

# Глубокий Анализ YouTube-Каналов: От Векторных Данных к Стратегии Роста

## Часть I: Анатомия Канала — Проектирование Системы Аналитических Инсайтов

В этом разделе излагается комплексный технический конвейер, необходимый для преобразования необработанных векторных данных в богатое, многоуровневое понимание контента канала и его сообщества. Методология выходит за рамки простого семантического поиска и представляет собой сложную систему тематического анализа.

### Раздел 1: Проектирование Аналитической Системы: От Векторов к Темам

#### 1.1. Аналитический Конвейер: Модульный Подход к Пониманию Контента

Для глубокого анализа структуры контента канала предлагается многоэтапный процесс: **Снижение размерности → Кластеризация → Тематическая маркировка**. Такая модульная архитектура позволяет независимо оптимизировать каждый этап, обеспечивая максимальную точность и интерпретируемость результатов.<sup>1</sup> Процесс начинается с запроса к векторной базе данных Quadrant для извлечения векторных представлений (эмбеддингов) всех транскриптов видео. Этот массив данных служит основой для последующего контент-анализа.

#### 1.2. Шаг 1: Снижение Размерности с Помощью UMAP (Uniform Manifold Approximation and Projection)

Текстовые эмбеддинги, полученные с помощью современных моделей, существуют в пространствах с сотнями или даже тысячами измерений. Прямая кластеризация в таких многомерных пространствах часто оказывается неэффективной из-за так называемого "проклятия размерности", при котором метрики расстояния теряют свою значимость, и

все точки становятся примерно равноудаленными друг от друга.<sup>2</sup>

Для решения этой проблемы используется алгоритм UMAP, который проецирует высокоразмерные векторы транскриптов в пространство меньшей размерности (например, от 5 до 50 измерений). В отличие от линейных методов, таких как анализ главных компонент (PCA), UMAP является техникой обучения многообразий, которая превосходно сохраняет как локальную, так и глобальную сложную, нелинейную структуру данных.<sup>4</sup> Это гарантирует, что видео, тематически близкие в исходном векторном пространстве, останутся сгруппированными и в пространстве с пониженной размерностью, что является критически важным для точной кластеризации. Кроме того, этот шаг является необходимым условием для осмысленной двух- или трехмерной визуализации ландшафта контента канала.<sup>5</sup>

### 1.3. Шаг 2: Тематическая Кластеризация с Помощью HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise)

Для группировки видео по темам настоятельно не рекомендуется использовать алгоритм K-Means. K-Means принудительно распределяет каждый элемент данных по заранее определенному количеству кластеров (k) и предполагает, что эти кластеры имеют сферическую форму, что крайне редко соответствует действительности для сложных тематических структур.<sup>7</sup> Такой подход может привести к формированию бессмысленных групп, где не связанные между собой видео оказываются в одном кластере, а целостные темы искусственно разделяются.

В качестве основного инструмента кластеризации рекомендуется использовать HDBSCAN. Его ключевые преимущества для данной задачи заключаются в следующем:

- **Отсутствие необходимости задавать количество кластеров:** HDBSCAN самостоятельно определяет естественное количество тематических кластеров, присутствующих в данных. Это крайне важно, поскольку априори неизвестно, сколько основных тем развивает автор на своем канале.<sup>9</sup>
- **Работа с кластерами произвольной формы:** Алгоритм способен идентифицировать кластеры различной формы и плотности, что позволяет точно моделировать как широкие, так и узкоспециализированные контентные темы.<sup>6</sup>
- **Идентификация выбросов как "шума":** HDBSCAN маркирует видео, которые не принадлежат ни к одному coherentному тематическому ядру, как "шум" или выбросы (присваивая им метку кластера -1).<sup>1</sup>

Идентификация "шумовых" точек — это не недостаток алгоритма, а мощный стратегический сигнал. Эти видео-выбросы не являются ошибками анализа; они представляют собой контент, который отклоняется от устоявшихся тематических ядер канала. Это могут быть единичные эксперименты, неудачные попытки запустить новые форматы или спонсорские интеграции, не совпадающие с основной контентной политикой. Выделив эти видео в отдельную категорию, можно провести их целенаправленный анализ. Обладают ли они аномально высоким или низким уровнем

вовлеченности? Выражают ли комментарии под ними замешательство или, наоборот, восторг от новизны? Анализ таких выбросов дает прямую обратную связь на инновационные попытки автора — это гораздо более глубокое понимание, чем принудительное распределение каждого видео по тематическим "коробкам", как это сделал бы K-Means.

