ИСКУССТВЕННЫЙ И ЕСТЕСТВЕННЫЙ ИНТЕЛЛЕКТ

В АНАЛИЗЕ НЕСТРУКТУРИРОВАННЫХ ТЕКСТОВЫХ ДАННЫХ ИЗ СОЦИАЛЬНЫХ МЕДИА

Ipsos

ОИРОМ
ОБЪЕДИНЕНИЕ ИССЛЕДОВАТЕЛЕЙ РЫНКА И ОБЩЕСТВЕННОГО МНЕНИЯ

«Люди не автоматы, их решения не всегда надежны. На их выводы сильно влияют разные факторы: сиюминутное настроение, погода, желание заморить червячка. В этом смысле человек сильно уступает искусственному интеллекту, действующему по алгоритму и защищённому от таких шумов.»

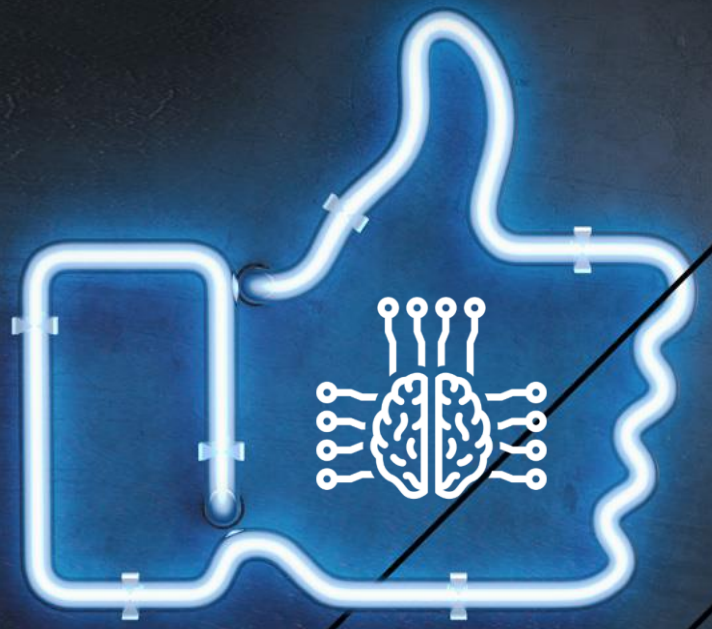
Даниэль Канеман

ПСИХОЛОГ, ЛАУРЕАТ НОБЕЛЕВСКОЙ ПРЕМИИ ПО ЭКОНОМИКЕ 2002 ГОДА

[redacted]

С ПОМОЩЬЮ ИИ

можно работать с данными соц. медиа
для решения широкого
круга бизнес задач



АНАЛИЗ ВОСПРИЯТИЯ БРЕНДА
(ИМИДЖ, СИТУАЦИИ ИСПОЛЬЗОВАНИЯ)

АНАЛИЗ ВОСПРИЯТИЯ КАТЕГОРИИ
(U&A)

