

Автоматизированный анализ репрезентации товара в цифровой рекламе образовательных услуг

А. Б. Углова, И. М. Богдановская, Б. А. Низомутдинов

Аннотация— В статье рассматриваются структурные и смысловые характеристики репрезентации товаров в цифровой рекламе образовательных услуг помогающих специалистов в социальных сетях (на примере социальной сети «ВКонтакте»). Исследуются возможности автоматизированных методов анализа в выявлении вербальных маркеров репрезентаций профессионального инструментария для дальнейшего создания систем мониторинга и этической оценки информации об электронных товарах. Разработана общая схема автоматизированного сбора информации об образовательных товарах и обработки полученных данных по тематическим сообществам (педагоги, психологи, врачи). Для тематического моделирования текстовых данных использовался метод LDA (Latent Dirichlet Allocation). С помощью метода TF-IDF (Term Frequency-Inverse Document Frequency) были выделены важные термины в каждом наборе описаний для анализа смысловых фреймов рекламных объявлений. Выявлены вербальные маркеры с опорой на которые авторы рекламных текстов внедряют в сознание читателя узловые элементы сценария покупки профессионального обучения для специалиста. Кластерная структура репрезентаций педагогических товаров включает узкоспециализированные описания, предназначенные для специалистов, работающих в общеобразовательных учреждениях; для репрезентаций психологических товаров характерна ориентация на широкую аудиторию, включающую неспециалистов, тематически смешанная структура; в репрезентациях медицинских товаров основная информация не связана с медициной, что может ввести пользователей в заблуждение. Методика автоматизированного кластерного анализа маркетинговых стратегий, продающих образовательный контент будет востребована широким кругом специалистов в сфере менеджмента образования, разработчиками контента образовательных интернет-каналов, образовательными учреждениями, педагогами, работающими в виртуальной среде.

Ключевые слова — социальные сети, парсинг, LDA, помогающие специалисты, образовательные товары, профессиональный инструментарий, вербальные маркеры.

I. ВВЕДЕНИЕ

Социальная сеть ВКонтакте самая большая и популярная социальная сеть в России, один из самых

посещаемых сайтов Рунета. Широкая аудитория, привлеченная удобством интерфейса, обширными социальными итерациями, разнообразным контентом, открыто пользуется всеми дополнительными возможностями сети, в том числе и получает информацию об электронных товарах для развлечения, заработка и обучения. Использование электронных товаров, связанных с обучением, можно рассматривать как получение нематериального блага, направленного на удовлетворение познавательных потребностей пользователя в интернет-среде. Электронные образовательные товары ускоряют коммуникацию, дают возможность получения образования в любой точке мира. По этой причине комплексными цифровыми площадками для взаимодействия с потенциальными потребителями таких услуг становятся социальные сети. Реклама товаров/услуг в социальных сетях позволяет пользователю быстро находить нужную информацию, дает возможность обратиться непосредственно к производителю, может включать ссылки на другие ресурсы, распространяться через подписчиков и т. п. Таким образом, структурная организация информации об образовательных товарах в социальных сетях во многом будет определять интерес пользователей. С другой стороны - важным является и содержательный аспект информации об электронных образовательных товарах, которые должны отвечать потребностям обучающихся в профессиональном саморазвитии. Тем не менее, существуют значительные ограничения в мониторинге качества информации, которая репрезентирует как сами товары, так и поставщика предоставляемых образовательных услуг [1]. В этой связи настоящее исследование направлено на анализ структурно-семантических характеристик информации об образовательных товарах помогающих специалистов в социальных сетях с помощью методов машинного обучения. Методика автоматизированного анализа информации об образовательных товарах будет востребована широким кругом специалистов в сфере менеджмента образования, разработчиками контента образовательных интернет-каналов, образовательными учреждениями, педагогами, работающими в виртуальной среде.

Изучение приемов распространения различных продуктов через социальные сети ведется в нескольких направлениях. В рамках социально-психологического дискурса проводятся широкомасштабные исследования мнений о формировании лидерства брендов в

Статья получена 20.04.2024 г.

А.Б. Углова, РГПУ им. А. И. Герцена (anna.uglova@list.ru)

И. М. Богдановская, РГПУ им. А. И. Герцена (ibogdanovs@herzen.spb.ru)

Б. А. Низомутдинов, Университет ИТМО (boris@itmo.ru)

социальных сетях для продвижения новых продуктов. Например, в работе N. Wang et al. описывается механизм влияния мнения отзывов о принятии решений другими потребителями [2]. Можно говорить о том, что популярные товары сами по себе становятся лидерами мнения, которые определяют тренд в освоении профессионального инструментария. Н. Zhang et al. указывают в своем исследовании, что на интерес к продуктам в социальных сетях сильно влияет содержание описания товара: чем больше в описании товара профессиональной лексики, указывающей на экспертность автора, тем более надежной и полезной воспринимается информация [3].