#### **1.4. Шаг 3: Тематическая Интерпретация с Помощью Больших Языковых Моделей (LLM)**

Процесс кластеризации на выходе дает группы идентификаторов видео, которые сами по себе не несут смысловой нагрузки. Необходимо понять содержательную суть каждого кластера. Для этого для каждого кластера отбирается репрезентативная выборка транскриптов (например, 5-10 видео, наиболее близких к центру масс кластера). Эти транскрипты подаются на вход мощной LLM с использованием структурированного, ролевого промпта.

##### **Шаблон Промпта: Персона "Канальный Аналитик"**

**\*\*Персона:\*\*** Ты — эксперт-стратег по контенту YouTube и аналитик с докторской степенью в области медиаисследований. Ты обладаешь острым умом, проницательностью, и твоя цель — преобразовать сложные паттерны контента в ясные, действенные темы.

**\*\*Контекст:\*\*** Тебе предоставлена коллекция транскриптов нескольких видео с YouTube, которые были алгоритмически сгруппированы на основе семантической близости. Твоя задача — проанализировать эти транскрипты, чтобы определить их основную, объединяющую тему.

**\*\*Задача:\*\***

1. Внимательно прочти все предоставленные транскрипты.
2. Определи центральную тему, основное послание или цель автора, а также типичный формат или стиль подачи (например, tutorial, документальный фильм, влог, дебаты).
3. Сгенерируй краткое, запоминающееся и описательное "Название темы" (3-5 слов) для этого кластера видео.
4. Напиши "Тематическое резюме" (80-100 слов), которое подробно объясняет тему. Оно должно описывать, о чем эти видео, на какие ключевые вопросы они отвечают, какую позицию занимает автор и какую ценность они представляют для зрителя.

**\*\*Формат:\*\*** Ответ ТОЛЬКО одним JSON-объектом с двумя ключами: "theme\_title" и "thematic\_summary".

Такой структурированный подход, использующий персону, контекст, задачу и формат, обеспечивает стабильно высокое качество и машиночитаемый вывод для каждого кластера.<sup>11</sup> Этот процесс преобразует необработанные кластеры в понятный для человека

**Тематический Ландшафт** канала.

## 1.5. Параллельный Анализ: Картирование Голоса Аудитории

Точно такой же аналитический конвейер (UMAP → HDBSCAN → LLM-маркировка) должен быть применен независимо к векторным эмбедингам всех **пользовательских комментариев**.

Это позволит выявить основные темы обсуждений внутри сообщества, которые могут совпадать, а могут и существенно отличаться от тем контента. В результате могут быть обнаружены кластеры, посвященные "Часто задаваемым вопросам", "Техническим проблемам", "Похвале ведущего" или "Спорам на тему X". Это создает второй, критически важный слой анализа: понимание того, о чем на самом деле говорит аудитория.

## Раздел 2: Измерение Истинного Резонанса: Взвешенный Коэффициент Вовлеченности (WES)

### 2.1. Уход от "Метрик Тщеславия"

Простые показатели, такие как количество лайков или даже просмотров, могут вводить в заблуждение. Видео может иметь много просмотров, но оказывать слабое влияние. Истинный резонанс контента проявляется в инвестициях времени и усилий со стороны аудитории.<sup>13</sup>

Для создания более тонкого показателя эффективности вводится **Взвешенный Коэффициент Вовлеченности (Weighted Engagement Score, WES)**. Принцип его расчета основан на том, что не все взаимодействия равноценны: комментарий требует больше усилий, чем лайк, а репост является публичной рекомендацией и актом поддержки.<sup>13</sup>

### 2.2. Формула WES

Предлагается настраиваемая формула для расчета WES:

$$WES = (wlikes \cdot \text{Лайки}) + (wcomments \cdot \text{Комментарии}) + (wreposts \cdot \text{Репосты})$$

**Рекомендуемые стартовые веса:** wlikes=1, wcomments=5, wreposts=10.

Важно понимать, что эти веса не являются константой; это стратегические рычаги.

