

фора



Сильно



# Message Tagger Overview

Details of Machine Learning Algorithm & Integration Steps

# Проект «Message Tagger» запущен для анализа отзывов в соц. сетях Фора и Сильпо (Justin in progress)

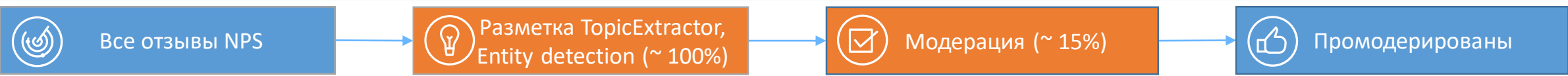
Цель и Бизнес задача проекта

Разработать систему с моделями ИИ для автоматической разметки отзывов (NPS) в социальных сетях (Сильпо, Фора, Justin). Автоматизированная передача размеченных сообщений бизнес группам холдинга.

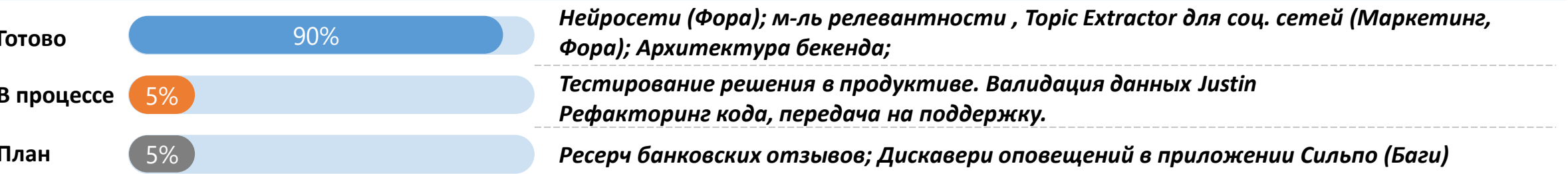
Было



Стало



Статусы

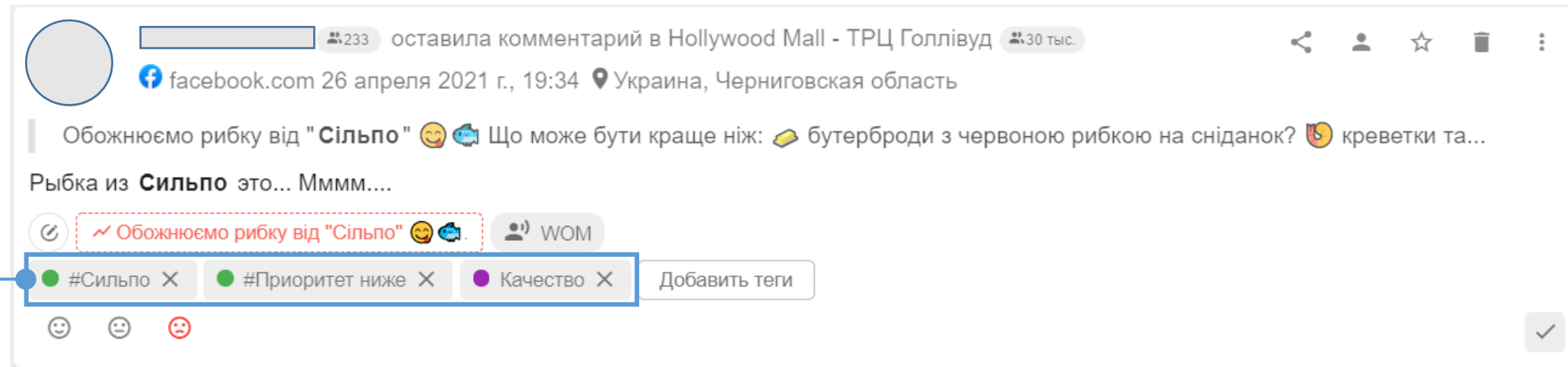


\*Подключение внутренних источников возможно после перехода на Hotline 2  
\*Открытый вопрос с Hotline 2 заморозил оптимизацию процесса коммуникации с филиалами и подключения внутренних источников Сильпо