Также интересной является теория «гравитационного поля», которая говорит о том, что потребители способны создать внутренние волнообразные изменения мнений в социальных сетях, изменяя уровень доверия к продуктам [4]. В отсутствие внешних источников официальной информации, которая могла бы помочь определить ликвидные образовательные курсы, информационный шум «раскрученных» товаров осложняет поиск для пользователей.

Ряд работ посвящен модернизации процесса внедрения продуктов на основе технологий нейронных сетей. L. Ottikunta в своей работе описывает модель анализа информации о социальных итерациях в социальных сетях и возможности использования социального контента для облегчения рекомендаций высокого качества [5].

N. Gozuacik et al. в своей работе показывают возможности использования нейронных сетей для анализа негативных и позитивных отзывов с целью выявления проблем, связанных с продуктом и нахождения инновационных идей [6]. X. Wang et al. в исследовании описывают модель оценки интенсивности спроса и поиска значимых характеристик продуктов на основе машинного обучения [7].

Стоит отметить, что большая часть исследований, посвященных изучению цифровых товаров направлена на реализацию маркетинговых технологий, связанных с возможностью повлиять на поведение потребителя при покупке разных товаров [8]. В то же время, работ описывающих этическое представление полезных для пользователей товаров практически нет.

Ряд работ описывает возможность реализации этических принципов, разработанных на основе искусственного интеллекта, в репрезентации медицинских товаров. M. Goigand et al. указывают на необходимость выстраивания открытого диалога с клиентами в момент принятия решения, который позволяет пользователю самостоятельно оценить качество услуги [9].

A. Anand et al. в своей работе рассматривают этический аспект построения виртуального маркетинга. Авторы указывают на нарушения этических кодексов поведения различными организациями, направленных на достижение целевых показателей доходов [10]. N. Toledano et al. в своих работах оценивают этический аспект создания некоммерческих образовательных проектов. Авторы указывают, что использование

специальной академической лексики, повышает ценность товаров и делает их более привлекательными. [11].

В связи с вышесказанным *цель* нашего исследования состояла в выявлении структурно-семантических характеристик репрезентации товара в цифровой рекламе образовательных услуг (на примере социальной сети «ВКонтакте»). Не менее важным аспектом стало исследование возможностей автоматизированных методов анализа в выявлении вербальных маркеров репрезентаций образовательных услуг помогающих специалистам в социальных сетях для дальнейшего создания систем мониторинга и этической оценки информации об электронных товарах.

II. СБОР И ОБРАБОТКА ИНФОРМАЦИИ

Раздел «Товары» в сообществах ВКонтакте - это функционал, который позволяет администраторам сообществ создавать каталоги товаров и продавать их прямо в социальной сети. Пользователи могут просматривать товары, добавлять их в корзину и оформлять покупку, не покидая страницу сообщества. Таким образом, товары в сообществах ВКонтакте предоставляют возможность удобного онлайн-шопинга для пользователей и расширяют возможности монетизации для администраторов сообществ.

При проведении данного исследования была поставлена цель - собрать профессиональные помогающие сообщества, автоматизированным методом, выгрузить все товары из выбранных сообществ и провести текстовую аналитику собранных данных, для тематического моделирования и выделения ключевых терминов описания.

Для парсинга описания товаров в ВКонтакте с помощью VK API можно использовать метод "market.get" из раздела "Товары" (market) API. Формат запроса к API будет выглядеть так: https://api.vk.com/method/market.get?owner_id=-123456&count=100&extended=1&access_token=YOUR_ACCESS_TOKEN&v=5.131

где:

- owner_id - идентификатор сообщества, товары которого нужно получить (если сообщество имеет id 123456, то owner_id будет равен -123456);
- count - количество товаров, которые нужно получить;
- extended - флаг, указывающий на необходимость получения дополнительной информации о товарах, включая описание;
- access_token - токен доступа к API Вконтакте;
- v - версия API.

После отправки запроса данные будут получены в формате JSON. В ответе будут содержаться объекты с информацией о каждом товаре, включая его описание, название, цену и другие характеристики.

Таким образом, с помощью метода "market.get" можно получить описание товаров в сообществе ВКонтакте вместе с другой полезной информацией о них.

Результатом сборки стал массив файлов JSON, содержащий описание товаров для 3 групп специалистов - врачи, психологи и педагоги.

На следующем этапе потребовалась обработка тестов методами, основанными на машинном обучении. Для выделения ключевых маркеров репрезентации образовательных услуг был проведен тематический анализ, выделены ключевые термины и составлено облако ключевых слов.

Общая схема сбора и проведения анализа текстовой информации приведена на рисунке 1.