АНАЛИЗ СИТУАЦИЙ ПОТРЕБЛЕНИЯ

АНАЛИЗ ТРЕНДОВ

В НАШЕМ АРСЕНАЛЕ СЕГОДНЯ БОЛЬШОЙ СПЕКТР МЕТОДОВ

- PLSA
- DOC2VEC
- LDA
- BERT
- Universal Sentence Encoders

НИ ОДИН ИЗ НИХ НЕ МОЖЕТ РАБОТАТЬ САМ ПО СЕБЕ

Ipsos

НАШ ПРИМЕР НА СЕГОДНЯ:
анализ обсуждений двух крупных алкогольных категорий

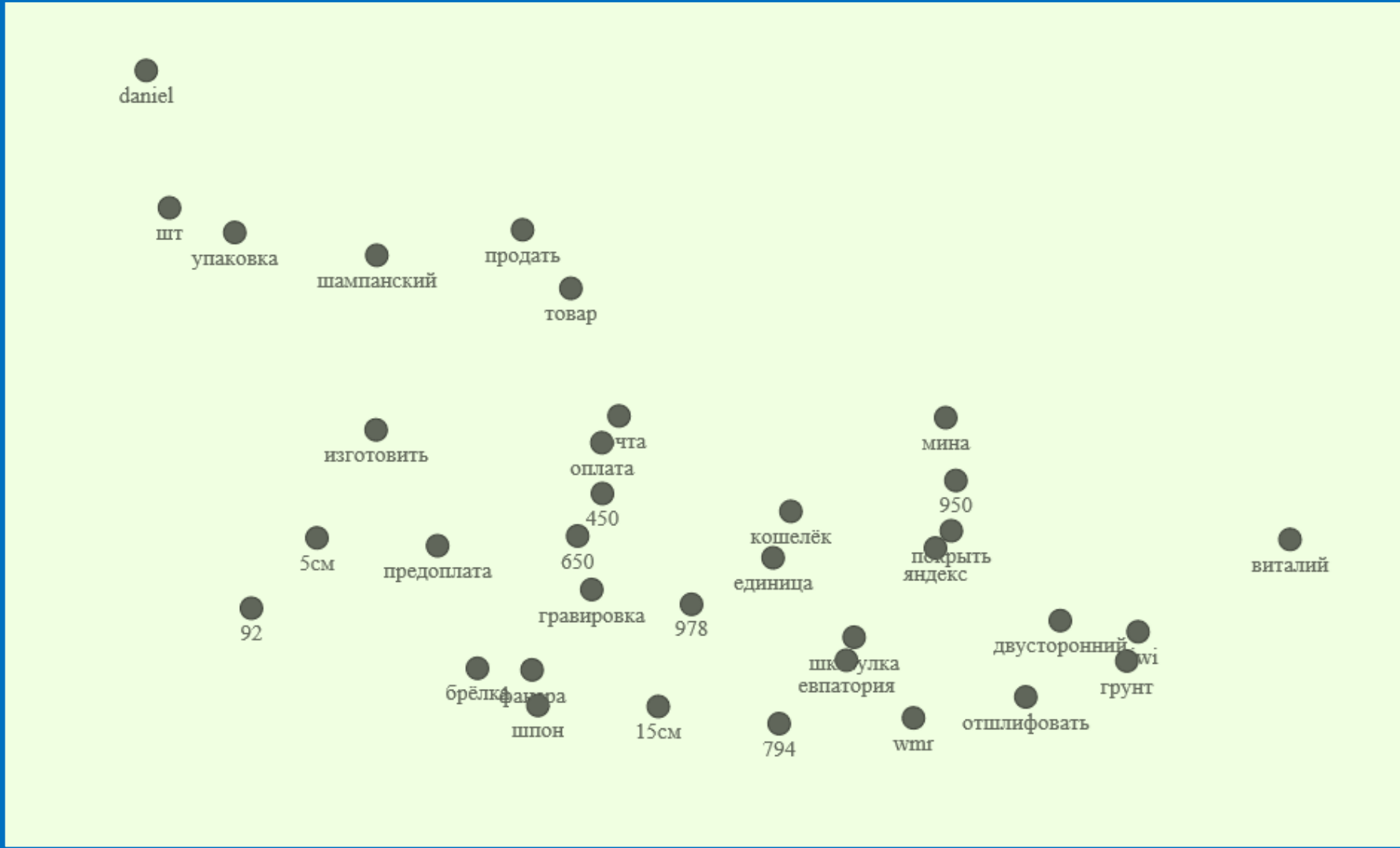
ANALYSIS 0588

26 БРЕНДОВ

ANALYSIS 0548

НЕСКОЛЬКО СОТЕН ТЫСЯЧ УПОМИНАНИЙ

Пример типичного сегмента на «сырых» данных



ПОЧЕМУ ТАК ПОЛУЧИЛОСЬ?