# Проект «Message Tagger» примеры реализации

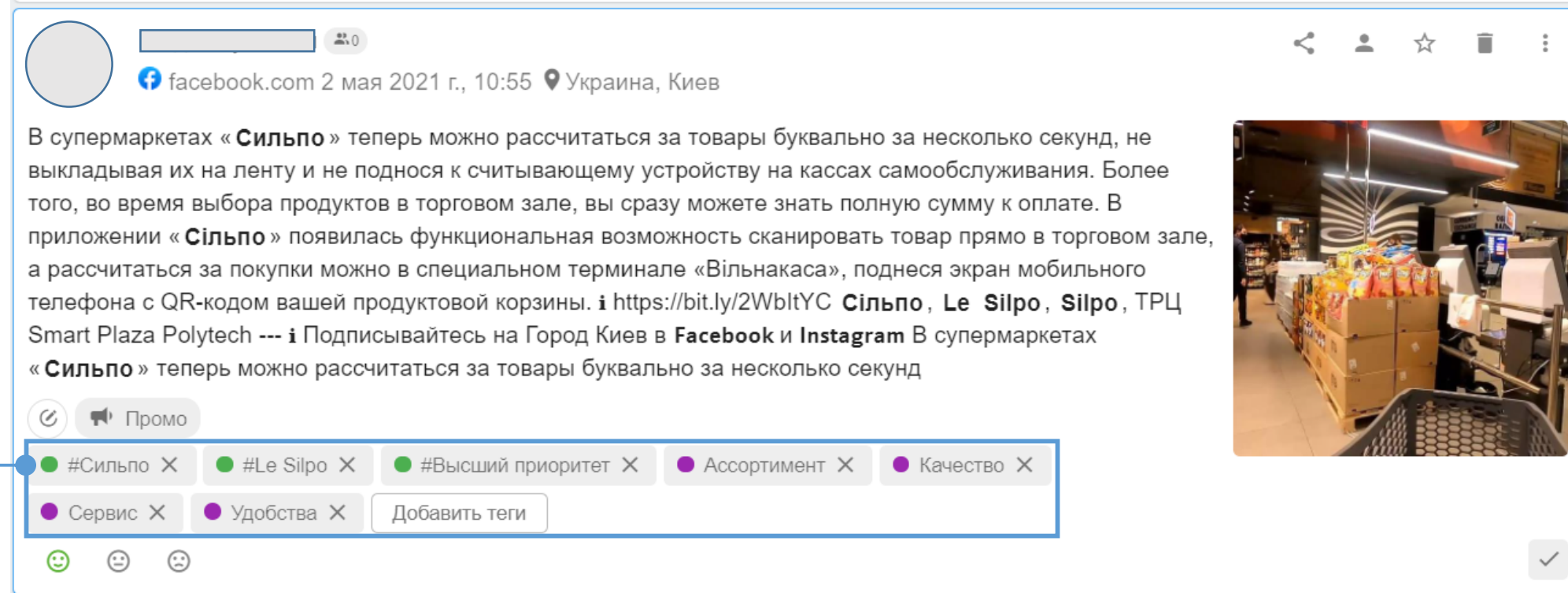
## Пример #1

Интегрировано  
в YouScan



## Пример #2

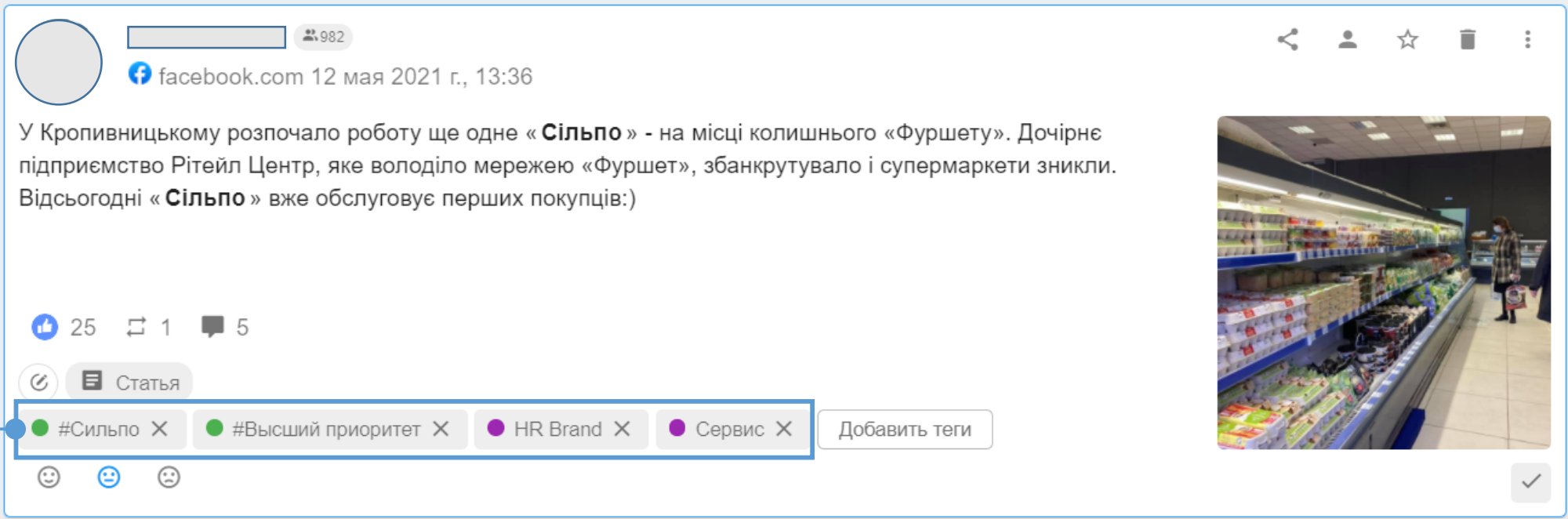
Интегрировано  
в YouScan



# Проект «Message Tagger» примеры реализации

## Пример #3

Интегрировано  
в YouScan