Рис. 1. Организационная схема проведения сбора и обработки информации

III. АНАЛИЗ СОБРАННЫХ ДАННЫХ И ТЕКСТОВАЯ АНАЛИТИКА

Для тематического моделирования был выбран метод LDA. Метод LDA (Latent Dirichlet Allocation) — это вероятностная модель тематического моделирования, которая используется для извлечения тем из коллекции текстовых документов. Он основан на предположении о том, что каждый документ может быть представлен как смесь различных тем, а каждая тема может быть представлена как распределение вероятностей по словам.

Метод LDA обладает рядом преимуществ по сравнению с другими методами для тематического моделирования по следующим причинам:

- Учитывает смешивание тем: LDA предполагает,

что каждый документ содержит несколько тем, что соответствует реальной природе текстовых данных, где документы могут быть связаны с несколькими темами одновременно.

- Модель LDA является генеративной: позволяет генерировать новые документы на основе изученных тем и их распределений.
- Вероятностный подход: метод LDA основан на вероятностной модели (байесовская статистика и распределение Дирихле), что делает его более устойчивым к шуму и позволяет учитывать неопределенность в данных. В рамках модели LDA каждый документ представляется как совокупность некоторого количества тем, а появление каждого слова связано с какой-либо из тем документа.

Таким образом, метод LDA является мощным инструментом для извлечения тематической структуры из текстовых данных и широко используется в области обработки естественного языка и анализа текстов.

В ходе анализа было проведено тематическое моделирование для каждого набора данных, итогом стали 3 файла, содержащие по 15 тем для каждого направления, со списком ключевых слов. С помощью метода MDS проведена визуализация данных. MDS (Multidimensional Scaling) — это метод визуализации данных, который может быть применен к результатам тематического моделирования для визуального представления тематической структуры коллекции текстовых документов. MDS позволяет отобразить многомерные данные (например, расстояния между темами) на двух- или трехмерное пространство с сохранением относительных расстояний между данными. Пример визуализации приведен на Рисунке 2.

Размер темы говорит о предельной вероятности, которая подсчитывается для всех тем. Этот параметр сообщает о предельной вероятности темы в корпусе (насколько сильно представлена тема в корпусе).

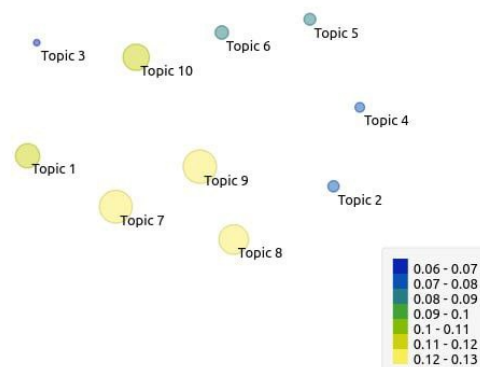


Рис. 2. Визуализации выделенных тем по LDA с помощью метода MDS

Следующим шагом стало выделение важных терминов в каждом наборе описаний. Для этого был задействован метод TF-IDF. Метод TF-IDF (Term Frequency-Inverse Document Frequency) — это статистическая мера, используемая в обработке

естественного языка для оценки важности терминов (слов) в документе относительно коллекции документов.

TF-IDF применяется для выявления ключевых слов или терминов в документах и может быть использован, например, для кластеризации документов или классификации текстов.

Важно отметить, что ключевые слова и важные термины имеют различие. Ключевые слова и важные термины в TF-IDF имеют схожие цели - выделение важной информации из текста, но подходят к этой задаче с разных точек зрения: ключевые слова обычно выбираются на основе содержания текста, в то время как важные термины в TF-IDF определяются статистической значимостью в контексте коллекции документов.

IV. АНАЛИЗ ПОЛУЧЕННЫХ ДАННЫХ

Выгруженные с помощью TF-IDF ключевые слова, составляют собой базовый смысловой фрейм, позволяющий составить рекламный текст репрезентирующий товар с использованием знакомых, стереотипных слов из семантического поля профессии и вызвать доверие у читателя [12].

Рассмотрим вербальную структуру рекламных текстов для образовательных товаров по психологии, педагогике и медицине, для выявления вербальных маркеров образовательных услуг помогающих специалистам в социальных сетях (таблица 1).

Таблица 1. Содержание и структура смыслового фрейма текстов-репрезентаций, выделенных на основе метода TF-IDF

