

УДК 159.9:004.8(045)

С.Ю. Жданова, А.Н. Рабчевский, Е.Г. Ашихмин

ОЦЕНКА И МОНИТОРИНГ ДЕСТРУКТИВНЫХ ФОРМ ПОВЕДЕНИЯ ОБУЧАЮЩИХСЯ С ПОМОЩЬЮ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Обеспечение безопасности обучающихся является необходимым организационно-педагогическим условием в современной системе образования. Цель исследования – использование искусственного интеллекта для оценки и мониторинга деструктивных форм поведения студентов вуза. Из пространства социальной сети «ВКонтакте» с помощью системы VK.BARKOV.NET и поисковой системы SEUS были отобраны пользователи социальной сети ВКонтакте, указавшие в качестве места учебы ПГНИУ (80000 человек). Далее с помощью нейросети осуществлялся поиск информации о вовлеченности пользователей в социальные группы, характеризующиеся деструктивными формами поведения. С помощью системы искусственного интеллекта на основе анализа данных профиля, содержания публикаций и деструктивных групп был выявлен один пользователь, характеризующийся тяжелыми формами деструктивного и саморазрушительного поведения; три пользователя, у которых есть явные признаки склонности к деструктивному поведению средней тяжести. Сравнительный анализ результатов, полученных нейросетью, с результатами анализа маркеров, выделенных экспертами-психологами, показал, что нейросеть выявила четырех пользователей, в свою очередь методика на основе выделенных экспертами-психологами маркеров – пять пользователей. Различие в результатах объясняется тем, что методика, используемая искусственным интеллектом, фиксировала деструктивный контент, в то время как методика экспертов-психологов вместе с тем включала маркеры, связанные с психическим состоянием обучающихся. Описанный подход позволяет оценить потенциальные риски и деструктивные формы поведения у индивида на основе его цифрового следа в социальных сетях.

Ключевые слова: социальные сети, искусственный интеллект, безопасность, деструктивное поведение, маркеры деструктивного поведения.

DOI: 10.35634/2412-9550-2025-35-1-48-59

Введение

В контексте проблем обеспечения психологической безопасности особую актуальность приобретают вопросы профилактики форм деструктивного поведения в интернет-среде.

Анализ литературы показывает, что рассматриваемая проблема относится к числу активно обсуждаемых. Авторы исследуют роль и влияние Интернета на психическое состояние и поведение молодежи, изучают личностные особенности подростков, склонных к интернет-зависимости и деструктивному поведению в Сети, рассматривают формы девиантного поведения в интернет-пространстве, факторы, их обуславливающие, обсуждают вопросы выявления и профилактики рисков в виртуальной среде, говорят о необходимости предотвращения угроз в Интернет-пространстве, разработке позитивного контента для молодежи в социальных медиа.

Анализируя влияние социальных сетей с деструктивным контентом на личность подростков и молодежи, авторы отмечают, что недостаточный уровень цифровой грамотности и позитивной цифровой культуры, а также возрастно-психологические особенности современных подростков и молодежи делают их наиболее уязвимыми к угрозам деструктивного поведения в Интернет-среде [10; 14].

Так, Д.В. Новиков считает, что регулярное использование цифровых и медиа-технологий (мессенджеров, социальных сетей) приводит не только к положительным аспектам, но и негативным – обуславливает деструктивные проявления у подростков и молодежи в сети Интернет [14].

Анализируя риски, обусловленные интернет-средой, Р.Е. Загоруйко полагает, что не критичное отношение к деструктивному контенту в социальных сетях способствует десоциализации современных подростков, становится предпосылкой их девиантного поведения. Это связано, как считает автор, с тем, что подростки воспринимают деструктивный контент как правдивый, активно взаимодействуют с ним, используют и руководствуются в своей жизни [6].

Данной точки зрения придерживается и Ю.Н. Сеницын [17]. Рассматривая роль и влияние Интернета на современных подростков и молодежь, автор пишет, что сегодня деструктивное поведение занимает одно из ведущих мест среди подрастающего поколения. Ю.Н. Сеницын обращает внимание на тенденцию увеличения количества и форм личностно-психологических, социально психологиче-

ских, социально-институциональных девиаций. Анализируя факторы, обуславливающие деструктивное поведение в интернет-среде, автор считает, что сегодня достаточно активно посредством социальных сетей Интернет среди молодежи распространяются негативные образы, примеры, идеалы поведения и деятельности. Исходя из этого автор полагает, что наряду с новыми возможностями и положительными моментами использования инновационных технологий, цифровизация повлекла за собой ряд серьезных проблем, обозначила перед педагогами и родителями новый вызов в виде растущей проблемы деструктивного поведения обучающихся в цифровой образовательной среде [17].