## Пример #4

Один из лучших магазинов. Всегда свежие овощи и фрукты, до безумия вкусная выпечка, кулинария. Большой ассортимент. В связи с карантином огромное желание наличия увлажнителя для пальцев в магазинах, чтобы открывать пакеты

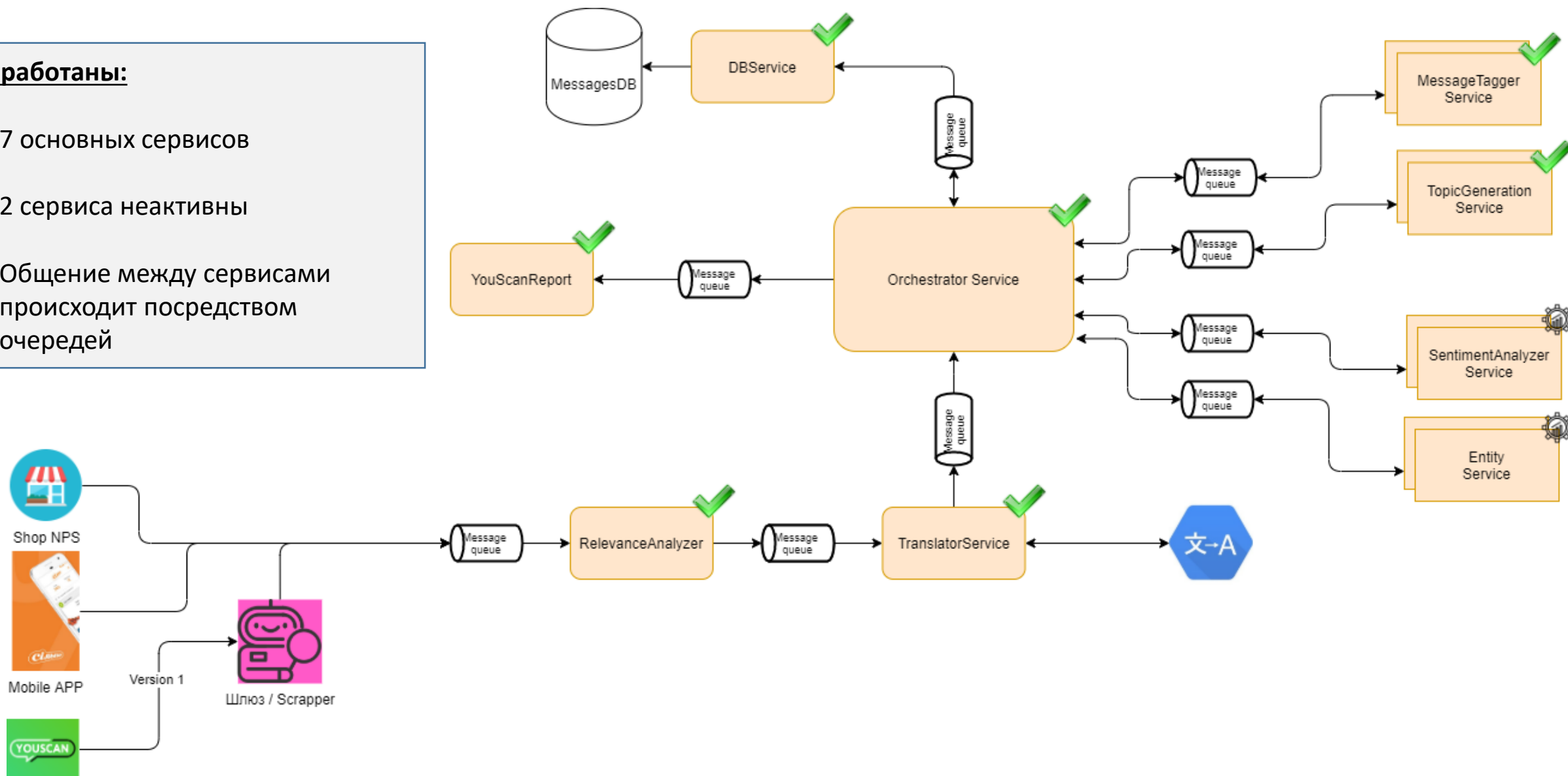
- "Качество" 56.58%
- "Сервис" 51.55%
- "Ассортимент" 49.16%
- "Карантин" 47.65%