Товары по психологии		Товары по педагогике		Товары по медицине	
Ключевые слова	Вес	Ключевые слова	Вес	Ключевые слова	Вес
Курс	0,039	Курс	0,047	Обучение	0,072
Занятие	0,035	Образование	0,040	Массаж	0,058
Психология	0,028	Профессиональный	0,038	Кожа	0,032
Психолог	0,024	Обучение	0,035	Проводить	0,032
Рубль	0,021	Программа	0,028	Косметология	0,029
Часы	0,020	Повышение	0,023	Курс	0,028
Развитие	0,019	Квалификация	0,023	Занятие	0,022
Программа	0,018	Дополнительный	0,018	Уход	0,020
Образование	0,013	Часы	0,018	Модуль	0,018
Обучение	0,012	ФГОС	0,018	Лицо	0,018
Стоимость	0,012	Работников	0,016	Процедуры	0,017
Профессиональный	0,011	Документ	0,016	Медицинский	0,016
Деятельность	0,011	Предназначен	0,016	Диплом	0,015
Получить	0,011	Диплом	0,015	Онлайн	0,014
Работа	0,011	Деятельность	0,015	Тело	0,014
Онлайн	0,010	Дети	0,015	Основы	0,013
Диплом	0,010	Дошкольный	0,014	Работа	0,013
Консультирование	0,010	Образца	0,014	Профессиональный	0,013
Модуль	0,010	Форма	0,014	Научиться	0,011
Дети	0,010	Педагогических	0,014	Знания	0,011
Педагог	0,009	Установленного	0,014	Программа	0,011
Квалификация	0,009	Переподготовка	0,013	Блок	0,011
Психологический	0,009	Организация	0,013	Данный	0,011
Данный	0,009			Эстетист	0,011
Научиться	0,009			Образец	0,011

Проанализировав структурно-семантическую организацию репрезентации товаров специалистов-психологов, можно выделить специфику представления данных. На первый план выходит пространственно-временная организация обучения (курс, занятия, часы, программа, образование, обучение, модуль), финансовый аспект коммуникации (рубль, стоимость), а также внедрение идей о возможности получить работу и профессионально «вырасти» (развитие, профессиональный, деятельность, работа, диплом, получить, квалификация). Также представлены основные направления профессиональной деятельности (консультирование, работа с детьми, педагогическая работа).

Можно говорить о том, что, опираясь на ключевые слова, авторы текстов внедряют в сознание читателя узловые элементы сценария покупки обучения для специалиста: пользователь побуждается постоянному

профессиональному саморазвитию в онлайн-формате, а в качестве основных направлений обучения выступает работа с детьми и консультирование. Данный сценарий направлен на широкую аудиторию и затрагивает базовые темы самореализации, развития, желания найти работу, тему денег и возможности быстрого обучения в онлайн-формате [13].

Педагогический тезаурус включает в себя также ключевые слова-маркеры, связанные с пространственно-временной организацией профессионального обучения (курс, образование, обучение, программа, повышение, квалификация, дополнительный, переподготовка, организация, часы), важным элементом является описание документации (документы, образец, форма, установленный, диплом), также раскрываются направления работы (ФГОС, дети, дошкольный, деятельность, педагогический). Представлен образ возможного профессионального статуса покупателя

товара (предназначен, работник, педагогический).

Опишем смысловую организацию сценария покупки профессиональных товаров для педагогов: успешный педагог ориентируется на формализованную, структурированную документацию и обучение, которое предназначено для профессиональных работников образовательной сферы. Данный сценарий призван отсеять часть аудитории путем использования вербальных маркеров принадлежности к профессиональному педагогическому сообществу [14].

Смысловой фрейм предложений обучения для медицинских работников включает целостное описание дополнительной области деятельности, связанное с эстетической медициной и косметологией (массаж, кожа, косметология, уход, лицо, процедуры, тело, эстетист, образец, сфера).

Представлены элементы описания пространственной организации обучения (обучение, проводить, курс, занятие, модуль, диплом, онлайн, программа, образование), финансовой организации (рубли,

стоимость).

Одно из направлений связано с демонстрацией научности подхода (медицинский, научиться, знание, основы, профессиональный, освоить). Опишем сценарий данных рекламных постов: дифференциация косметических и косметологических услуг, требующих медицинской лицензии, описание конкретного предмета обучения, его финансовой валидности для успешной профессиональной самореализации.

На втором этапе нами были проанализированы результаты автоматизированной кластеризации информации о курсах обучения и переподготовки. Для каждого направления были выделены 10–15 основных кластеров, объединяющих ключевые слова-маркеры, описывающие направления популярного профессионального обучения.

Рассмотрим кластерную структуру направлений обучения педагогов (таблица 2).

Таблица 2. Кластерная структура репрезентаций профессионального инструментария для педагогов

Название кластера	Ключевые слова	Вес кластера
Повышение квалификации для учителей по работе с ФГОС	ФГОС, педагогический, повышение, предназначен, квалификация, работник, образование, форма, класс, данный	0,173
Коммерческие императивные предложения обучения	Необходимый, приобрести, профессия, старт, оканчивать, нужный, документ, образование, навыки, образец	0,12
Популярные направления обучения	Образование, программа, водитель, деятельность, средства, основы, НАПО, английский, документ, дополнительный	0,117
Переподготовка в сфере работы с детьми	Дети, тип, дополнительный, переподготовка, программа, педагог, образование, профессиональный, цель, образование	0,096
Финансовая и временная организация обучения	Формат, семестр, материалы, рубль, обучение, работа, задание, часть, учебный	0,095
Переподготовка на преподавателя-технолога	Преподавание, курс, специальность, ученик, разный, техника, шугаринг	0,088
Переподготовка в сфере права	Вид, новый, переподготовка, итоговый, право, диплом, день, образование	0,081
Повышение квалификации для учителей по работе с новыми стандартами образования	Министерство, образовательный, стандарт, профессиональный, педагогический, РФ, программа, защита, образование, модуль	0,078
Обучение в сфере дополнительного музыкального образования	Дети, ученики, занятия, развитие, музыкальный, особенности, доступный, дополнительный	0,071

