



Инструменты генеративного ИИ для мониторинга информационного пространства и проведения социологических и маркетинговых исследований

Акульшин М.В., **SURVEYSTUDIO**

Принципы использования



- **Масштаб и скорость.** LLM читает тексты как внимательный человек, но круглосуточно и в любом объёме — тысячи сообщений в час, без усталости и очереди на разметку.
- **Онлайн-мониторинг.** Группы и каналы, отзывы, публикации и СМИ, форумы — непрерывное наблюдение вместо периодических «замеров» и ручных сводок.
- **Понимание, а не подсчёт слов.** Сарказм, смешанные оценки, контекст: модель точно идентифицирует, реферирует и классифицирует события и мнения.
- **Структура на выходе.** Классификация по заданной таксономии, извлечение проверяемых фактов, готовые черновики реакций — сразу в базу и на дашборд.
- **Исследовательский контур.** Открытые вопросы анкет, интервью, фокус-группы: качественные данные обрабатываются в количественном масштабе.
- **Экономика.** Разбор одного документа стоит копейки — мониторинг из разового проекта превращается в фоновый процесс.

Пример: сеть кофеен «Зерно сомнения»



Четыре кофейни, отзывы на четырёх площадках: Яндекс.Карты, 2ГИС, ВКонтакте, доставка. За две недели — 36 сообщений. Кто-то должен их прочитать, рассортировать, на часть ответить, а из потока сделать вывод для управляющего.

На практике маркетолог просматривает отзывы «когда есть минутка», отвечает самым громким, а сводку делает раз в квартал по памяти. Сигналы тонут: жалоба на цены внутри четырёхзвёздочного отзыва, постоянный гость, который уже дважды ушёл без покупки, три жалобы на очередь за неделю в одной и той же точке.

Датасет синтетический: название кофейни вымышленное, промпт, код и цены — настоящие.

4

кофейни в Петербурге

4

площадки с отзывами

36

сообщений за 14 дней

20

требуют ответа

Анализ тональности: где словарь ошибается



Наивный словарный анализатор совпал с разметкой модели лишь в **44%** случаев (16 из 36). Ошибки структурные:

«Ну спасибо за „быструю подачу" — всего 25 минут за одним капучино отстояла. Отдельный комплимент кондиционеру...»

Словарь **позитивный** → LLM **негативный** сарказм

Словарь увидел «спасибо» и «комплимент». Сарказм читается только смыслом.

«Отдельный респект за систему „один бариста на весь Невский".
Бесплатная медитация в очереди... Дзен обрела, капучино — минут через двадцать.»

Словарь **позитивный** → LLM **негативный** скорость-

«Респект», «бесплатно», «дзен» — три позитивных слова в жалобе на очередь.

«Апдейт: написали, извинились, вернули деньги. Приехало за 35 минут, горячее... Вот так надо работать с косяками!»

Словарь **негативный** → LLM **позитивный** доставка+, персонал+

Гостя хвалит работу с жалобой, но слово «косяки» утащило отзыв в негатив.

«Парковку у ТЦ перекрыли, охрана хамит... Кофейня, в общем-то, ни при чём, но настроение испорчено, звёзды снимаю.»

Словарь **негативный** → LLM **не по адресу**

Претензия к торговому центру: LLM исключила отзыв из метрик кофейни.

Что добавляет LLM



Языковая модель читает отзыв как внимательный сотрудник, а возвращает то, что нужно базе данных. Весь «интеллект» системы — **один системный промпт** со схемой ответа.

- Фиксированный список из восьми аспектов — иначе агрегировать будет нечего
- Несколько аспектов с разной тональностью у одного отзыва
- В факты — только то, что можно процитировать
- Каждый аспект подтверждается цитатой из текста
- Тон бренда и запрет канцелярита для черновиков ответов

```
{  
  "overall": "позитивный | негативный |  
              смешанный | нейтральный",  
  "aspects": [{"aspect": "скорость",  
                    "sentiment": "негатив",  
                    "quote": "25 минут за одним капучино"}],  
  "urgency": "низкая | средняя | высокая",  
  "sarcasm": true,  
  "facts": ["только проверяемая конкретика"],  
  "insight": "вывод для команды",  
  "reply_draft": "ответ в тоне бренда"  
}
```

Как устроено



Ни одного экзотического компонента: сбор, один вызов LLM на отзыв со строгой JSON-схемой, валидация и обычная BI-витрина поверх.