Канал, ориентированный на построение сильного сообщества, должен увеличить вес комментариев (wcomments). Канал, нацеленный на максимальный охват и узнаваемость бренда, должен повысить вес репостов (wreposts).<sup>16</sup>

Чтобы сделать этот показатель сравнимым между видео с разным количеством просмотров, его необходимо нормализовать, например, на 1000 просмотров. Это создает метрику **WES на 1000 просмотров**:

$WES1k = \text{Просмотры} \cdot WES \cdot 1000$

### 2.3. Тепловая Карта Эффективности: Связь Тем с Вовлеченностью

На этом этапе объединяются два потока анализа: тематические кластеры из Раздела 1 и рассчитанный для каждого видео WES. Путем усреднения показателя WES1k для всех видео в рамках каждого тематического кластера можно напрямую измерить, какие темы вызывают наиболее осмысленную реакцию аудитории.

#### Ключевой Инструмент: Тепловая Карта Тематической Эффективности

Эта таблица является центральным диагностическим инструментом всего анализа. Она предоставляет единое, ясное представление о том, какие темы контента "работают", какие — нет, и где могут скрываться "неограниченные алмазы". Она напрямую отвечает на главный вопрос пользователя: "какой контент и в каком формате находит наибольший отклик у аудитории канала".

Переход от разрозненных данных (транскрипты и цифры вовлеченности) к стратегическому пониманию происходит следующим образом: кластеризация структурирует контент по темам, WES придает вовлеченности осмысленное значение, а данная таблица соединяет эти два элемента. Это позволяет перейти от тактических наблюдений в духе "Видео X получилось удачным" к стратегическому выводу: "Тема Y, характеризующаяся [резюме], стабильно генерирует глубокую вовлеченность аудитории".

Название темы	Тематическое резюме	Количество видео	Среднее кол-во просмотров	Средний WES1k	Доминирующая тональность комментариев
Продвинутое создание виджетов	Углубленные tutorиалы по созданию сложных виджетов с использованием нестандартных техник.	12	150,000	245.8	Положительная

Обзоры новых гаджетов	Первые впечатления и детальные разборы последних технологических новинок.	25	220,000	180.2	Смешанная
История технологий	Документальные видео об эволюции ключевых технологий и их влиянии на общество.	8	95,000	155.1	Положительная
Эксперименты с форматами	Видео, не вписывающиеся в основные рубрики (влоги, интервью, стримы).	5 (выбросы)	40,000	110.5	Нейтральная
Новости индустрии	Еженедельные дайджесты новостей и событий в технологической сфере.	30	75,000	95.4	Нейтральная

Таблица будет отсортирована по убыванию среднего WES1k, чтобы немедленно выделить наиболее эффективные столпы контента.

## Часть II: Стратегический План — Преобразование Аналитики в Действия

Этот раздел переводит диагностические выводы из Части I в предписывающую и действенную стратегию для создателя контента. Он служит мостом от данных к решениям.

### Раздел 3: Проект Контент-Стратегии на Основе Данных

### **3.1. "Столпы Усиления": Масштабирование Успешных Направлений**

Этот раздел фокусируется на темах из верхнего квартиля Тепловой Карты Эффективности. Это проверенные "победители" — контентные направления, которые стабильно показывают высокий резонанс. Рекомендации должны быть конкретными и основываться на комплексном анализе. Например: "Тема 'Продвинутое создание виджетов' имеет самый высокий показатель WES1k. Рекомендуется создать серию из трех частей, посвященную подтемам X, Y и Z, которые были определены как часто задаваемые вопросы в ходе анализа комментариев к видео этого кластера". Такой подход не просто предлагает "делать больше того же", а использует данные из комментариев для углубления и расширения самой успешной темы.

### **3.2. "Зона Экспериментов": Обнаружение Скрытых Возможностей**

Здесь внимание уделяется темам, которые могут иметь невысокие средние просмотры, но исключительно высокий WES1k. Это нишевые темы, которые находят отклик у очень страстной и вовлеченной части аудитории.

