

МЕТОДИКА ИССЛЕДОВАНИЯ МАРКЕТИНГОВЫХ ДАННЫХ В ЦИФРОВОЙ СРЕДЕ ДЛЯ СОВЕРШЕНСТВОВАНИЯ КОНТЕНТ-СТРАТЕГИИ КОМПАНИИ

Диджитализация всех сфер жизни потребителя требует от бизнеса соразмерных ответов. Чем большее количество стадий принятия решения о покупке совершаются в диджитал среде, тем более критичным для компаний становится вопрос разработки полноценной стратегии представленности в цифровой среде. Первым этапом разработки подобной стратегии неизбежно должен выступить анализ текущего положения компании в цифровой среде, для осуществления которого необходима новая методология маркетинговых исследований, базирующаяся на новых методах сбора, обработки и анализа маркетинговой информации. «Старые», традиционные методы маркетинговых исследований, такие как эмпирические исследования, анализ CRM-систем и т.д. плохо подходят для решения данной задачи.

На рис. 1 представлена авторская классификация возможных источников маркетинговых данных в цифровой среде.



Рис.1. Возможные источники маркетинговых данных в цифровой среде
Источник: составлено авторами

Для извлечения маркетинговых данных подобного типа необходимо использовать новый метод, отсутствующий в классической методологии маркетинговых исследований – метод парсинга цифровой среды. **Парсинг** представляет собой процесс автоматизированного извлечения информации из цифровой среды на базе кода, написанного на одном из языков программирования (например, Python).

Маркетинговые данные в цифровой среде возможно проклассифицировать с точки зрения степени их структурированности (см. рис. 2).

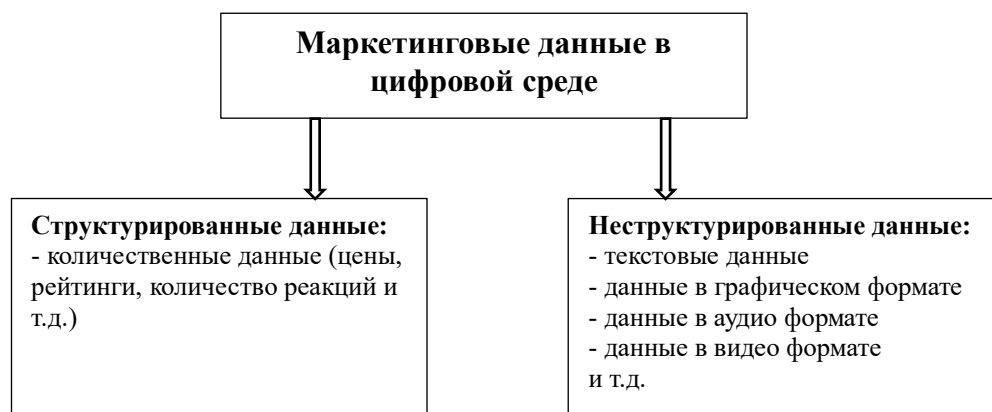


Рис. 2. Классификация маркетинговых данных в цифровой среде
Источник: составлено авторами

Как правило, неструктурированные маркетинговые данные в цифровой среде представлены в виде естественной цифровой информации (ЕЦИ). **Естественная цифровая информация** представляет собой неструктурированный массив данных, являющийся результатом вовлеченности пользователя во взаимодействие с цифровой средой. В рамках данного исследования нас интересует та часть естественной цифровой информации, которая касается взаимодействия пользователя с компанией в цифровой среде, примерами чего могут быть:

- отзывы потребителей на товары / услуги (в карточках на маркетплейсах, в геолокационных сервисах, на сайтах-отзовиках, в социальных сетях и т.д.);
- комментарии пользователей к публикациям компании в социальных сетях;
- дискуссии, обсуждения на тематических форумах, касающиеся бренда, а также ответы на комментарии других пользователей и т.д.

Ключевой особенностью ЕЦИ является то, что она находится в открытом доступе и может быть свободно исследована.

Что касается дальнейшей процедуры действий со спарсенными данными, то методы их обработки и анализа логично разделить с точки зрения типа данных: качественные, количественные либо смешанные. Количественными, например, могут быть статистические данные о состоянии рынков и отраслей, цены конкурентов на маркетплейсах и их динамика, рейтинги товаров, параметры вовлеченности целевой аудитории во взаимодействие с брендом и т.д. Методы обработки и анализа таких данных относятся к области data mining и зависят от их объема, разнообразия, степени структурированности и скорости приращения [3], то есть по сути относятся ли подобные спарсенные данные к big data и если да, к какой из ее составляющих («Fast data», «Big analytics», «Deep insight» или др.) [2].

Однако наибольший интерес в контексте данного исследования представляют качественные данные, извлеченные из цифровой среды, то есть представленные в текстовом, аудио, видео и др. подобных форматах. Методы обработки и анализа подобных данных лежат в области компьютерной лингвистики, анализа изображений и схожих областей, то есть дополняют традиционные методы data mining количественных маркетинговых данных. Среди популярных методов обработки естественной цифровой информации можно выделить кластеризацию текстовых массивов, поиск закономерностей, анализ аномалий и выявление выбросов в данных, оценку тематической, эмоциональной и контекстуальной составляющих ЕЦИ. Также стоит отметить роль нейросетевого подхода в анализе естественной цифровой информации, в частности, такие методы как автоматизированный поиск тематического контента, анализ плотности вероятности присутствия тематических сигналов [1], анализ тональности реакций целевых аудиторий, методы семантического моделирования и векторного представления слов [4], составление

графов тематических связей и взаимовлияний текстовой информации и оценка их структурной сложности [5] и т.д.

