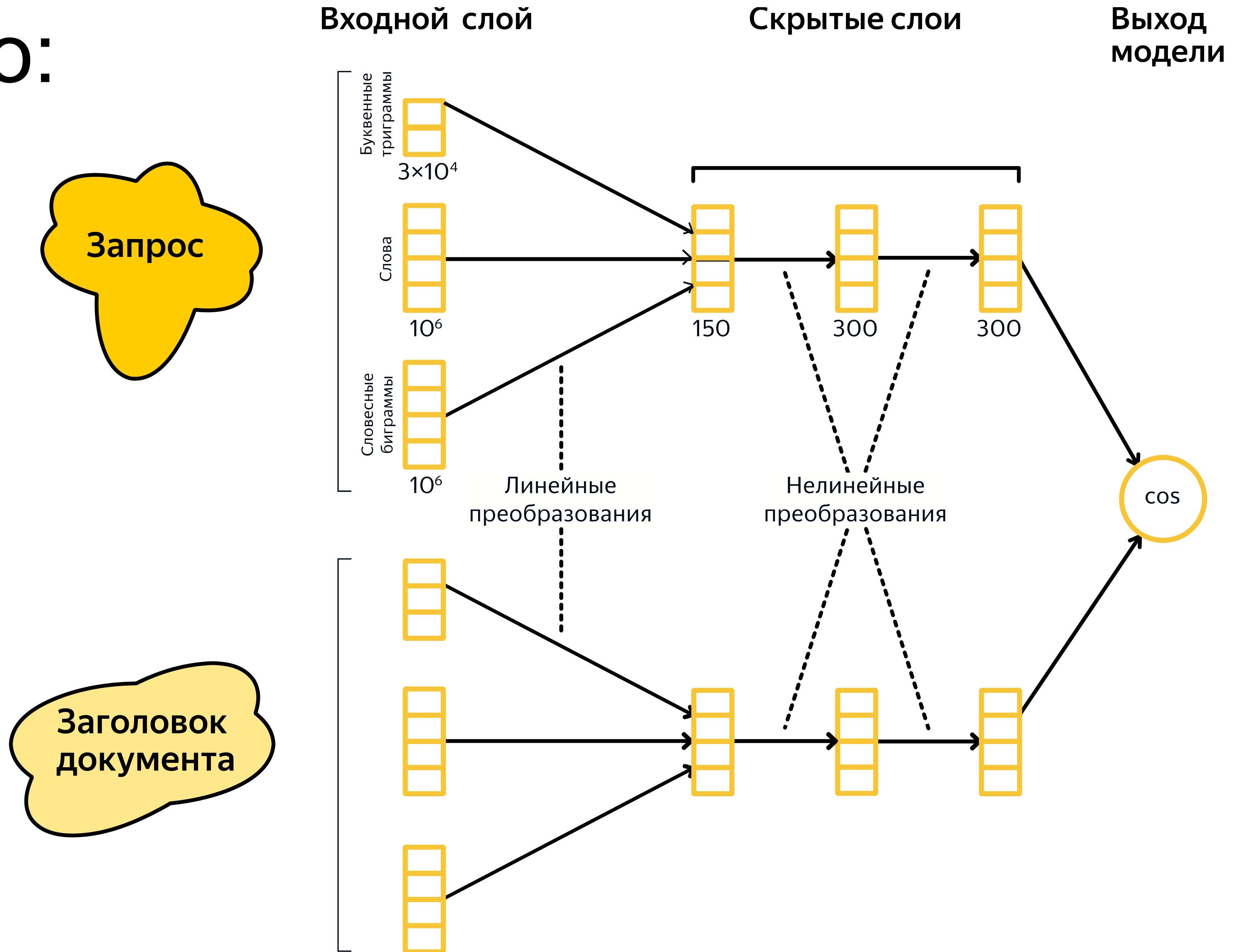


Известны реализации — Word2vec, DSSM, ...

Алгоритм «Палех»