Наглядная  
разметка  
TopicExtractor

# Основные реализованные сервисы архитектуры

## Разработаны:

- 7 основных сервисов
- 2 сервиса неактивны
- Общение между сервисами происходит посредством очередей



# Месседж Таггер покрывает все возможные аспекты текстовой информации с помощью 5 ключевых моделей ИИ

## Ключевые сервисы с моделями ИИ



### Relevance Analyzer

- Прескоринг до сервиса перевода
- Модель расчёта Релевантности



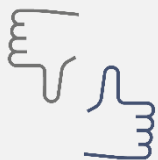
### Tagger Service

- Определение бизнеса (Тэг)
- Определение бизнес сущности



### Topic Extractor

- Распознавание тематики входящего упоминания (любого текста)



### Sentiment Analyzer

- Определение тональности упоминания



### Entity Recognition

- Выделение ключевых сущностей в упоминании

## Комментарий

**Модель Прескоринга удаляет самые нерелевантные сообщения на этапе попадания сообщений в систему. После перевода модель расчёта Релевантности группирует сообщения по приоритету обработки.**

**Модель определяет Тэги к которым относится упоминание.**  
Тэги формируются на основе бизнесов, бизнес сущностей, которые требуют отдельного рассмотрения. (например Сильпо, Фора, «Лавка Традиций»)




**Модель распознаёт о чём именно (тематика) написано в отзыве гостя.**  
Список тематик формируется на основе всех доступных данных (история)

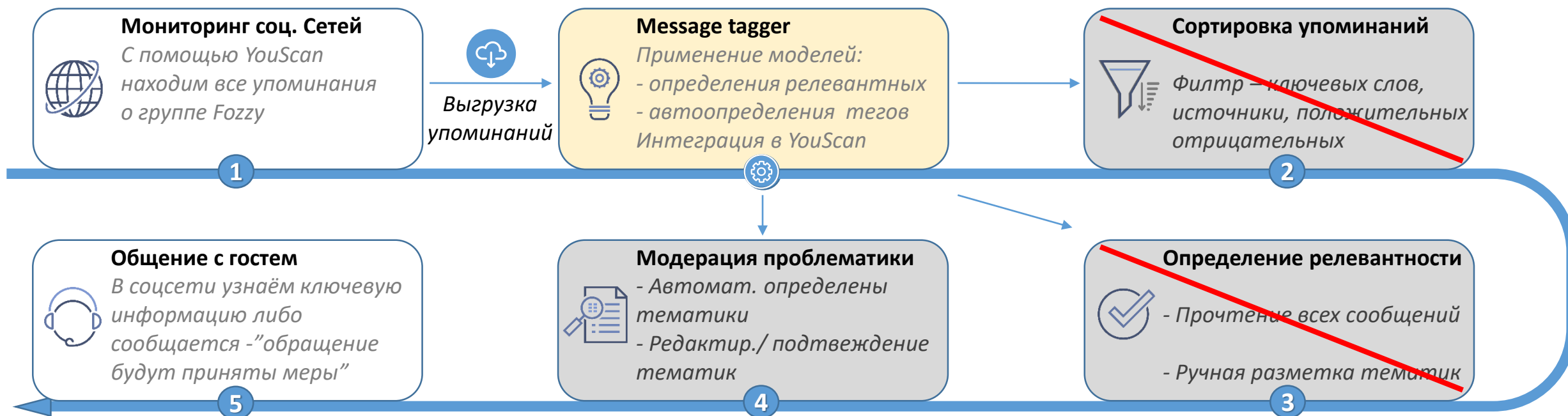
**Модель Тональности определяет как именно отзываются**  
(хорошо/плохо/нейтрально).  
**Неактивен в рамках проекта** (высокая точность тона входящих данных)