Вопросы влияния социальных сетей на подростков в контексте психологической безопасности рассматривает также Ю.А. Захарова. Анализируя возможные риски, которым могут подвергнуться подростки под влиянием информации на страницах сообществ и профилей социальных сетей, автор говорит о том, что наличие потенциальных рисков для несовершеннолетних пользователей сети Интернет составляет угрозу психологической безопасности их личности [7].

Рассматривая формы деструктивного поведения в Сети и факторы, их обуславливающие, исследователи обсуждают в основном такие формы деструктивного поведения в интернет-среде, как киберзависимость, кибербуллинг, злоупотребление временем в социальных сетях, участие подростков и молодежи в опасных онлайн-играх, говорят о неадекватном использовании социальных сетей, размещении вредоносной информации, интернет-фейков. Вместе с тем, практически не освещенными остаются вопросы размещения в социальных сетях и мессенджерах деструктивного и аутодеструктивного контента, связанного с угрозой жизни, причинением вреда себе и другим. Исследования, посвященные данной проблеме, являются крайне малочисленными [8; 9; 13; 16; 18].

Так, в исследовании Е.М. Тимошиной показано негативное влияние деструктивных интернет-субкультур на личность подростка [18]. Выделяя два вида деструктивного контента: противоправный (суицидальные сообщества, нацизм) и аморальный (шок-контент, радикальный феминизм, ЛГБТ-пропаганда), автор говорит о негативном влиянии данного контента на психику подростков, считает, что воздействие деструктивного контента может привести к виктимизации и криминализации подростков и молодежи [18].

Особый интерес представляет вопрос о предикторах, обуславливающих деструктивное поведение в интернет-пространстве. Рассматривая в качестве важных факторов и причин деструктивного поведения в виртуальной среде свойства личности, исследователи отмечают, что подростки, проявляющие интерес к деструктивному контенту в социальных сетях, имеют низкую самооценку и конфликты с окружающими [6]. Они в большей мере ориентированы на виртуальную реальность и общение в интернет-пространстве, в то время как в реальном мире чувствуют себя неуверенно, проявляют пассивность, не видят перспектив своего будущего [1; 6; 21].

Так, в исследовании М.Н. Черкасовой и А.В. Тактаровой было показано, что для лиц, характеризующихся деструктивным поведением в социальных сетях, характерны нарушения психического состояния, перенос игры в реальность, изменения в эмоциональной сфере [20]. Авторы считают, что социальные сети могут стимулировать гедонистическую агрессию, способствовать созданию искусственного образа, согласно которому понятия «жизнь» и «смерть» теряют свою значимость в виртуальном мире, а агрессия воспринимается как нечто необходимое и конструктивное [20].

Рассматривая деструктивное поведение как разрушительное поведение человека, характеризующееся отклонением от существующих норм поведения, приводящее к снижению критичности, когнитивным искажениям восприятия и понимания происходящего, эмоциональным нарушениям, наносящее вред самому себе или окружающим [14; 17], авторы говорят о необходимости его профилактики [2-7; 10; 12; 14; 17; 19].

В качестве мер профилактики рисков в онлайн-пространстве исследователи предлагают: мониторинг активности учащихся в цифровом пространстве, ограничение доступа к цифровой среде, различные программы, направленные на предотвращение антисоциального использования Интернета; считают, что в ходе программ профилактики особое внимание следует уделять формированию у обучающихся самоуважения, уважения к другим, нравственных ценностей, критического мышления, социальной и гражданской ответственности [6; 7; 10; 14; 17].

Ряд авторов считает, что профилактику деструктивного поведения учащихся в цифровой образовательной среде важно осуществлять посредством обучения цифровой грамотности, развития навыков использования цифровых технологий, осознанного поведения в социальных сетях, проверки достоверности информации в Интернете [17].

В качестве форм, средств и способов профилактики авторы предлагают организацию тренингов, групповых и индивидуальных сессий, консультаций по профилактике цифрового насилия, развитию навыков эмоционального контроля и укреплению психического здоровья учащихся [11; 17].

О необходимости профилактики и минимизации деструктивного поведения студентов высшей школы говорит Д.В. Новиков [14]. Автор пишет о необходимости организации просветительной работы, направленной на формирование у молодежи целостного позитивного мировоззрения, пропаганду здорового образа жизни, профилактику правонарушений, повышение культуры информационной безопасности. В качестве мер противодействия проявлениям деструктивного поведения в цифровой среде и их профилактики автор предлагает формировать у обучающихся цифровые компетенции, а также осуществлять наблюдение за ходом образовательного процесса и деятельностью учащихся [14]. Профилактика деструктивного поведения в виртуальной среде может осуществляться, согласно точке зрения Д.В. Новикова, также за счет реализации образовательной программы, которая включает в себя лекции и практикумы по информационной безопасности и ведению социальных сетей, мессенджеров; практические занятия по ведению деловой переписки и цифрового этикета в сети Интернет; занятия на тему: «Фактчекинг: способы проверки достоверности информации»; мастер-классы на темы социальной инженерии и кибербуллинга [14].