В результате анализа кластерной структуры образовательных предложений для педагогов можно выделить узкоспециализированные направления обучения, которые возможно успешно распознать по ключевым словам. Наибольший вес имеют темы связанные с повышением квалификации для учителей по работе с ФГОС, показывающий актуальность данной темы. Большой вес имеет кластер, связанный с реализацией суггестивных посланий, о необходимости приобретения данных курсов. Так как в рамках профессиональной деятельности педагогов необходимо

постоянное повышение квалификации, коммерческое давление с посланием «необходимости приобретения навыков», будет успешно привлекать клиентов.

Также были выделены частные темы, связанные с популярными видами обучения: обучение вождению и английскому языку; обучение в сфере работы с детьми; переподготовка в сфере права, музыкального образования, на преподавателя-технолога, в сфере работы с новыми стандартами образования. Рассмотрим кластерную структуру репрезентаций образовательных услуг в рамках психологии (таблица 3).

Таблица 3. Кластерная структура репрезентаций профессионального инструментария по психологии

Название кластера	Ключевые слова	Вес кластера
Обучение в сфере семьи и брака, ориентированный на женщин.	Игра, тип, родитель, штука, замуж, характер, доступ, женщина, урок	0,099
Ораторское мастерство в работе психолога	Блок, рубль, психология, мощность, практический, звук, голос, работа, нюанс, mail	0,094
Обучение в направлении «Mindfulness»	Скуре, удобный, судьба, вопрос, занятие, рубль, освоить, осознанность	0,084

Переподготовка в сфере консультативной психологии	Проходить, сфера, переподготовка, психолог, курс, консультирование, научиться, окончание, документ, образование	0,078
Обучение работе с метафорическими картами и таро	Обучение, МАК (метафорические ассоциативные карты), помощь, психология, разобраться, предназначение, профессиональный, таро	0,073
Общепсихологические знания	Влиять, впадать, правда, поступать, воспоминание, наш, исчезать, жить, называть, мозг	0,073
Ускоренное, дистанционное обучение и содействие в трудоустройстве	Данный, специальность, трудоустройство, дистанционный, модуль, быстро, диагностика, содействие, кратчайший, психолог	0,065
Психология денег: темы, связанные с финансовым поведением и психологией денег.	Сессия, деньги, финансы, особенности, сообщить, возраст, руководитель, трансформация, психология	0,061
Обучение работе со сновидениями в психологии	Символ, сновидение, работа, фантазия, личность, взаимодействие, клиент, занятие	0,058
Переподготовка в сфере педагогической психологии	Обучение, переподготовка, образовательный, педагог, психолог, академический, являться, цель	0,057
Повышение квалификации в сфере арт-терапии в работе с детьми	Рисунок, 5 лет, доступ, ключ, особенности, квалификация, ДОО, бессрочный, повышение	0,054
Перинатальная психология, сопровождение беременности	Рубль, модуль, беременность, средний, программа, сопровождение, психологический, медицинский, факультет	0,053
Обучение работе с архетипами	Научиться, словарь, архетип, реализовать, психология, курс, делать, профессия	0,052
Обучение дошкольной и школьной диагностике (готовность к обучению в школе, школьная зрелость)	Готовность, школьный, материалы, зрелость, школьник, обучение, тест, занятие, волевой, психолог	0,043

Наибольший вес имеют кластеры, не связанные с профессиональным обучением психологии: обучение, направленное на о женщин, стремящихся создать семью; курсы обучения ораторскому мастерству; обучение эзотерическим практикам работы с картами таро. Помимо этого, представлены кластеры связанные с обучением в сфере психологии: обучение в направлении «Mindfulness»; переподготовка в сфере консультативной психологии; предоставление общепсихологических знаний; переподготовка в сфере педагогической психологии; обучение работе со сновидениями в психологии; повышение квалификации в сфере арт-терапии в работе с детьми; перинатальная психология, сопровождение беременности. Также представлена тема ускоренного, дистанционного обучения и содействие в трудоустройстве, которая актуальна для молодых специалистов и привлекает клиентов для обучения.

Можно сказать, что контент, представленный в в рекламе образовательных услуг по психологии в первую очередь ориентирован на обычных пользователей и популяризирует «развлекательную» психологию и эзотерические околонунаучные практики.