**Модель выделяет ключевые сущности, например Магазин по адресу..., время события, задействованные стороны (кассир)**  
**Неактивен в рамках проекта** (отсутстве интерфейса для внедрения)

# Реализованный процесс обработки сообщений с использованием интерфейса YouScan позволит ускорить обработку отзывов

## Процесс обработки сообщений с помощью YouScan после интеграции решения

-  Youscan – измененный процесс
-  Social networks – ресёрч интеграций
-  Tagger – внедряемый алгоритм



## Комментарий

- Менеджеры по работе с гостями смогут обработать **положительные и отрицательные** отзывы
- В реализованном процессе менеджеры только **подтверждают/редактируют** разметку «тематик»
- Message Tagger** автоматически выгружает и обрабатывает сообщения и сохраняет их (для дальнейшего улучшения моделей)
- Через созданный аккаунт **удаляет нерелевантные сообщения как пользователь**
- Расставляет **приоритеты релевантности** сгруппированных сообщений и **размечает тематики**



# Message Tagger покрывает 3 ключевых бизнес кейса

## Ключевые результаты



~75% ▼

### Сокращение человекочасов

На 65 – 80 %, или ~ 80 часов/неделю,  
или 2-3 дополнительных модератора

За счёт сокращения доли нерелевантных  
отзывов требуется меньше ресурсов для  
работы с отзывами.

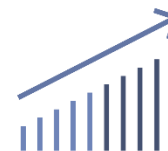


x2 ▲

### Увеличена скорость обработки

В 1.5 - 2 раза ускорена обработка  
(модерация) 1 упоминания

Модель генерации тематик позволяет  
быстрее определять острые отзывы, а  
также предварительно размечает -  
уменьшая мануальную разметку



x8 ▲

### Увеличен охват

Увеличен охват просматриваемых  
лояльностью сообщений в 6 – 8 раз.

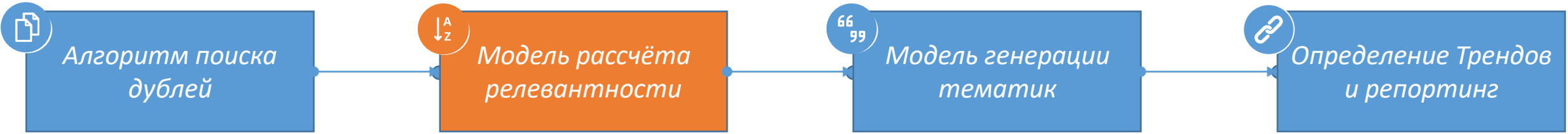
Часть модераторов успевает работать с  
положительными отзывами, также  
теперь просматривает **весь негатив**.