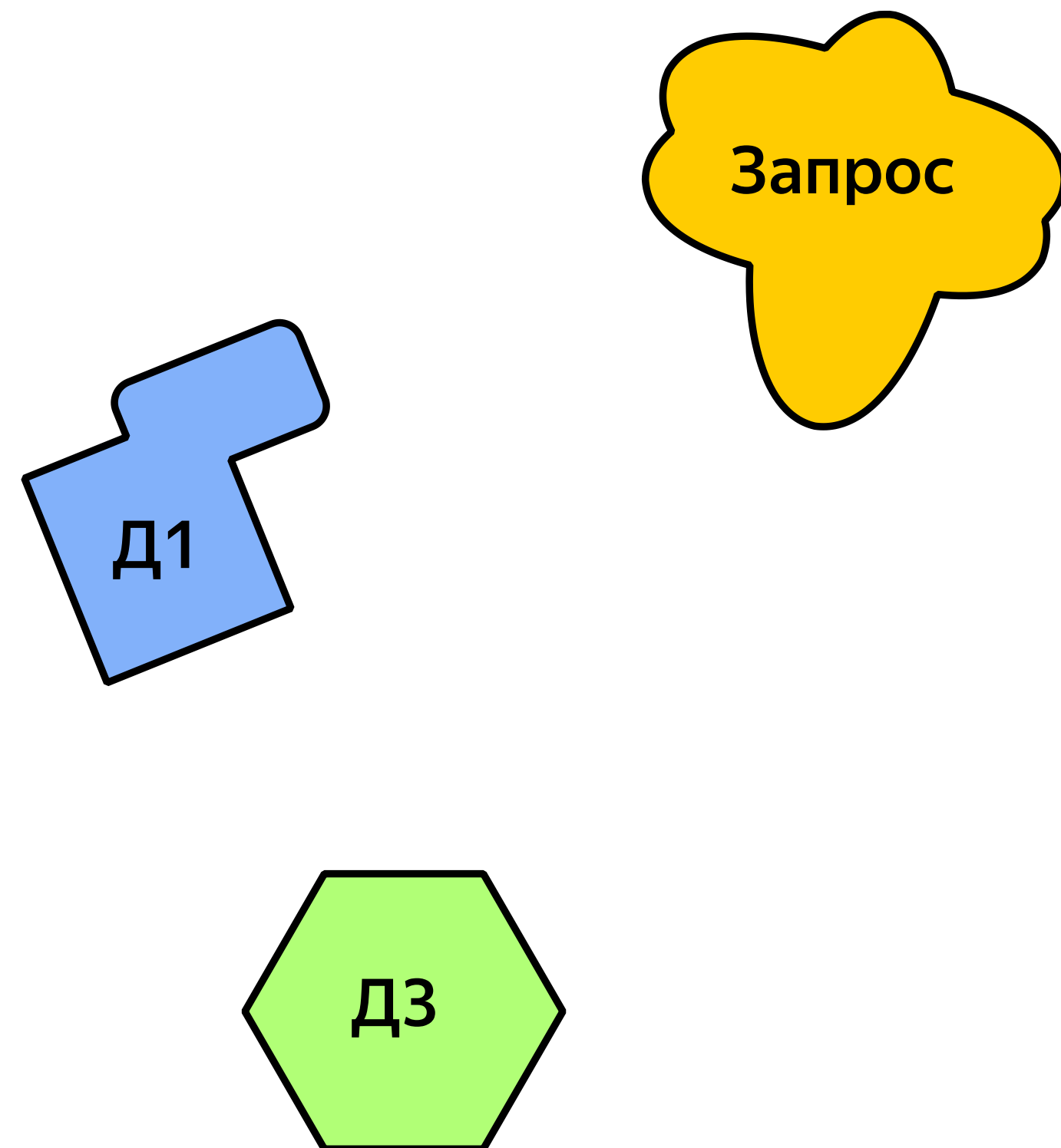


Устройство:

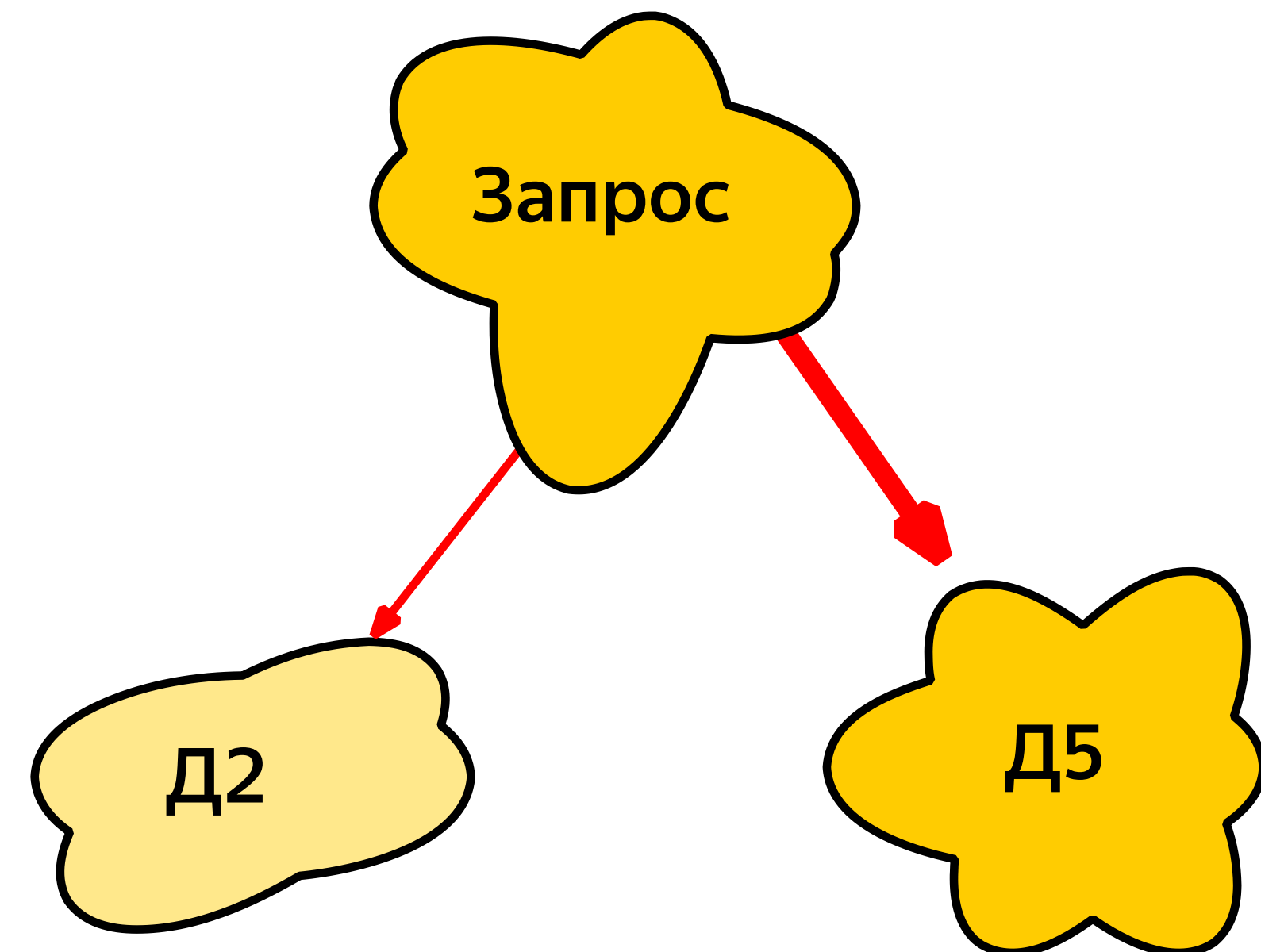


Нейросеть надо учить: нужны примеры!

Отрицательные



Положительные

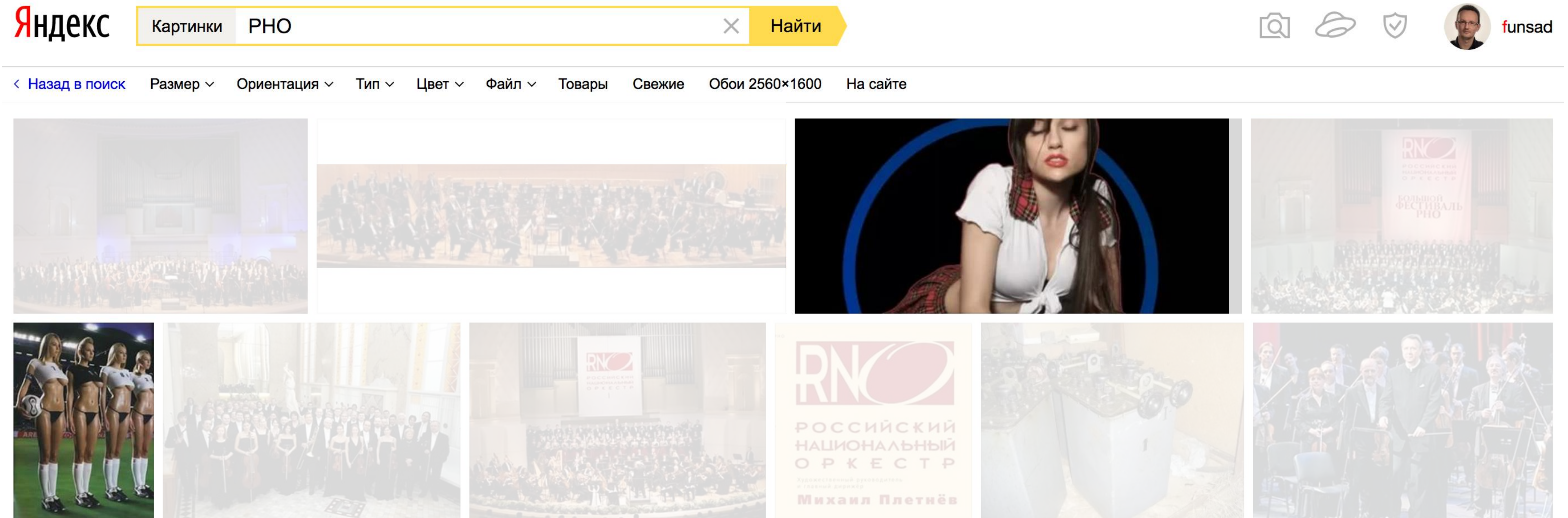


На чём можно учиться?