Анализ ключевых слов показал, что одним из важных маркеров выделенных кластеров является акцент на триггерных психологических темах, важных для обычных интернет-пользователей, который можно считать приемом манипулятивного воздействия. Смешанная жанровая структура, объединяющая вербальные маркеры псевдонаучных подходов, приводит к дискредитации направления обучения для потенциальных учеников.

Рассмотрим кластерную структуру направлений медицинского обучения (таблица 3).

Таблица 4. Кластерная структура профессионального инструментария в сфере медицине

Название кластера	Ключевые слова	Вес кластера
Обучение массажу	Тело, массаж, гуаша, основы, восточный, коррекция, сеанс, рубль	0,199
Обучение проведению косметических процедур на лице	Проводить, кожа, блок, ресница, косметический, работа, брови, комплексный, методики, массаж	0,125
Дополнительное обучение санитарным инструкциям	Модуль, медицинский, окончание, САН, образец, заказать, лицо, занятие, ВУС	0,109
Дополнительное обучение для медицинских сестер	Дело, сестринский, первый, специалист, медицинский, помощь, рубль	0,1
Обучение косметическим процедурам на теле	Проводить, красота, процедура, курсы, кожа, массаж, приобретать, курс, эпиляция	0,082
Коммерческие императивные предложения официального обучения	Университет, деятельность, срок, самый, сфера, официальный, данный, помочь, начать	0,078
Обучение работе с различными детскими заболеваниями	Дети, проведение, массаж, энцефалопатия, аутизм, перинатальный, сколиоз, плоскостопие, косопалость, ДЦП	0,073
Общий курс косметологии	Кожа, рубль, документы, урок, косметология, уровень, базовый, оперативный, технический, ваш	0,062
Дополнительное обучение для младшего медицинского персонала	Оказание, реанимация, заболевание, регистрация, пациент, помощь, младший, медицинский, синдром	0,05

Большая часть информации, представленная в описании товаров посвящена различным видам

косметологического обучения. Данные предложения не относятся к медицинским услугам, однако, предоставляются пользователям по запросу о медицинском обучении, что может ввести потенциальных учеников в заблуждение, так как опираясь на исследования S. H. Lee, H. Ro можно сказать, что под влиянием информатизации у потребителей сформировалось достаточно высокое доверие к информации, представленной в социальных сетях [15]. Также представлены различные виды обучения для младшего медицинского персонала по работе с документацией, освоению санитарно-гигиенических норм. Стоит отметить, что данный вид профессионально-маркированного инструментария очень структурирован. В медицинских товарах четко прописывается содержание, что усиливает впечатление и позволяет влиять на мнение читателя, так как вне зависимости от того, относится ли предложенная информация к медицинскому профилю, «научный» контекст делает информацию приемлемой [16].

После ручной разметки кластерной структуры экспертами-психологами, для проверки валидности выделенных названий, нами была проведена проверка с помощью бота на основе технологии GPT-3. Тематики, выделенные в автоматизированном формате на 70% совпали с выделенными в ручном режиме. Данная проверка дает возможность в дальнейшем полностью автоматизировать процесс анализа.

V. ЗАКЛЮЧЕНИЕ

В работе проанализирована структурно-семантическая организация товара в цифровой рекламе образовательных услуг. Выявлены наиболее популярные категории профессионального обучения, которые можно успешно распознавать по ключевым словам. Описаны базовые смысловые фреймы, задающие сценарий восприятия рекламных предложений в рамках педагогики, психологии и медицины. Полученные результаты могут послужить основой для мониторинга качества информации, которая репрезентирует как сами товары, так и поставщика предоставляемых услуг. Показаны возможности использования автоматизированного тематического моделирования (метод LDA), а также метода TF-IDF. Методика автоматизированного кластерного анализа маркетинговых стратегий, продающих образовательный контент будет востребована широким кругом специалистов в сфере менеджмента образования, разработчиками контента образовательных интернет-каналов, образовательными учреждениями, педагогами, работающими в виртуальной среде.

БЛАГОДАРНОСТИ

Исследование выполнено при финансовой поддержке гранта Российского научного фонда № 22-78-10047 «Конструктивные и деструктивные коммуникативные практики специалистов помогающих профессий в цифровых медиа» (<https://rscf.ru/project/22-78-10047/>).

