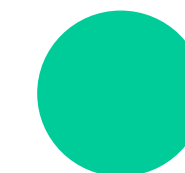




T R A F L O W

Нейросети

актуальность, кейсы
и куда бежать,
если очень надо



Актуальность и проблемы бизнеса

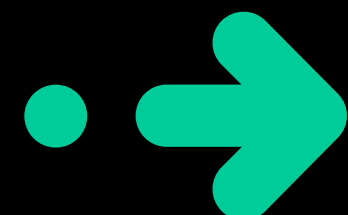
- **Рост объемов user-generated контента (отзывов, постов, комментариев, графического контента)**
- **Дорогостоящая обработка**
- **Риски, связанные с пропуском «токсичных» данных**



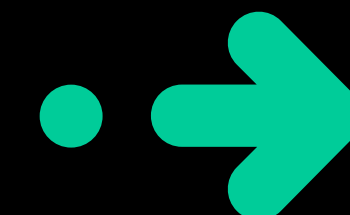
TRAFLOW

Кейсы

**Определение
тональности**



**Модерация
графического
контента**



**Персонализация
контентного
сайта**

Данные

- **Комментарии к постам**
- **Обратная связь от клиентов**
- **Информационный фон вокруг бренда или события**

Кейс 1: Определение ТОНАЛЬНОСТИ

Подходы

- 1 Ручная обработка**
- 2 «Экспертный» подход**
- 3 Машинное обучение**
- 4 Активное машинное обучение**
- 5 AutoML**

Что поменялось?

Раньше: вручную созданные наборы правил обработки данных на основе статистики и эвристики

Сейчас: машинное обучение – это процесс, в ходе которого система обрабатывает большое число примеров, выявляет закономерности и использует их, чтобы прогнозировать характеристики новых данных

Машина лучше знает как решить задачу





TRAFL © W

Задача

Оградить пользователей портала от фото и видеоконтента нарушающего законодательство или правила платформы (AppStore & Google Play)

Кейс 2: Модерация графического контента

Подходы

- 1 Отсутствие автоматизации**
- 2 Предфилترация с уменьшением штата модераторов**
- 3 Автоматизация + реакция на жалобы**



T R A F F L O W

Цель

Повысить показатели вовлеченности и удержания портала, показывая пользователям интересующий их контент

Кейс 3:

Персонализация контентного сайта

Подходы

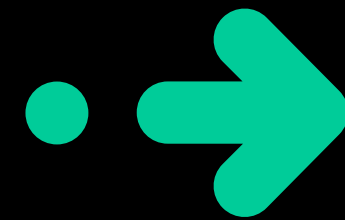
- 1** Отсутствие автоматизации
- 2** Разработка формулы и автоматическое ранжирование
- 3** Разработка метрики близости и поиск похожих исходя из нее
- 4** Использование «выученной» нелинейной метрики близости

Нейросети – инструмент, а не магия и не панацея

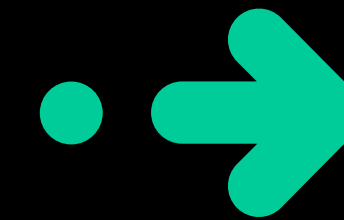
- Плохая интерпретируемость
- Чувствительность к объему и качеству данных
- Необходимость подстраиваться под изменения данных
- Затраты на инфраструктуру
- Adversarial атаки
- Непредсказуемые сроки разработки для малоизученных применений
- «Кража» моделей через публичный API

Что стоит делать если вы хотите применять нейросети?