SEGMENT 05500

Человек делает коробки для Jack Daniels и написал об этом несколько тысяч сообщений

Отделяем релевантные сообщения от шума с помощью AI

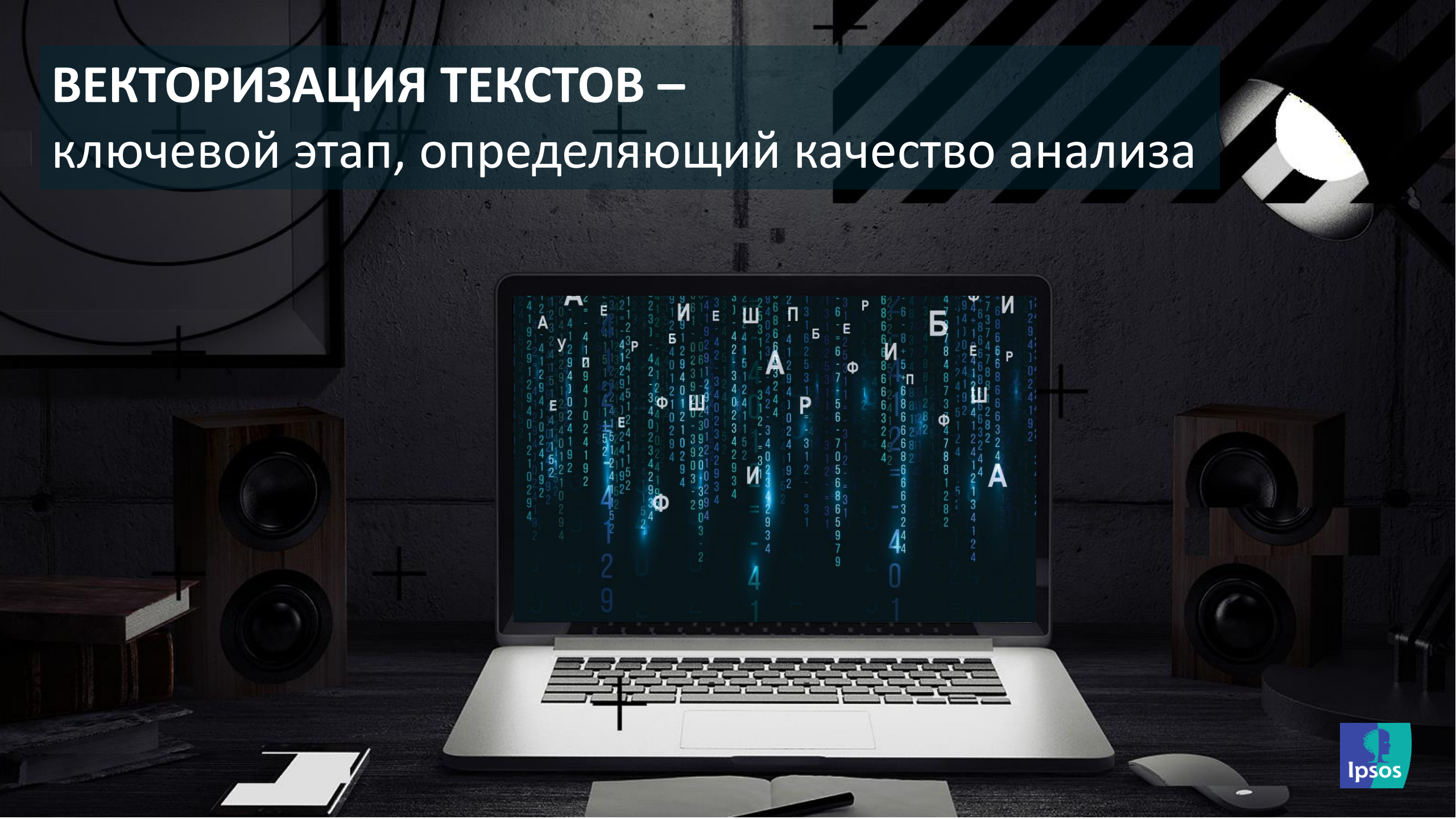
RELEVANT

Блэк бакарди, танцы в моей кровати.
Не говори мне хватит, снимая своё платье. 😋
Девичник удался 😂

SPAM

Жесть, когда в детском центре в ЦУМе играет песня "Блэк Бакарди"
и это хором начинают петь девочки. Чувствуешь, что Россия встаёт с колен.