- 1 Источники.** Яндекс.Карты, 2ГИС, VK, агрегаторы доставки
- 2 Сбор.** API площадок и выгрузки; новые отзывы — каждые 30 минут
- 3 LLM-разбор.** Один вызов на отзыв, строгая JSON-схема, температура 0
- 4 Валидация.** Проверка схемы; невалидный JSON — повторный запрос
- 5 Витрина.** База + BI-дашборд: тренды, аспекты, точки
- 6 Действия.** Алерты в Telegram, очередь ответов с черновиками

Главный «код» системы — системный промпт (~3 тыс. знаков). Открыт целиком: demo.smyslograf.ru/about



8 против 1

негативных упоминаний «скорости» на Невском,
неделя к неделе

21 августа

сработал алерт «скорость · Невский»

19

инсайтов для продукта и маркетинга за две недели

С 17 августа в отзывах точки на Невском стал повторяться аспект «скорость»: «один бариста на смене», «очередь до двери». По отдельности — обычные плохие дни. В агрегате — 8 негативных упоминаний за шесть дней против одного неделей ранее.

Причина оказалась операционной — из утренней смены ушли люди. Но заметил её **не управляющий, а мониторинг отзывов**.

Побочный продукт — поле insight: от спроса на абонемент до просьбы вернуть облепиховый чай к осени. Раньше это тонуло в потоке; теперь — готовая повестка продуктовой встречи.

Тональность: **все** позитивный смешанный негативный нейтральный

Аспект: **все** скорость напитки еда персонал цены доставка чистота атмосфера

Точка: **все** Невский Васильевский Петроградка Московский только требующие ответа

Показано 36 из 36

Яндекс.Карты ★★★★★ Инга А. · 23 авг · Васильевский

Привела родителей показать «мою» кофейню. Папа, человек старой закалки и растворимого кофе, попробовал флэт уайт и сказал: «Ну, за такое и жить можно». Считаю это высшей наградой для вас.

позитивный срочность: низкая нужен ответ напитки атмосфера [показать разбор](#)

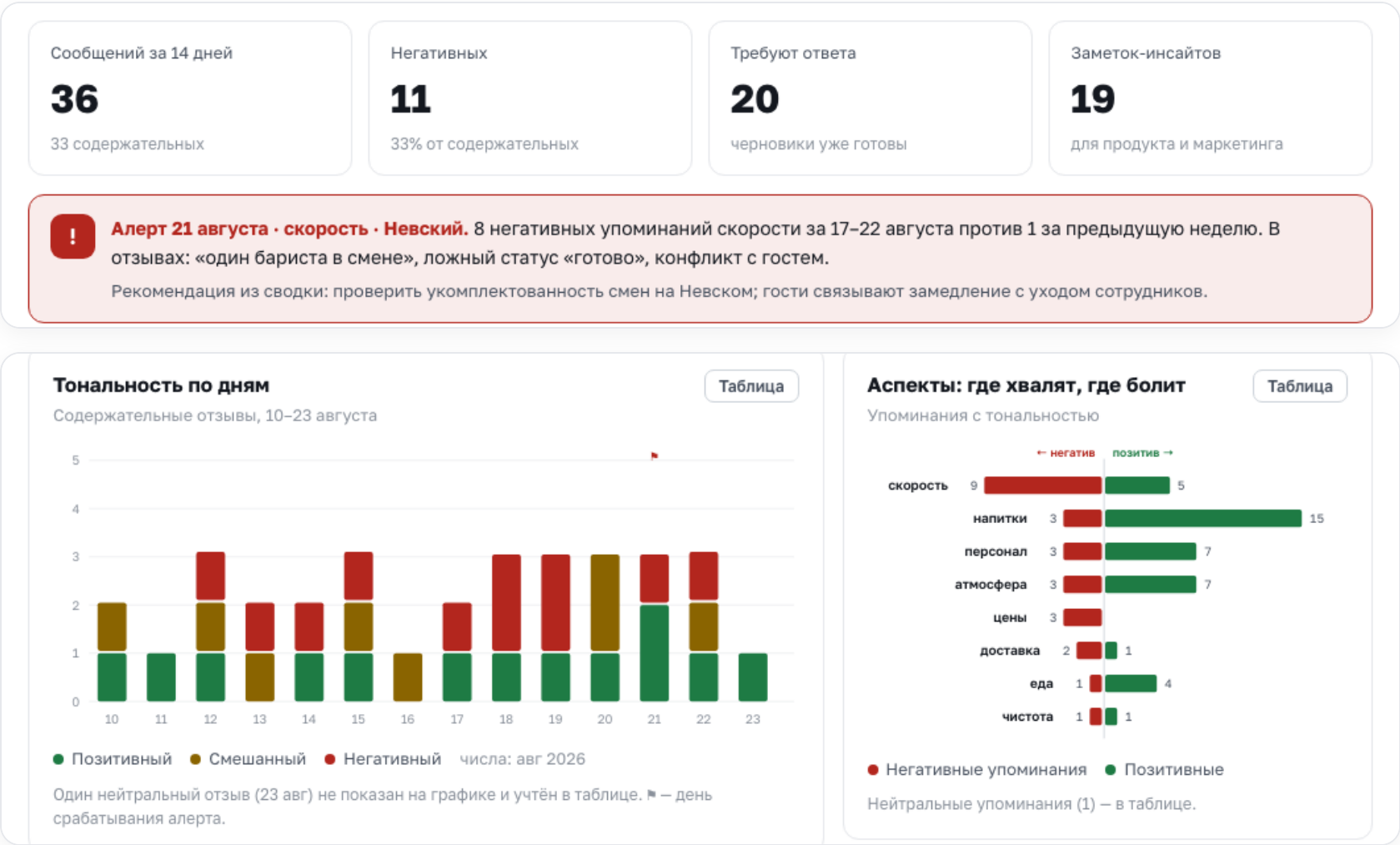
Яндекс.Карты ★★★☆☆ Вера И. · 23 авг · Невский