Бизнес идея



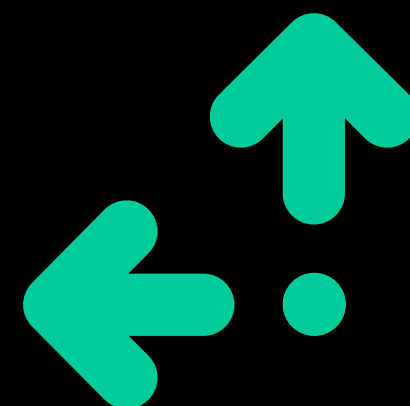
**Сбор, разметка
и подготовка
данных**



Моделирование



**Применение
модели**



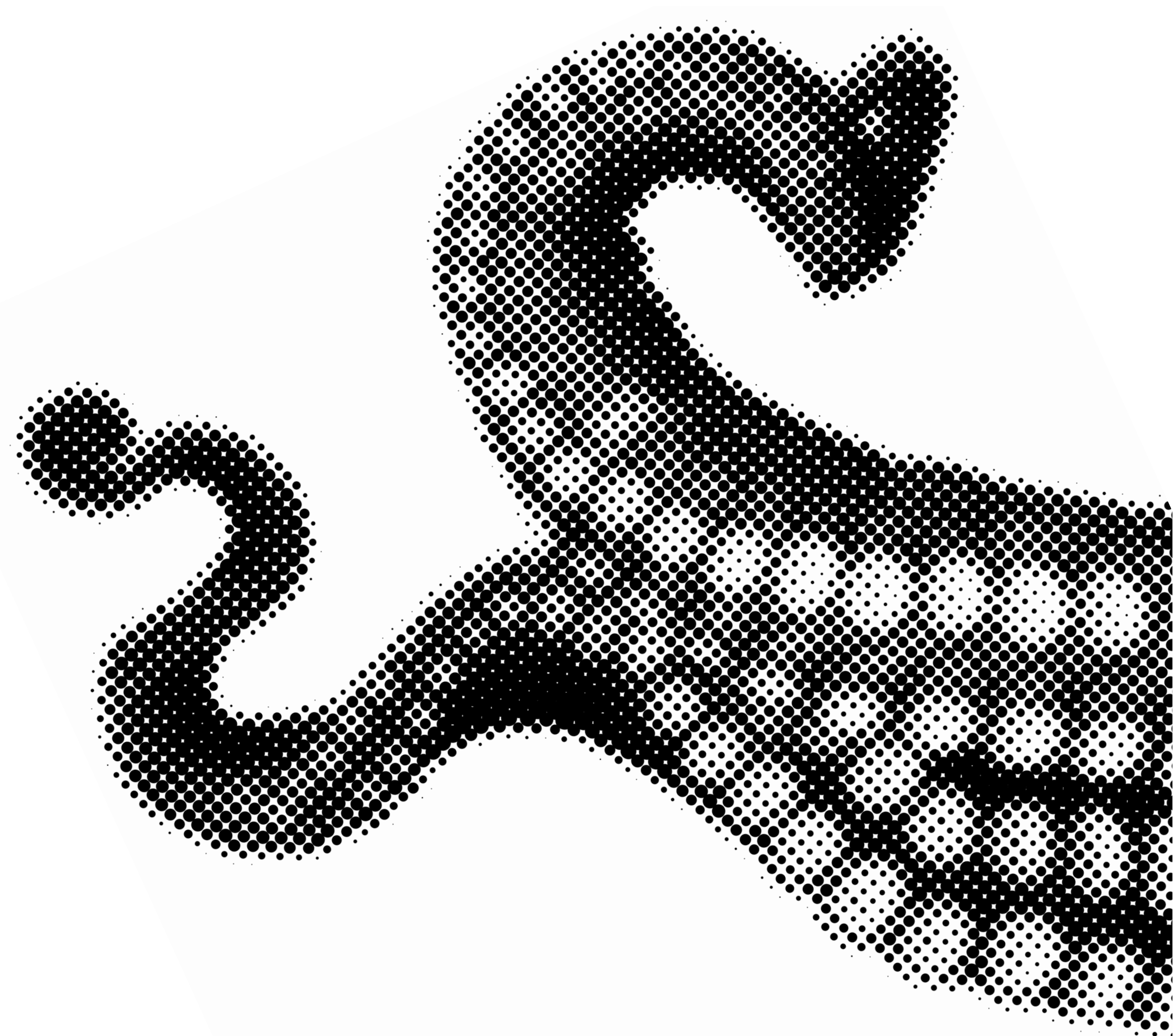
Тестирование

Какие варианты и сколько это стоит?

- \$** Есть возможность применить готовое решение
- \$\$\$** Есть идея, данных нет
- \$\$** Есть бизнес-задача и данные
- \$?** Есть желание разобраться самостоятельно

- ✓ **Доверие к Data Driven решениям растет**
- ✓ **Рутинная работа постепенно автоматизируется**
- ✓ **Данные становятся важным активом и инструментом для компаний**
- ✓ **Порог входа снижается благодаря облачным сервисам (IaaS/PaaS/SaaS)**

Текущий тренд





T R A F L O W

Спасибо за внимание!