Обычный продюсер, видя видео с низким количеством просмотров, может счесть его неудачей. Однако метрика WES может рассказать другую историю. Видео с 10 000 просмотров, но с показателем WES1k в три раза выше среднего по каналу, указывает на существование гипервовлеченной, но недостаточно обслуживаемой аудитории. Это не провал, а успешный пилотный выпуск для новой, потенциально очень лояльной контентной вертикали. Рекомендация в данном случае — не просто "сделать еще одно такое видео", а "рассматривать эту нишу как основу для потенциального спин-офф сериала или эксклюзивного контента для платных подписчиков". Это меняет понимание успеха с чистого охвата на глубину аудитории и потенциал монетизации.

Также в этом разделе анализируются видео-выбросы, идентифицированные HDBSCAN. Если такое "аномальное" видео демонстрирует высокий WES, это является сильным сигналом об удачном эксперименте, который можно развить в новую полноценную тему.

### **3.3. "Список Деприоритизации": Стратегическое Распределение Ресурсов**

В этом разделе анализируются темы из нижнего квартиля Тепловой Карты — контент, который стабильно не может сгенерировать ни просмотры, ни глубокую вовлеченность. Рекомендация здесь — не "немедленно прекратить", а "деприоритизировать инвестиции". Такой подход, подкрепленный данными, позволяет автору уверенно перераспределять время и производственный бюджет с низкоэффективного контента на высокоэффективные направления из "Столпов Усиления" и "Зоны Экспериментов".

## Раздел 4: Совершенствование Формата, Тональности и Подачи

### 4.1. Определение Фирменного Стиля Автора

Для анализа стиля и тональности необходимо изолировать транскрипты из самых эффективных тематических кластеров ("Столпы Усиления"). Эти тексты будут проанализированы LLM с использованием специализированного промпта, нацеленного на стилистический и тональный анализ.

**Шаблон Промпта: Персона "Лингвистический Аналитик"**

**\*\*Персона:\*\*** Ты — вычислительный лингвист и эксперт по стилистике. Ты анализируешь тексты, чтобы деконструировать уникальный авторский голос.

**\*\*Контекст:\*\*** Тебе дан транскрипт очень успешного видео с YouTube. Твоя цель — проанализировать его лингвистические свойства, чтобы создать руководство по стилю для автора.

**\*\*Задача:\*\*** Проанализируй предоставленный транскрипт и оцени следующие атрибуты по шкале от 1 до 5 (1=Низкий, 5=Высокий). Для каждого атрибута предоставь одно предложение с обоснованием и примерами из текста.

1. **\*\*Формальность:\*\*** (1=Очень неформальный, 5=Очень формальный)
2. **\*\*Энергия/Темп:\*\*** (1=Медленный и размеренный, 5=Быстрый и энергичный)
3. **\*\*Юмор:\*\*** (1=Серьезный/Сухой, 5=Частые шутки/Остроумие)
4. **\*\*Лексическая плотность:\*\*** (1=Простой язык, 5=Сложный/Насыщенный жаргоном)
5. **\*\*Ясность инструкций:\*\*** (1=Абстрактный/Концептуальный, 5=Четкий, пошаговый)
6. **\*\*Связь с аудиторией:\*\*** (например, использование "вы", "мы", риторические вопросы).

**\*\*Формат:\*\*** Ответь JSON-объектом, где каждый ключ — это атрибут (например, "Formality"), а значение — другой объект, содержащий "score" и "justification".

Этот промпт предоставляет количественную и качественную разбивку наиболее успешного стиля коммуникации автора, позволяя объективно оценить его ключевые характеристики.<sup>17</sup>

### 4.2. Создание Руководства по Тональности и Стилю



Усреднив оценки из лингвистического анализа по всем высокоэффективным видео, можно создать окончательное руководство по стилю.

#### **Ключевой Инструмент: Руководство по Фирменному Стилю Автора**

Авторы часто развивают свой стиль интуитивно. Эта таблица делает неявное явным. Она кодифицирует их "магическую формулу", позволяя быть более последовательными, обучать других авторов или редакторов и сознательно изменять свой стиль при экспериментах с новыми форматами. Она дает структурированный и основанный на данных ответ на вопрос "какая тональность автора работает, а какая нет".