# Модель релевантности с прескорингом для Сильпо

## Реализованный алгоритм последовательных моделей

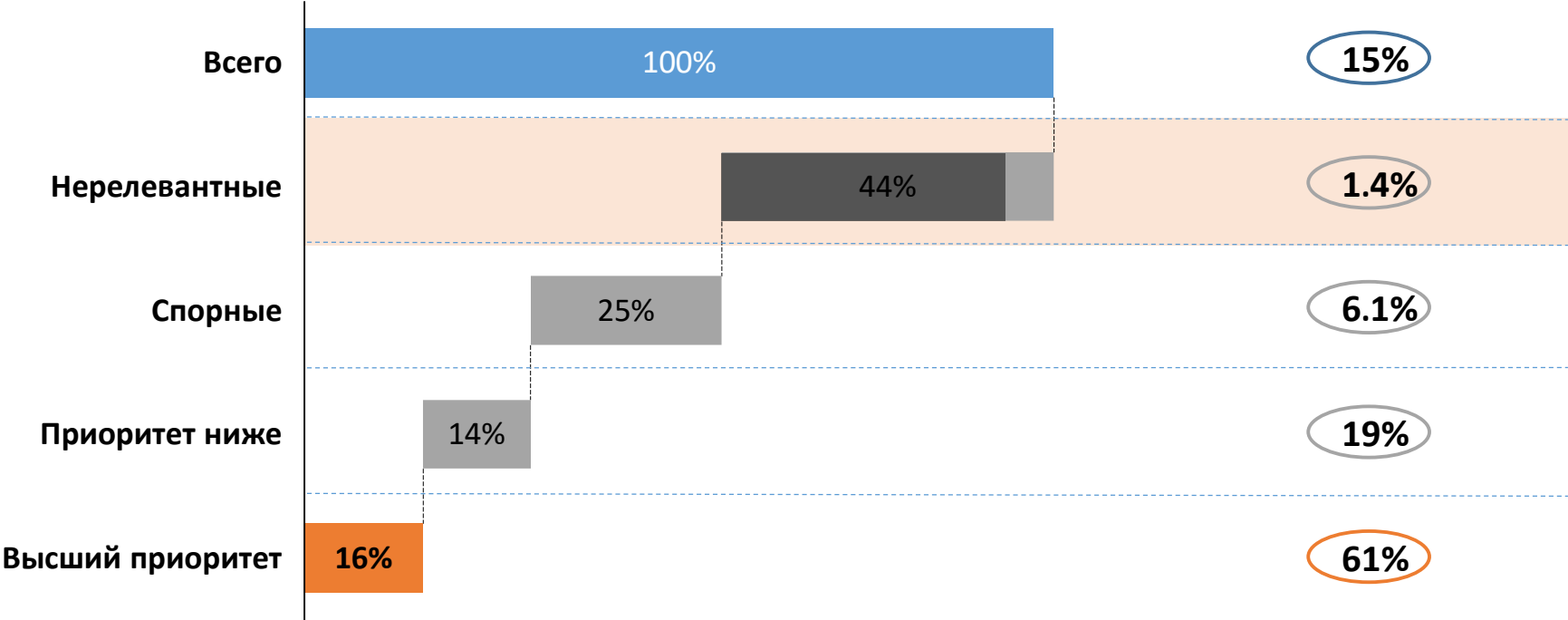
Детализировано ниже    Удаление до перевода



### Группы приоритетности на валидационных сообщениях, %

### Доля релевантных, %

### Комментарий



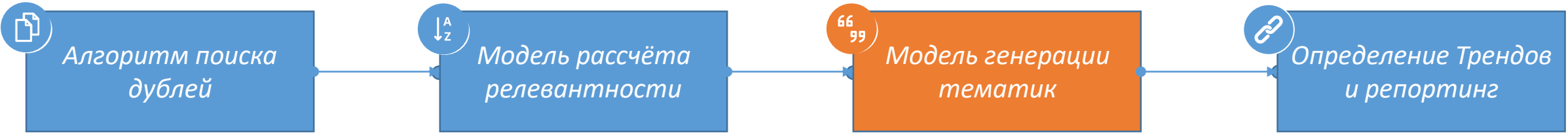
- В группе «Нерелевантных» 80% нерелевантных сообщений удаляются до перевода
- Перекалибровка даёт **основные 4 группы** (методом квантилей)
- Группа «Спорные» может быть рассмотрена к обработке с **наименьшим приоритетом** из-за низкой доли важных сообщений
- «Высший приоритет» релевантная группа первоочердной обработки

Источники: Данные размеченные менеджерами по работе с клиентами после запуска Message Tagger в пилот

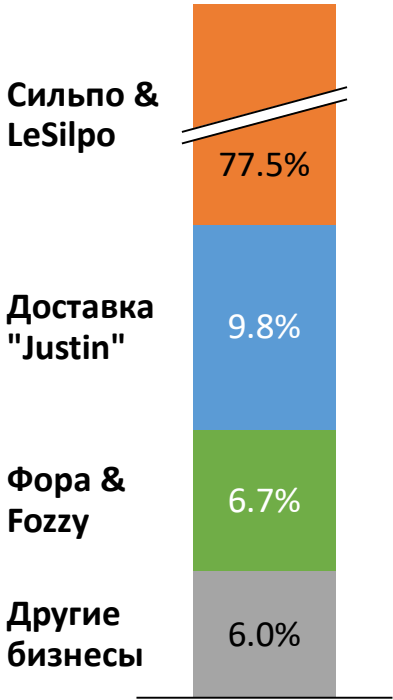
Система позволяют генерировать тематики автоматически, требуя лишь модерацию для дальнейшего улучшения точности Сильпо