ВЕКТОРИЗАЦИЯ ТЕКСТОВ –
ключевой этап, определяющий качество анализа



ОСНОВНЫЕ ПОДХОДЫ К ВЫДЕЛЕНИЮ ТЕМ

МОНОСЕГМЕНТАЦИЯ

выделение ключевой идеи сообщения

МУЛЬТИСЕГМЕНТАЦИЯ

выделение основных аспектов в сообщении

“Советы по выбору”

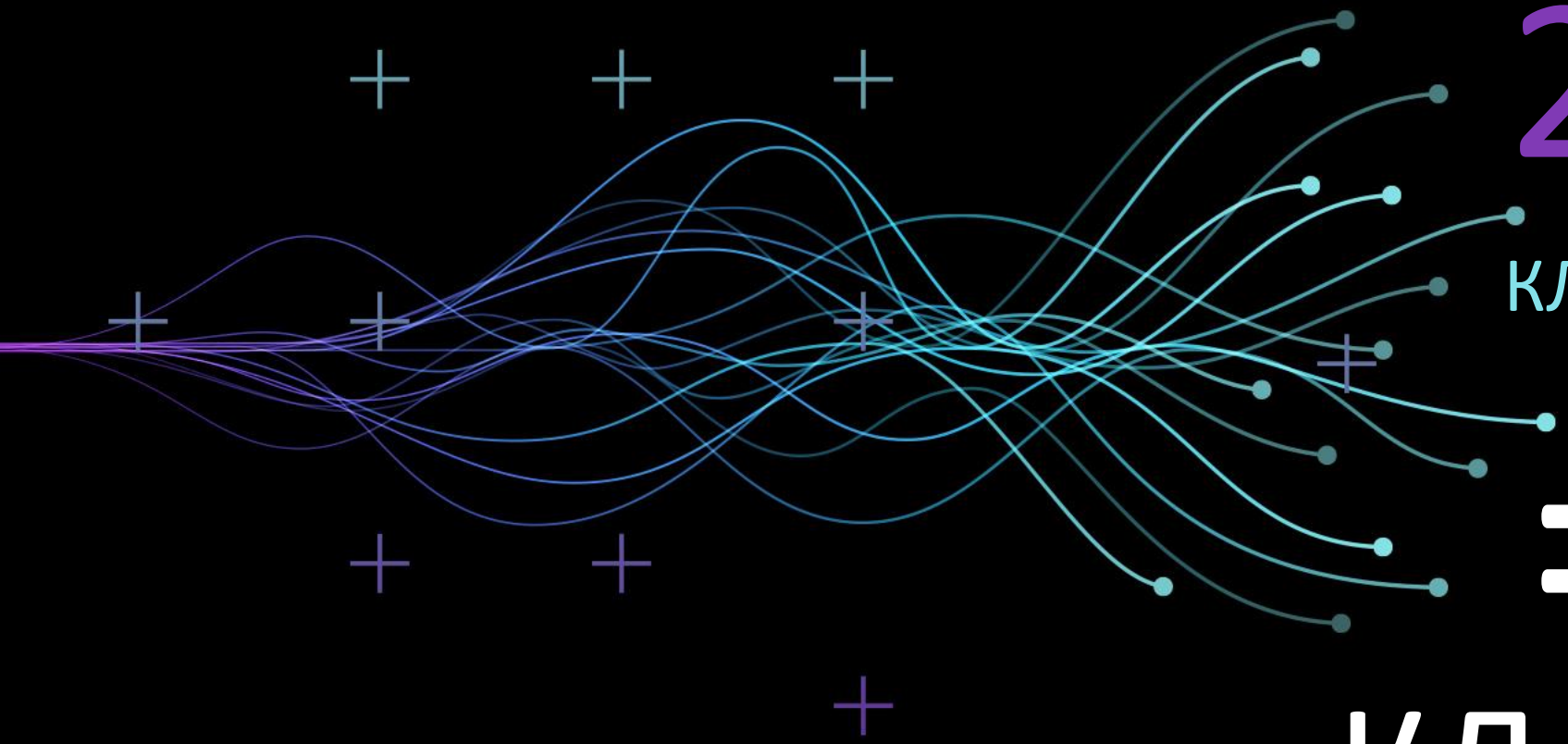
я вообще к блендам отношусь с некоторым недоверием. Гленик по моему, жестковат. Если брать бленды, то рекомендую из доступных обыкновенный Макаллан 12 летний. Или искать свои вкус в односолодовых. Кто-то любит Ирландию кто-то Шотландию. Тут реально дело вкуса и подбора под собственные ощущения.