Результаты исследования цифровой среды предприятия могут быть использованы при принятии и реализации маркетинговых решений, например, формировании спроса через формирование информационной повестки, адаптации ассортимента и разработки упаковки, совершенствовании стратегий позиционирования, контент-стратегий и стратегий комьюнити-менеджмента, настройки инструментов привлечения платного трафика и мн. др.

В рамках данной статьи хотелось бы продемонстрировать пример методики исследования цифровой среды с целью совершенствования контент-стратегии компании в социальных сетях (см. рис. 3).

Сбор данных		Обработка и анализ данных	Использование данных
Источник данных	Объект парсинга		
Новостная повестка (публикации в СМИ)	News API Google News API и аналоги	Выявление самых часто встречающихся слов Кластеризация тематик Поиск взаимовлияний между новостными тематиками Выявление «хайповых» и «нехайповых» тем во временном разрезе Оценка эмоциональной и контекстуальной составляющих упоминаний компании и конкурентов	Корректировка матрицы контента Определение частоты постинга и составление календарного графика
Запросы потребителей в поисковых системах	Яндекс Wordstat и Google Trends	Выявление самых часто встречающихся слов Кластеризация тематик Поиск закономерностей между поисковыми запросами (в т.ч. во временном разрезе)	
Отзывы потребителей на независимых площадках	Например, Яндекс карты, Google Maps и др.	Выявление самых часто встречающихся слов Кластеризация тематик и анализ плотности присутствия тематических сигналов (например, для выявления основных проблем и жалоб) Оценка тональности отзывов Комментарии на отзывы со стороны других пользователей	
Социальные сети компании	Официальные и (при необходимости) неофициальные сообщества ВК, Телеграм и др.	Выявление самых часто встречающихся слов Кластеризация тематик Анализ реакций подписчиков (колич. данные) Контент комментариев на публикации со стороны подписчиков + Контент комментариев на комментарии: анализ плотности присутствия тематических сигналов (например, для выявления основных проблем и жалоб) Оценка тональности комментариев	
Социальные сети конкурентов	Официальные и (при необходимости) неофициальные сообщества ВК, Телеграм и др.	Выявление самых часто встречающихся слов Кластеризация тематик Оценка тональности публикаций и комментариев к ним Анализ реакций подписчиков (колич. данные) Контент комментариев на публикации со стороны подписчиков + Контент комментариев на комментарии: анализ плотности присутствия тематических сигналов (например, для выявления основных проблем и жалоб)	

Рис. 3. Методика исследования цифровой среды с целью совершенствования контент-стратегии компании в социальных сетях

Источник: составлено авторами

Представленная на рис. 3 методика исследования цифровой среды может быть использована для выявления аномалий в различных источниках данных (что может, например, помочь выявить «купленные» отзывы и «накрученные» посты конкурентов), для составления графов тематических связей между разными источниками данных, для составления разговорного словаря потребителей, для депозиционирования от конкурентов по тематике контента в социальных сетях, что в конечном итоге будет представлять собой полноценную стратегию совершенствования (корректировки) контента компании в социальных сетях как реакции на изменение внешней среды.

По результатам уже проведенной корректировки контент-стратегии компании имеет смысл проанализировать динамику матрицы посевов. Данный шаг будет являться первичным для автоматизации данной методики, так как анализ изменений частоты и содержания реакций целевой аудитории на скорректированную контент стратегию компании может запустить цикл ее автоматической корректировки, также реализованный через программный код на Python.

Развитие данной методики лежит в области подключения других источников маркетинговых данных, в частности, касающихся целевой аудитории, позиционирования, tone-of-voice и других бренд-идентификаторов, что может в итоге создать полноценный полностью автоматизированный инструмент анализа текущей контент-стратегии компании и ее совершенствования на базе извлечения (парсинга) маркетинговых данных из цифровой среды.

Список использованных источников:

1. Chen, J., Zhang, A., Li, M., Smola, A., Yang, D. (2023) A cheaper and better diffusion language model with soft-masked noise - arXiv preprint arXiv:2304.04746.
2. Hinchcliffe, D. Big Data, The Moving Parts: Fast Data, Big Analytics, and Deep Insight. Available at: <https://www.flickr.com/photos/dionh/7550578346/in/photostream/> (accessed 20.01.2025)
3. Kitchin, R., McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*. 3 (1): 205395171663113.
4. Knott, A., Pedreschi, D., Jitsuzumi, T., Leavy, S., Eysers, D. (2024) AI content detection in the emerging information ecosystem: new obligations for media and tech companies. *Ethics and information technology*.
5. Tufte, E.R. (1998) *Envisioning Information*. Graphics Press, Cheshire, Connecticut. Sixth printing.