Апдейт: в воскресенье утром за стойкой снова двое, очередь прошла минут за пять. Если так будет и в будни — вернусь в ряды утренних постоянников. Пока нейтрально, наблюдаю.

нейтральный срочность: низкая нужен ответ скорость [показать разбор](#)

Фильтры — те самые поля, которые извлекла LLM: до разбора такой фильтрации не существовало.

Дашборд



Все цифры посчитаны по структурированным разборам — ни одна не размечалась вручную.

МОДЕЛЬ	ОДИН ОТЗЫВ	1000 ОТЗЫВОВ В МЕСЯЦ
Лёгкая модель	≈ 25 коп.	≈ 250 Р
Модель классом выше	≈ 50 коп.	≈ 500 Р
Лёгкая модель через Batch API	≈ 13 коп.	≈ 125 Р

Один отзыв ≈ 1300 токенов на входе и 350 на выходе.

≈ 250 Р

месяц мониторинга тысячи отзывов на лёгкой модели

Капучино стоит 300–350 рублей: **месяц мониторинга дешевле одной чашки**. Считать нужно не токены, а время маркетолога — полчаса в день на чтение потока превращаются в пять минут на сводку.

Подводные камни



Модель может ошибиться или додумать

Предохранители: температура 0 и жёсткая схема, программная сверка цитат, выборочный ручной контроль 5–10% разборов еженедельно.

Черновик — не публикация

Ответы гостям публикует человек; маркетолог из автора становится редактором.

Данные и право

Имена гостей маскировать до отправки в API, сырые тексты хранить у себя, правила площадок и 152-ФЗ проверить с юристом.

Промпт — это код

Версионирование и регрессионный набор из 30–40 сложных отзывов перед каждой правкой.

С чего начать: неделя до работающего контура



- 1 Выгрузите отзывы за 2–3 месяца с двух главных площадок — хватит CSV
- 2 Возьмите промпт из демо, адаптируйте аспекты (8–10) и тон бренда
- 3 Прогоните массив через лёгкую модель, посмотрите 30 разборов глазами, зафиксируйте регресс-набор
- 4 Поставьте разбор на расписание: тренд по аспектам, очередь ответов, лента инсайтов
- 5 Добавьте один алерт: всплеск негатива по аспекту неделя к неделе

Вопросы



- ? А если модель выдумает то, чего в отзыве нет?
- ? Что с персональными данными?
- ? Почему не готовый сервис мониторинга?
- ? Кто нужен, чтобы это запустить?
- ? Публичные ответы тоже пишет модель?
- ? А если в отзыве напишут инструкции для модели?
- ? Почему не дообучить маленькую модель классификации?
- ? А мониторинг Telegram-чатов и групп?

Ответы — на demo.smyslograf.ru/about, раздел «Вопросы с задних рядов».

Спасибо за внимание

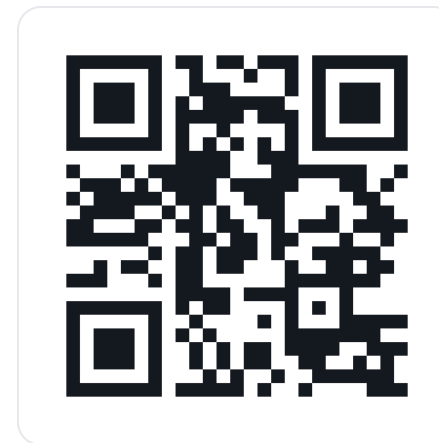


Максим Акульшин

генеральный директор
ООО «Системы информации и связи»

SURVEYSTUDIO

info@survey-studio.ru · +7 921 918-89-12



demo.smyslograf.ru

Здесь можно ознакомиться с этой презентацией,
скачать исходный промпт и поработать со стендом
живого разбора отзывов — попробуйте сами.