Наивный подход:

- положительные примеры — кликнутые документы,
- отрицательные — документы без клика

На что тут кликнут?



А пользователь искал расшифровку РНО — «работа над ошибками»

Кликнут ли тут?

Яндекс

сколько кг весила щука которую поймал путин



Найти

ПОИСК

КАРТИНКИ

ВИДЕО

КАРТЫ

МАРКЕТ

НОВОСТИ

ПЕРЕВОДЧИК

ЕЩЁ



Вес щуки Путина. Сколько весит щука...

[bolshoyvopros.ru](#) > ...**ves...putina-skolko...schuka...pojmal...** ▼

Сколько весит щука, которую поймал Путин? тэги: вес, масса, путин, рыба, щука. ...

Вес щуки Путина - 21 кг. Владимир Путин на молодежном форуме «Селигер-2013» подтвердил, что тувинская щука весила 21 кг.

Нашлось 18

[Дать объявл](#)

Мы научились находить
верные примеры

Положительные примеры

- › Те, где модель предсказывает, что пользователь задержится на сайте надолго или останется там навсегда
- › И др. :)

Отрицательные примеры

- › Случайный документ
- › Случайный документ с включением слов запроса
- › Hard negative mining

Результаты



Сравним нейронную сеть и bm25:

[фильм про дяденьку который родился старым]

«Загадочная история Бенджамина Баттона (2009)»

bm = 0.00, **нейронная сеть = 0.80**

[фильм про парня который играл на барабанах]

«Рецензия на фильм «Одержимость». Кино на Фильм.ру»

bm = 0.12, **нейронная сеть = 0.69**

| А в коммерческих запросах?

[сеточка на кран купить]:

Водосберегающие экономные насадки аэраторы

bm = 0.00, **нейронная сеть = 0.59**

Что такое аэратор для смесителя?

bm = 0.00, **нейронная сеть = 0.56**

[болгарка]:

Угловая шлифовальная машина — Википедия

bm = 0.00, **нейронная сеть = 0.81**

Углошлифовальные машины (УШМ) — купить по выгодной цене...

bm = 0.00, **нейронная сеть = 0.78**

[слайсер для домашнего использования недорого]:

■ Ломтерезки - купить в Москве в интернет магазине. Ломтерезки электрические для дома низкие цены, отзывы, большой выбор

■ $bm = 0.03$, **нейронная сеть = 0.79**

■ Ломтерезки Ritter Германия, купить для дома в Москве, электрические, выгодные цены, отзывы, видео, бесплатная доставка купить в Москве и Санкт-Петербурге

■ $bm = 0.03$, **нейронная сеть = 0.75**

[орбитреки интернет магазин]:

■ Эллиптические тренажеры - купить с доставкой, цены на эллиптические кардиотренажеры в интернет-магазине Спортмастер

■ $bm = 0.20$, нейронная сеть = 0.53

■ Орбитреки - Интернет-магазин спортивных товаров

■ $bm = 0.91$, нейронная сеть = 0.83

[штука на двух колесах на которой стоишь и едешь]:

Как называется транспортное средство на двух колесах с длинной ручкой?

bm = 0.19, **нейронная сеть = 0.79**

Электроскутер Segway (тест электро самоката Сегвей)

цены и модификации segway - Колеса.ру

bm = 0.17, **нейронная сеть = 0.76**

Fan-igra.ru. Как называется сигвей без руля? И где купить его? - Fan-igra.ru

bm = 0.00, **нейронная сеть = 0.44**

[громоотвод для дачного дома]:

Как сделать громоотвод в частном доме своими руками

bm = 0.64, нейронная сеть = 0.86

Молниезащита частного дома: устройство, монтаж

bm = 0.07, **нейронная сеть = 0.82**

[штука в машину для музыки]:

MP3 плеер вставляемый в прикуриватель автомобиля

bm = 0.00, **нейронная сеть = 0.67**

Музыка с телефона в машине 🎵 — бортжурнал Smart fortwo
Clever🍀 | DRIVE2

bm = 0.43, **нейронная сеть = 0.74**

Частые вопросы про Палех:

■ Типы запросов?

Все.

■ Языки и регионы?

Все.

■ Эффективность?

rFound1 +1.6% (на запросах длинного хвоста).

■ У меня упал вырос трафик! Это Палех?

Возможно.

■ Есть ли зоны, где Палех малоэффективен?

Это редкость, но есть – поиск цитат.

Палех думает о смысле.

Палех умный.

Будь как **Палех**!



Спасибо!

Александр Садовский

Поиск :: Руководитель продуктов

funsad@yandex-team.ru