Атрибут	Оптимальная оценка (среднее по топ-видео)	Описание и лучшие практики
Формальность	2.5	Используйте разговорный, но уважительный тон. Избегайте как излишнего официоза, так и фамильярности.
Энергия/Темп	4.2	Поддерживайте высокий темп повествования с короткими, динамичными предложениями. Используйте интонационные акценты.
Юмор	3.8	Включайте ситуативный юмор и остроумные замечания, связанные с темой. Избегайте отвлеченных шуток.
Лексическая плотность	3.0	Объясняйте сложные термины простыми словами, но не бойтесь использовать профессиональную лексику, когда это уместно.

### **4.3. Корреляция Формата с Успехом**

Анализ метаданных (описаний видео, названий) высокоэффективных по WES видео позволяет выявить ключевые слова, связанные с форматом (например, "Тutorial", "Обзор", "Интервью", "Прямой эфир"). Этот анализ определяет, какие форматы подачи материала наиболее тесно коррелируют с высокой вовлеченностью, что приводит к четким рекомендациям, таким как: "Видео в формате тьюториала продолжительностью до 15 минут стабильно превосходят другие форматы по уровню вовлеченности".

## Раздел 5: Развитие Связи с Сообществом

### 5.1. От Комментариев к Контенту: Маховик Обратной Связи

Этот раздел использует результаты анализа кластеров комментариев аудитории из Раздела 1.5. Превращение этих данных в контент-идеи замыкает цикл обратной связи и демонстрирует аудитории, что ее голос услышан.<sup>20</sup>

#### Действенные выводы:

- **Кластер "Часто задаваемые вопросы":** Прямая основа для создания отдельного FAQ-видео или серии "Разрушители мифов".
- **Кластер "Запросы на новые функции":** Если канал посвящен продукту или услуге, это готовая дорожная карта для разработки.
- **Кластер "Распространенная критика":** Уникальная возможность создать видео, которое прямо и прозрачно отвечает на опасения аудитории, что способствует укреплению доверия.
- **Кластер "Внутренние шутки/Мемы":** Использование этих элементов в контенте усиливает связь с сообществом и дает лояльным зрителям почувствовать себя "своими".

### 5.2. Побуждение к Ценному Взаимодействию

На основе проведенного анализа можно рекомендовать автору конкретные призывы к действию (Call-to-Action, CTA) для использования в видео. Вместо общего "Оставляйте комментарии" CTA может быть адаптирован: "Данные показывают, что наши самые вовлекающие видео — это tutorиалы. Напишите в комментариях, какую именно технику из этого видео вы собираетесь попробовать в первую очередь". Это стимулирует более качественные, тематические комментарии, что повышает WES и предоставляет более ценные данные для будущего анализа.

## Часть III: Продюсерская Версия — Продвинутые Стратегические Горизонты

Этот заключительный раздел представляет экспертные выводы, которые отличают стандартный отчет от высокоценной стратегической консультации. Он фокусируется на динамическом анализе и долгосрочном здоровье канала.

## Раздел 6: Выявление Скрытых Возможностей и Рисков

### 6.1. Временная Динамика: Эволюция и Усталость Тем

Построив график среднего WES1k для каждой темы во времени (по дате публикации видео), можно получить динамическую картину интереса аудитории.

**Потенциальные выводы:**

- **Усталость от темы:** Некогда популярная тема показывает со временем снижение вовлеченности, что сигнализирует о необходимости инноваций или смены фокуса.
- **Возникающие тренды:** Внезапно появляется новая тема, которая быстро набирает популярность.
- **Сезонность:** Определенные темы лучше "заходят" в конкретное время года.

Этот анализ превращает статичный снимок состояния канала в динамическое, продольное представление о его здоровье.

### 6.2. Анализ Траектории Тональности

Вместо того чтобы присваивать один балл тональности каждому комментарию, можно использовать LLM для анализа тональности *ответов* в ветке комментариев.

Стандартный анализ тональности говорит, является ли комментарий положительным или отрицательным. Анализ траектории тональности показывает, становится ли разговор *лучше* или *хуже*. Отфильтровав ветки, в которых отвечает сам автор, можно измерить его влияние на дискуссию. Способствует ли вмешательство автора деэскалации негативных веток, переводя их в нейтральное или позитивное русло? Провоцируют ли его ответы в позитивных ветках еще большую волну поддержки? Это прямой, измеримый показатель навыков автора в управлении сообществом и его способности формировать культуру канала. Это метрика здоровья самих отношений между автором и аудиторией.