БИБЛИОГРАФИЯ

- [1] Uglova A., Nizomudinov B., Model for monitoring the psychological safety of young people based on automated analysis of queries in search engines // E3S Web of Conferences. 2023. Vol. 419. P. 02030.
- [2] Wang N., Xie W., Qiu Y. Accelerating new product diffusion: How lead users serve as opinion leaders in social networks // Journal of Retailing and Consumer Services. 2023. Vol. 72. P. 103297.
- [3] Zhang H., Wang Z., Guo C. Product recommendation in online social networking communities: An empirical study of antecedents and a mediator // Information & Management. 2019. Vol. 56. I. 2. P. 185-195.
- [4] Wei X., Gong H., Song L. Product diffusion in dynamic online social networks: A multi-agent simulation based on gravity theory // Expert Systems with Applications. 2022. Vol. 213.
- [5] Ottikunta L. Improved constrained social network rating-based neural network technique for recommending products in E-commerce environment // International Journal of Intelligent Networks. 2022. V. 3. P. 80-86.
- [6] Gozuacik N., Sakar C. O., Ozcan S. Social media-based opinion retrieval for product analysis using multi-task deep neural networks // Expert Systems with Applications. 2021. Vol. 183. P. 1-13.
- [7] Wang X., Wang Y., Liu A. Determining Customer-Focused Product Features through Social Network Analysis // Procedia CIRP. 2020. Vol. 91. P. 704-709.
- [8] Hinz O., Schulze C., Takac C. New product adoption in social networks: Why direction matters // Journal of Business Research. 2014. V. 67, I. 1. P. 2836-2844.
- [9] Goirand M., Austin E., Clay-Williams R. A systems approach to implementing ethics in a COVID-19 AI application: A qualitative study Accelerating Strategic Changes for Digital Transformation in the Healthcare Industry // Information Technologies in Healthcare Industry. 2023. P. 201-218.
- [10] Anand A., Bowen M., Rangarajan D. The role of ethics in business-to-business marketing: An exploratory review and research agenda // Industrial Marketing Management. 2023. Vol. 115 (2). P. 421-438.
- [11] Toledano N., Gonzalez-Sanz J. D. Beyond the good and the right: Rethinking the ethics of academic entrepreneurship from a relational perspective // Research Policy. 2023. Vol. 53. Is. 2. P. 104944.
- [12] Ossokina D.O., Smith O. On the experience of English e-learning terminology frame modeling // Humanity Space. 2021. Vol. 10. № 7. P. 939-949.
- [13] Teplyashina A. N. Media discourse in social networks: lexical and semantic options // Humanitarian Vector. 2019. Vol. 14. № 5. P. 110-118.
- [14] Третьякова Т. П., Спиридонова В. А. Опыт когнитивно-дискурсивного анализа континуума интернет-комментариев // Язык и культура в глобальном мире. Санкт-Петербург, 2023. С. 318-322.
- [15] Lee S. H., Ro H. The impact of online reviews on attitude changes: The differential effects of review attributes and consumer knowledge // International Journal of Hospitality Management. 2016. Vol. 56. P. 1-9.
- [16] Cheng Y., Ho H. Social influence's impact on reader perceptions of online reviews // Journal of Business Research. 2015. Vol. 68. Is. 4. P. 883-887.

Углова Анна Борисовна, кандидат психологических наук, доцент кафедры психологии профессиональной деятельности и информационных технологий в образовании. Федеральное государственное бюджетное образовательное учреждение высшего образования «Российский государственный педагогический университет им. А. И. Герцена». email: anna.uglova@list.ru, ORCID: [0000-0002-8072-0539](https://orcid.org/0000-0002-8072-0539).

Богдановская Ирина Марковна, кандидат психологических наук, доцент кафедры психологии профессиональной деятельности и информационных технологий в образовании. Федеральное государственное бюджетное образовательное учреждение высшего образования «Российский государственный педагогический университет им. А. И. Герцена». email: ibogdanovs@herzen.spb.ru, ORCID: [0000-0001-7303-615X](https://orcid.org/0000-0001-7303-615X).

Низомудинов Борис Абдуллохонвич, ведущий аналитик Центра технологий электронного правительства Института дизайна и урбанистики Университета ИТМО, Санкт-Петербург (<https://itmo.ru/>), email: boris@itmo.ru, elibrary.ru: [authorid=794641](https://elibrary.ru/authorid/794641), ORCID: [0000-0002-4090-9564](https://orcid.org/0000-0002-4090-9564).

Automated Analysis of Product Representation in Digital Advertising of Educational Services

Anna B. Uglova, Irina M. Bogdanovskaya, Boris A. Nizamutdinov

Abstract — The article examines the structural and semantic characteristics of the representation of goods in digital advertising of educational services of helping specialists in social networks (using the example of the social network Vkontakte). The possibilities of automated analysis methods in identifying verbal markers of representations of professional tools for further creation of monitoring systems and ethical evaluation of information about electronic goods are being investigated. A general scheme has been developed for the automated collection of information about educational products and processing of the data obtained by thematic communities (teachers, psychologists, doctors). The LDA (Latent Dirichlet Allocation) method was used for thematic modeling of text data. Using the TF-IDF (Term Frequency-Inverse Document Frequency) method, important terms were identified in each set of descriptions for analyzing semantic frames of advertisements. Verbal markers have been identified, based on which the authors of advertising texts introduce the nodal elements of the professional training purchase scenario for a specialist into the reader's mind. The cluster structure of representations of pedagogical goods includes highly specialized descriptions intended for specialists working in general education institutions; representations of psychological goods are characterized by orientation to a wide audience, including non-specialists, and a thematically mixed structure; in the representations of medical products, the main information is not related to medicine, which can mislead users. The methodology of automated cluster analysis of marketing strategies selling educational content will be in demand by a wide range of specialists in the field of education management, content developers of educational Internet channels, educational institutions, teachers working in a virtual environment.