Реализованный алгоритм последовательных моделей

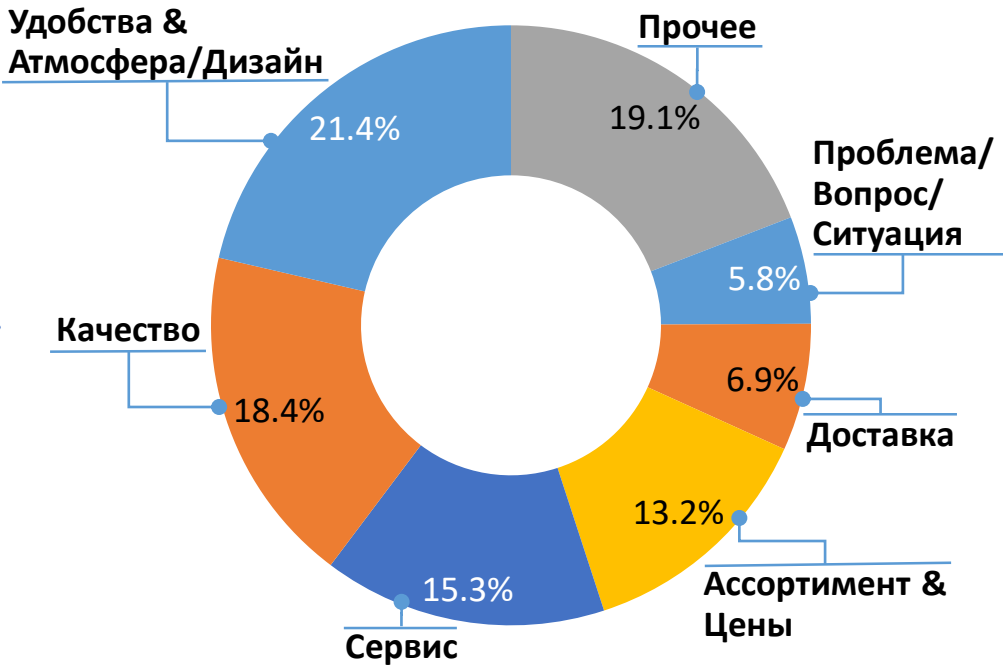
Детализировано ниже



Тегированные сообщения, %



Сгенерированные тематики сообщений в YouScan, %



Комментарий

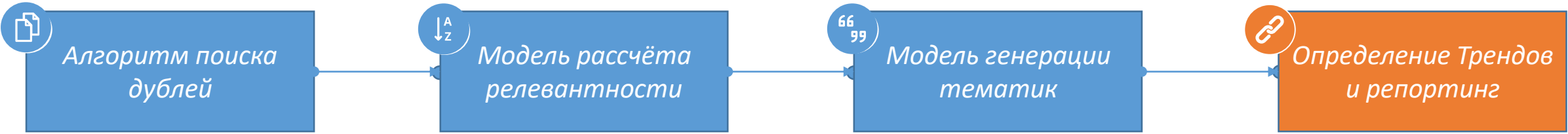
- В текущей версии модели настроены основные 17 тэгов, а также 17 тематик (топиков)
- Теги «Сильпо» и «ЛяСильпо» (как и «Фора» с «Fozzy») объединены для иллюстративных целей, самой моделью размечаются отдельно
- Темы «Удобства» объединена с «Атмосфера/Дизайн» (как и «Ассортимент» с «Цены») для иллюстративных целей, самой моделью размечаются отдельно

Источники: Данные размеченные моделью за промежуток недели после запуска Message Tagger в пилот