### 6.3. Картирование Экосистемы: Сотрудничество и Позиционирование

Программное извлечение и анализ всех исходящих ссылок из описания канала и описаний видео позволяет составить карту "экосистемы", в которой существует автор. Какие другие каналы, инструменты или веб-сайты он часто упоминает? Эта карта может выявить "смежных" авторов для потенциальных коллабораций или подсветить пробелы, где автор мог бы укрепить свой авторитет, ссылаясь на ценные, но малоизвестные ресурсы.

## Раздел 7: Система Непрерывного Совершенствования

### 7.1. Концептуальная Панель Управления Автора

На основе предложенной методологии можно спроектировать простую, но мощную панель управления (дашборд), которую автор мог бы использовать для мониторинга здоровья своего канала.

**Ключевые виджеты:**

- **Топ-5 тем по эффективности** (по WES1k за последние 90 дней).
- **Новые темы в комментариях** (новые кластеры, появляющиеся в обсуждениях).
- **Тренд общей тональности аудитории** (график изменения во времени).
- **Оповещения о видео-выбросах** (новые видео, помеченные HDBSCAN как "шум", для ручного анализа).

### 7.2. Предстартовая Оценка Контента: От Реактивного к Проактивному

Конечная цель этого фреймворка — перейти от анализа прошлого к прогнозированию будущего. Можно описать процесс, при котором автор берет сценарий или подробный план нового видео, генерирует для него векторный эмбединг и использует базу данных Quadrant для поиска его ближайших "соседей" среди уже опубликованных видео. После этого автор может увидеть средний WES1k и тональность комментариев для наиболее тематически схожих видео, которые он уже выпускал. Это дает основанную на данных "оценку потенциала" еще до начала съемок, позволяя доработать идеи, которые, по прогнозам, покажут низкую эффективность, и, наоборот, сделать ставку на те, что с высокой вероятностью найдут сильный отклик. Такой подход превращает всю аналитическую систему из инструмента ретроспективной отчетности в проактивного творческого партнера.

#### Источники

1. 3. Clustering - BERTopic - Maarten Grootendorst, дата последнего обращения: августа 23, 2025, [https://maartengr.github.io/BERTopic/getting\\_started/clustering/clustering.html](https://maartengr.github.io/BERTopic/getting_started/clustering/clustering.html)
2. Comparing Clustering Algorithms - Hacker News, дата последнего обращения: августа 23, 2025, <https://news.ycombinator.com/item?id=11609882>
3. [P] Clustering approach for multi-dimensional vectors : r/MachineLearning - Reddit, дата последнего обращения: августа 23, 2025, [https://www.reddit.com/r/MachineLearning/comments/15alpxe/p\\_clustering\\_approach\\_for\\_multidimensional\\_vectors/](https://www.reddit.com/r/MachineLearning/comments/15alpxe/p_clustering_approach_for_multidimensional_vectors/)
4. Unsupervised clustering of transformers-derived embeddings -- what clustering

and visualization algorithms to try after k-means + PCA, and is it just HDBSCAN + UMAP these days? : r/datascience - Reddit, дата последнего обращения: августа 23, 2025,

[https://www.reddit.com/r/datascience/comments/1elx6r6/unsupervised\\_clustering\\_of\\_transformersderived/](https://www.reddit.com/r/datascience/comments/1elx6r6/unsupervised_clustering_of_transformersderived/)