Keywords— social networks, parsing, LDA, helping specialists, educational products, professional tools, verbal markers.

REFERENCES

- [1] Uglova A., Nizamutdinov B., Model for monitoring the psychological safety of young people based on automated analysis of queries in search engines // E3S Web of Conferences. 2023. Vol. 419. P. 02030.
- [2] Wang N., Xie W., Qiu Y. Accelerating new product diffusion: How lead users serve as opinion leaders in social networks // Journal of Retailing and Consumer Services. 2023. Vol. 72. P. 103297.
- [3] Zhang H., Wang Z., Guo C. Product recommendation in online social networking communities: An empirical study of antecedents and a mediator // Information & Management. 2019. Vol. 56. I. 2. P. 185-195.
- [4] Wei X., Gong H., Song L. Product diffusion in dynamic online social networks: A multi-agent simulation based on gravity theory // Expert Systems with Applications. 2022. Vol. 213.
- [5] Ottikunta L. Improved constrained social network rating-based neural network technique for recommending products in E-commerce environment // International Journal of Intelligent Networks. 2022. V. 3. P. 80-86.
- [6] Gozuacik N., Sakar C. O., Ozcan S. Social media-based opinion retrieval for product analysis using multi-task deep neural networks // Expert Systems with Applications. 2021. Vol. 183. P. 1-13.
- [7] Wang X., Wang Y., Liu A. Determining Customer-Focused Product Features through Social Network Analysis // Procedia CIRP. 2020. V. 91. P. 704-709.
- [8] Hinz O., Schulze C., Takac C. New product adoption in social networks: Why direction matters // Journal of Business Research. 2014. Vol. 67, I. 1. P. 2836-2844.
- [9] Goirand M., Austin E., Clay-Williams R. A systems approach to implementing ethics in a COVID-19 AI application: A qualitative study Accelerating Strategic Changes for Digital Transformation in the Healthcare Industry // Information Technologies in Healthcare Industry. 2023. P. 201-218.
- [10] Anand A., Bowen M., Rangarajan D. The role of ethics in business-to-business marketing: An exploratory review and research agenda // Industrial Marketing Management. 2023. Vol. 115(2). P. 421-438.
- [11] Toledano N., Gonzalez-Sanz J. D. Beyond the good and the right: Rethinking the ethics of academic entrepreneurship from a relational perspective // Research Policy. 2023. V. 53. Is. 2. P. 104944.
- [12] Ossokina D.O., Smith O. On the experience of English e-learning terminology frame modeling // Humanity Space. 2021. Vol. 10. № 7. P. 939-949.
- [13] Teplyashina A. N. Media discourse in social networks: lexical and semantic options // Humanitarian Vector. 2019. Vol. 14. № 5. P. 110-118.
- [14] Tret'yakova T. P., Spiridonova V. A. Opyt kognitivno-diskursivnogo analiza kontinuumu internet kommentarijev // YAzyk i kul'tura v global'nom mire. Sankt-Peterburg, 2023. S. 318-322.
- [15] Lee S. H., Ro H. The impact of online reviews on attitude changes: The differential effects of review attributes and consumer knowledge // International Journal of Hospitality Management. 2016. Vol. 56. P. 1-9.
- [16] Cheng Y., Ho H. Social influence's impact on reader perceptions of online reviews // Journal of Business Research. 2015. Vol. 68. Is. 4. P. 883-887.

Boris A. Nizamutdinov, Leading Analyst of E-Governance Center, Institute of Design and Urban Studies, ITMO University (<http://itmo.ru/>), Saint-Petersburg, email: boris@itmo.ru, elibrary.ru: [authorid=794641](https://orcid.org/0000-0002-4090-956), [scopus.com: authorId=56938763900](https://orcid.org/0000-0002-4090-956), [ORCID: orcidID=0000-0002-4090-956](https://orcid.org/0000-0002-4090-956)

Anna B. Uglova, Candidate of Psychological Sciences, Associate Professor of the Department of Psychology of Professional Activity and Information Technologies in Education. The Herzen State Pedagogical University of Russia. email address: anna.uglova@list.ru, [ORCID: orcidID= 0000-0002-8072-0539](https://orcid.org/0000-0002-8072-0539).

Irina M. Bogdanovskaya, Candidate of Psychological Sciences, Associate Professor of the Department of Psychology of Professional Activity and Information Technologies in Education. The Herzen State Pedagogical University of Russia. email address: ibogdanovskaya@gmail.com, [ORCID: orcidID= 0000-0001-7303-615X](https://orcid.org/0000-0001-7303-615X)