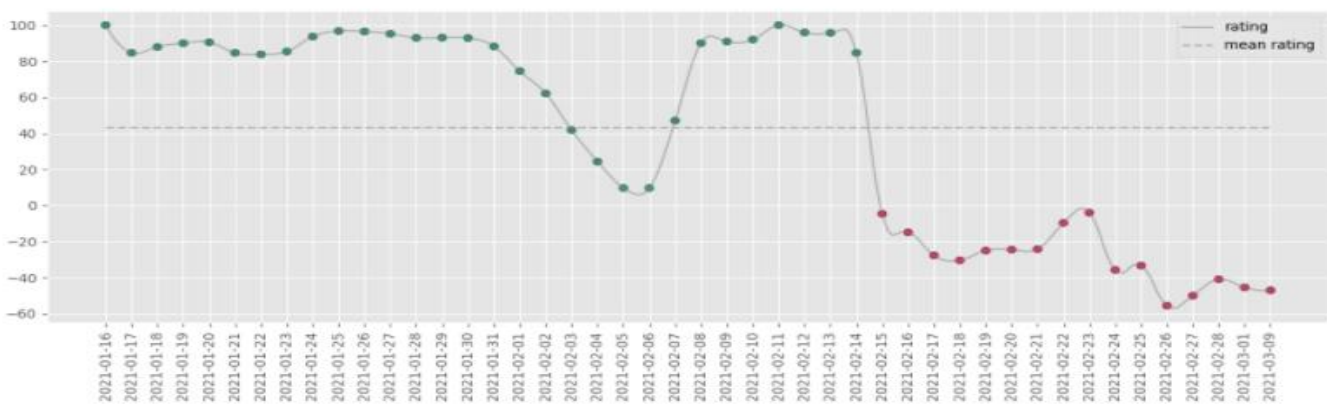
# Репортинге позволяет анализировать трендовые обсуждения и динамику Лояльности гостей по бизнесам и тематикам

## Реализованный алгоритм последовательных моделей

Детализировано ниже Удаление до перевода



Пример расчёта , %

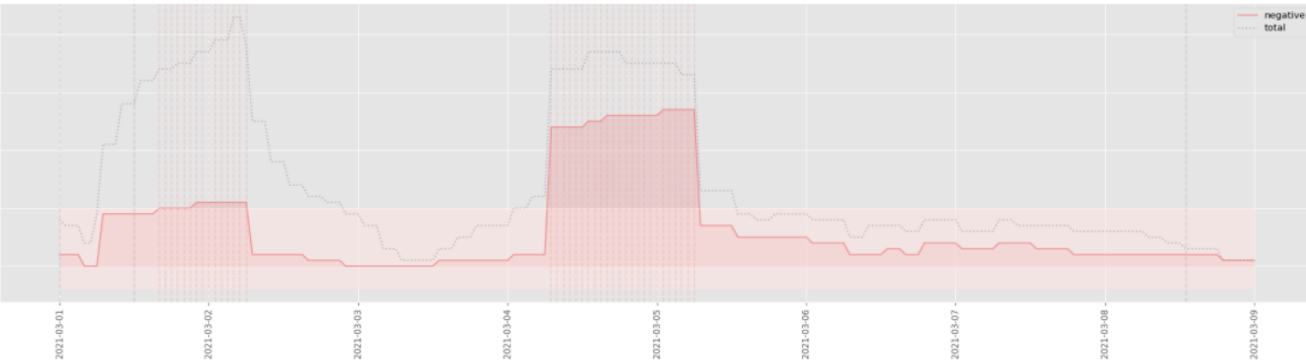


Пример метрик, %

Tag: 1 Fozzy Group  
Avg NSR: 43.29  
Current NPS: -28.57  
Current NSR: -47.06  
Confidence: 0.56  
Samples Number: 112

Комментарий

Стандартизация упоминаний позволяет мониторинг динамики по ключевым метрикам (верхний график) и находить Тренды обсуждения (пример нижний график)



«Жаль. ...а вы бы ходили в Сильпо ? Сельпо НЕ БУДЕТ покупать Фуршет во Львове, Житомире, Василькове и еще трех городах Сеть сельпо отказалась покупать магазины Фуршета в шести городах через угрозы ограничения конкуренции. Читать на русском Компания Сильпо-Фуд (Fozzy Group) не будет...»  
<https://rau.ua/новыни/новини-компаний/сильпо-не-буде-купувати-фуршет/>

# Для применения Месседж Тагер необходимо выполнение основных критериев

## Требования к применимости системы

1



### **Текстовые данные**

Регулярный поток текста-письма, документы отзывы  
**Вскором добавится голос!**

2



### **Рабочий процесс**

Формализован процесс использования текстовой информации

3



### **Исторические данные**

Наличие уже размеченных примеров или подобной базы знаний

4



### **Проведение Feasibility study**

Анализ применимости системы к процессу и типу данных

## Этапы внедрения, если выполняются пункты выше

- Создание нового домена в Месседж Таггер
- Выбор способа обмена данными
- Определение ключевых показателей и требований
- Конфигурация работы системы для нового домена
- Разработка моделей ИИ
- Подготовка MVP