Тип виски
Ценовой сегмент
Страна происхождения

Ipsos

Метод, который лучше сработает с конкретным массивом данных, невозможно определить заранее.

Использование нескольких методов в различной размерности дает огромный объем данных



20 РЕШЕНИЙ

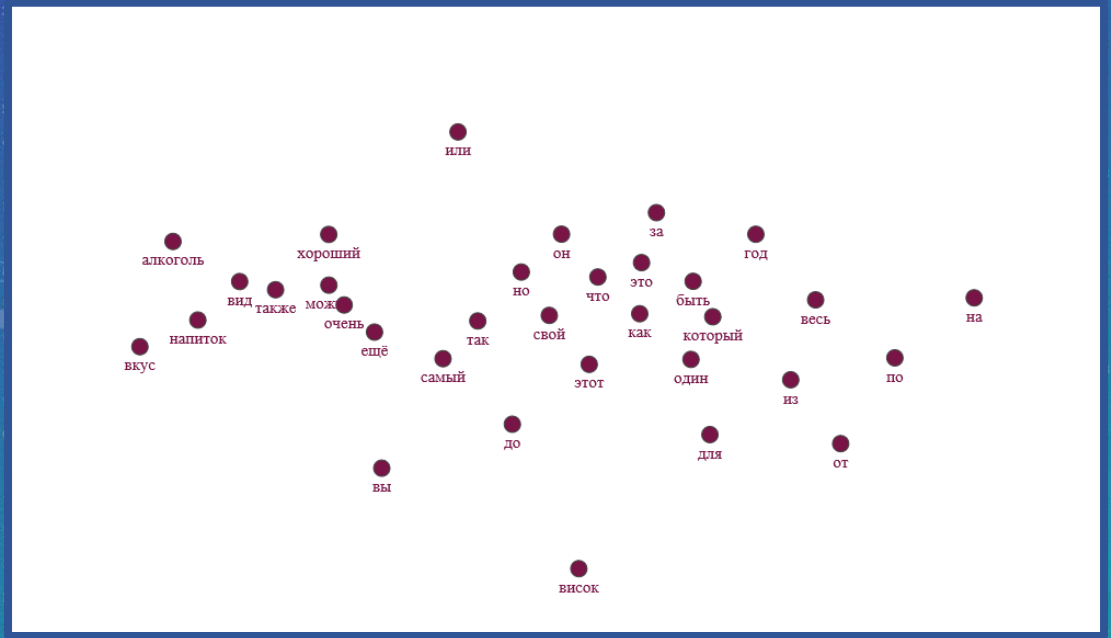
КЛАСТЕРОВ 20

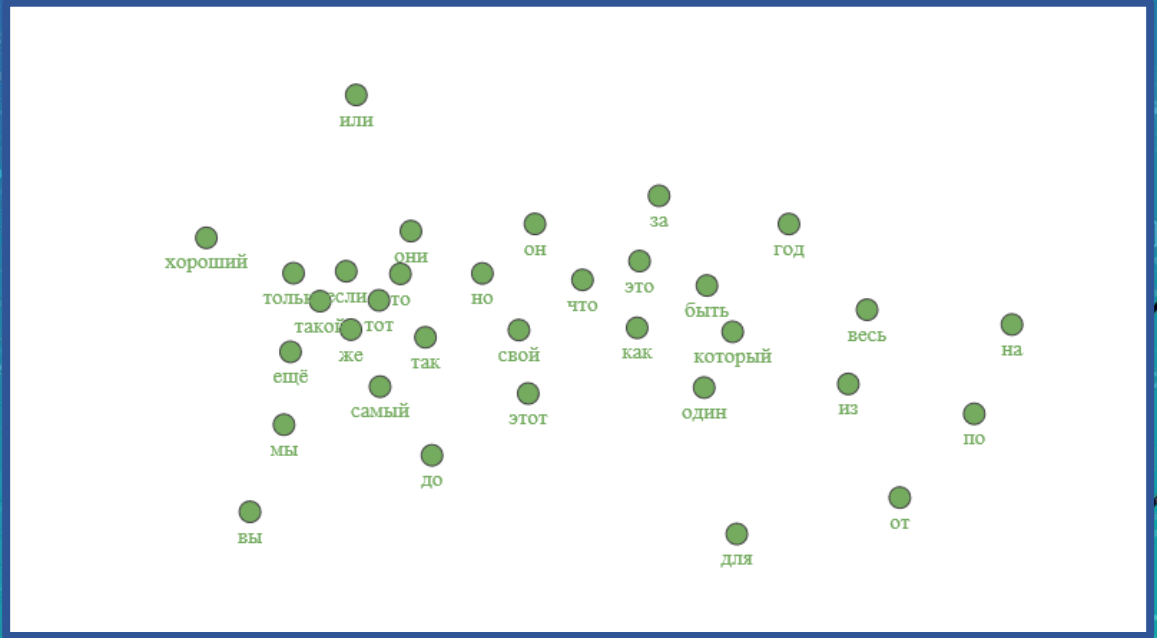
=400

КЛАСТЕРОВ

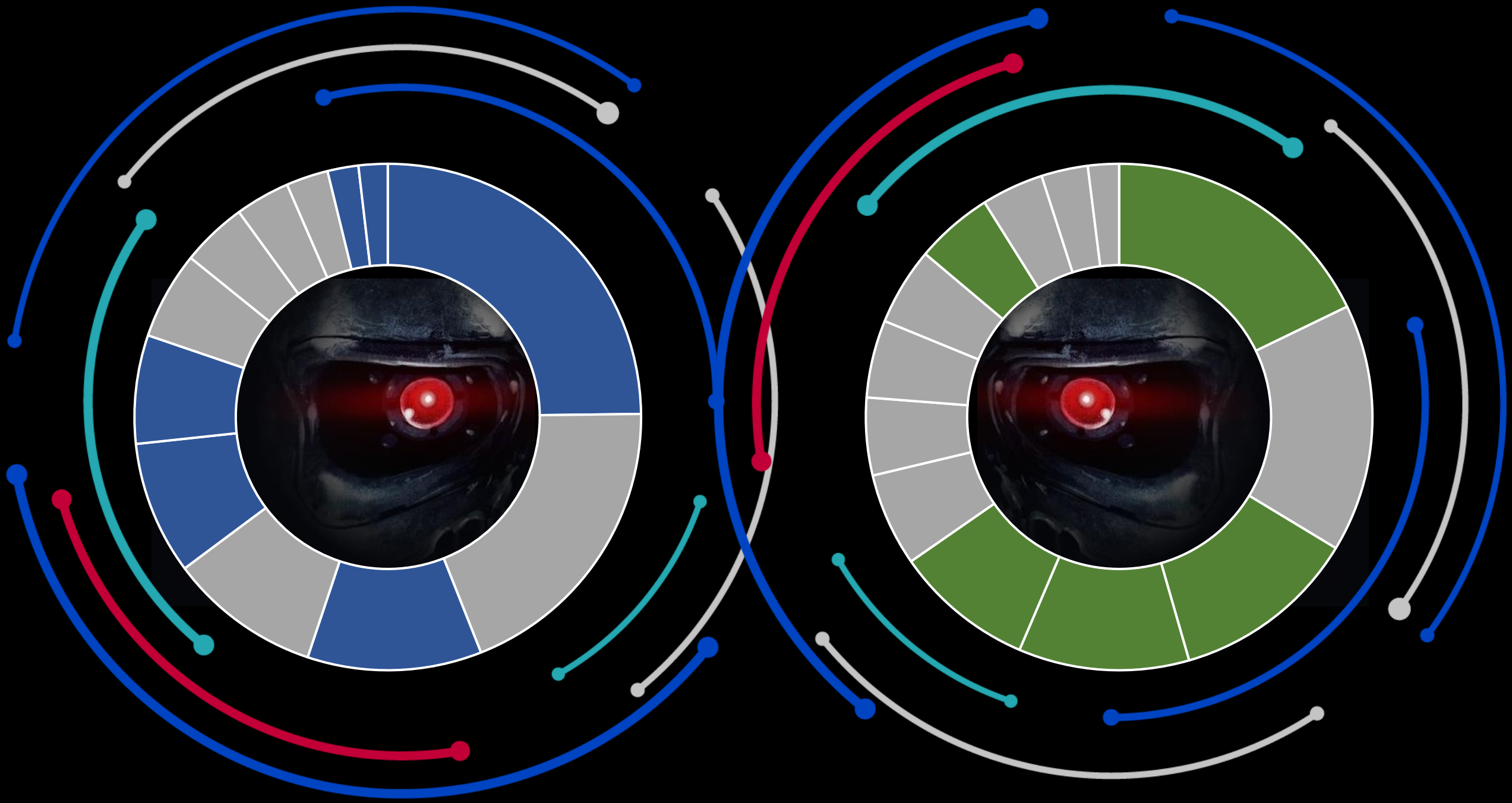