5. Mastering Data Clustering with Embedding Models | Towards Dev - Medium, дата последнего обращения: августа 23, 2025,  
<https://medium.com/towardsdev/mastering-data-clustering-with-embedding-models-87a228d67405>
6. HDBSCAN Clustering with Milvus, дата последнего обращения: августа 23, 2025, [https://milvus.io/docs/hdbscan\\_clustering\\_with\\_milvus.md](https://milvus.io/docs/hdbscan_clustering_with_milvus.md)
7. 2.3. Clustering — scikit-learn 1.7.1 documentation, дата последнего обращения: августа 23, 2025, <https://scikit-learn.org/stable/modules/clustering.html>
8. Comparing The-State-of-The-Art Clustering Algorithms | by Sina Nazeri | Medium, дата последнего обращения: августа 23, 2025,  
<https://medium.com/@sina.nazeri/comparing-the-state-of-the-art-clustering-algorithms-1e65a08157a1>
9. Comparing Python Clustering Algorithms — hdbscan 0.8.1 documentation, дата последнего обращения: августа 23, 2025,  
[https://hdbscan.readthedocs.io/en/latest/comparing\\_clustering\\_algorithms.html](https://hdbscan.readthedocs.io/en/latest/comparing_clustering_algorithms.html)
10. Topic Modelling vs Clustering - When & Why? : r/MLQuestions - Reddit, дата последнего обращения: августа 23, 2025,  
[https://www.reddit.com/r/MLQuestions/comments/yzgw44/topic\\_modelling\\_vs\\_clustering\\_when\\_why/](https://www.reddit.com/r/MLQuestions/comments/yzgw44/topic_modelling_vs_clustering_when_why/)
11. The Simple Prompt I Use to Get Meaningful Responses From LLMs | by Fadi Boulos, дата последнего обращения: августа 23, 2025,  
<https://medium.com/@fadiboulos/the-simple-prompt-i-use-to-get-meaningful-responses-from-llms-1aa6ec0c8e7a>
12. Mastering role prompting: How to get the best responses from LLMs - Portkey, дата последнего обращения: августа 23, 2025,  
<https://portkey.ai/blog/role-prompting-for-llms>
13. What is the correct formula to measure engagement rate on any social media platform?, дата последнего обращения: августа 23, 2025,  
[https://www.reddit.com/r/socialmedia/comments/vytx5c/what\\_is\\_the\\_correct\\_formula\\_to\\_measure\\_engagement/](https://www.reddit.com/r/socialmedia/comments/vytx5c/what_is_the_correct_formula_to_measure_engagement/)
14. Descriptive statistics of weighted engagement (WE) according to content type - ResearchGate, дата последнего обращения: августа 23, 2025,  
[https://www.researchgate.net/figure/Descriptive-statistics-of-weighted-engagement-WE-according-to-content-type\\_tbl1\\_339253551](https://www.researchgate.net/figure/Descriptive-statistics-of-weighted-engagement-WE-according-to-content-type_tbl1_339253551)
15. Weighted Engagement - Fanpage Karma Academy, дата последнего обращения: августа 23, 2025,  
[https://academy.fanpagekarma.com/en/hrf\\_faq/weighted-engagement/](https://academy.fanpagekarma.com/en/hrf_faq/weighted-engagement/)
16. How To Calculate Ad Engagement Rate - growth-onomics, дата последнего обращения: августа 23, 2025,  
<https://growth-onomics.com/how-to-calculate-ad-engagement-rate/>

17. Tone Of Voice & Style Prompt Descriptions For Better AI Outputs - Word.Studio, дата последнего обращения: августа 23, 2025, <https://word.studio/tone-of-voice-style-prompts-descriptions/>
18. Controlling Tone and Style With Prompt Engineering for AI Assistants, дата последнего обращения: августа 23, 2025, <https://promptlearnings.com/controlling-tone-and-style-with-prompt-engineering/>
19. I've been tweaking ChatGPT's writing style for specific tasks lately. If you have a go-to writing task (like weekly emails or blog posts), comment below and I'll share a system prompt to help ChatGPT stick to a consistent tone/style each time you write. : r/PromptEngineering - Reddit, дата последнего обращения: августа 23, 2025, [https://www.reddit.com/r/PromptEngineering/comments/1iapund/ive\\_been\\_tweaking\\_chatgpts\\_writing\\_style\\_for/](https://www.reddit.com/r/PromptEngineering/comments/1iapund/ive_been_tweaking_chatgpts_writing_style_for/)
20. YouTube Comment Analysis with LLMs | Kilian Merkelbach, дата последнего обращения: августа 23, 2025, <https://www.kmerkelbach.com/blog/2025/yt-comments/>
21. 10 ChatGPT Prompts For Analyzing Customer Feedback | by Shushant Lakhyani | Medium, дата последнего обращения: августа 23, 2025, <https://medium.com/@slakhyani20/10-chatgpt-prompts-for-analyzing-customer-feedback-c00a5f0d0795>