Поэтому мы применяем ряд математических методов для оценки качества сегментов

Но не каждое математически идеальное решение имеет смысл





ТОЛЬКО ЕСТЕСТВЕННЫЙ ИНТЕЛЛЕКТ позволяет сделать окончательный выбор и отшлифовать итоговое решение



ПРИМЕР ОПИСАНИЯ СЕГМЕНТА

Наша схема анализа данных

БАЗОВАЯ ЧИСТКА

ЕСТЕСТВЕННЫЙ ИНТЕЛЛЕКТ

ИИ ЧИСТКА

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

ВЕКТОРИЗАЦИЯ ТЕКСТОВ

ИСКУССТВЕННЫЙ + ЕСТЕСТВЕННЫЙ ИНТЕЛЛЕКТ

ТЕМАТИЧЕСКОЕ КОДИРОВАНИЕ

Сегментационные алгоритмы + естественный интеллект

ВЫБОР И ОПИСАНИЕ РЕШЕНИЯ

ЕСТЕСТВЕННЫЙ ИНТЕЛЛЕКТ

МАШИНА НЕ МОЖЕТ ПОНЯТЬ ВСЕЙ ПОЛНОТЫ КАРТИНЫ,

но идеально решает частные задачи

ЧЕЛОВЕЧЕСКИЙ ИНТЕЛЛЕКТ

больше не способен охватить каждую единицу данных, но способен увидеть картину целиком

НО ВМЕСТЕ ОНИ СПОСОБНЫ НА ВСЕ!