Захарова Ю.А. считает, что основным направлением работы по профилактике деструктивного поведения у подростков и молодежи в социальных сетях является развитие у них эмоционального интеллекта и критического мышления, высокий уровень которых включает в себя способность противостоять манипулятивному влиянию, умение анализировать информацию, делать конструктивные выводы, отслеживать и контролировать свою вовлеченность в интернет-сообщества и время в соцсетях. Развивая данные навыки у обучающихся, по мнению автора, возможно предостеречь различные потенциальные риски, с которыми сегодня сталкиваются активные пользователи социальных сетей [7].

Рассматривая специфику профилактики деструктивных форм проявления у обучающихся в вузе, следует отметить, что в качестве одного из способов профилактики деструктивного поведения традиционно рассматривается социально-психологическое тестирование, направленное на раннее выявление незаконного потребления наркотических средств и психотропных веществ. Данный вид психологического тестирования определен в законе об образовании (пп. 15.1, п.3, ст. 28, Гл.3, Федерального закона «Об образовании в Российской Федерации» от 29.12.2012 № 273-ФЗ).

Однако в вузе реализовать данный вид тестирования достаточно сложно [15]. Большой объем обучающихся, например, в ПГНИУ количество студентов составляет 16000 человек, не позволяет провести одновременное тестирование в течение одной-двух недель, как это регламентировано в нормативных документах. Кроме того, поскольку тестирование является добровольным, не все обучающиеся могут принимать в нем участие. Это приводит к тому, что часть учащихся, не принявших участие в тестировании, могут не получить квалифицированную помощь и поддержку со стороны преподавателей, кураторов, тьюторов, психологов.

В этой связи эффективным способом профилактики деструктивных форм поведения обучающихся является мониторинг вовлеченности студентов в группы, интернет-сообщества, склонные к деструктивному поведению на основе данных из открытых источников. Такими источниками могут быть социальные сети, онлайн-мессенджеры, сайты, блоги и т. д.

В рамках реализации совместного проекта Центра психолого-педагогической помощи Пермского государственного национально-исследовательского университета (ПГНИУ) и Института компьютерных наук и технологий (ИКНТ) ПГНИУ, направленного на создание портала информационной безопасности и обеспечение информационно-психологического сопровождения обучающихся ПГНИУ, с помощью систем искусственного интеллекта нами было реализовано совместное исследование.

Цель исследования – использование искусственного интеллекта для оценки и мониторинга деструктивных форм поведения студентов вуза.

Научная новизна исследования связана с разработкой новых методов анализа и изучением деструктивного контента в социальных сетях; расширением междисциплинарного взаимодействия для диагностики и выявления форм деструктивного поведения у обучающихся в вузе; применением для поиска форм деструктивного поведения в социальных сетях систем искусственного интеллекта.

Теоретическое значение исследования заключается в обогащении психологического знания, связанного с методологией исследования деструктивного поведения, специфики его проявления у обучающихся в студенческом возрасте. Исследование развивает представление о маркерах деструк-

тивного поведения в социальной сети; вносит вклад в изучение роли систем искусственного интеллекта при анализе данных о деструктивных формах поведения в социальных сетях.

Практическое значение работы заключается в том, что описанная в статье методология исследования, принципы, критерии формирования базы данных, маркеры деструктивного поведения в социальных сетях, выделенные на основе экспертной оценки и системы искусственного интеллекта, позволяют комплексно оценить потенциальные риски вовлечения индивида в деструктивные интернет-сообщества в социальных сетях на основе цифрового следа, выстраивать воспитательную и профилактическую работу; осуществлять психологическое сопровождение студентов, с целью обеспечения психологической безопасности личности и образовательного пространства вуза.

Методы

С целью изучения возможности формирования отчета о вовлеченности студентов в деструктивные интернет-сообщества на основе данных, получаемых из открытых источников из социальных сетей с применением искусственного интеллекта, нами было реализовано решение следующих задач:

1. Формирование базы данных пользователей из пространства социальной сети «ВКонтакте»;
2. Анализ анкетных данных, содержащихся в профилях пользователей;
3. Анализ содержания профилей и опубликованных пользователями документов;
4. Поиск информации о включенности пользователей (студентов ПГНИУ) в деструктивные социальные явления на основе цифрового следа в социальных сетях с помощью системы искусственного интеллекта.

5. Сравнение результатов автоматического анализа, выполненного нейросетью, с результатами анализа маркеров деструктивного поведения в социальных сетях, выделенных экспертами-психологами.

В рамках реализации первого этапа исследования с помощью технических программных средств формировалась база данных. Для этого из пространства социальной сети «ВКонтакте» с помощью технических программных средств были отобраны пользователи социальной сети ВКонтакте, указавшие в качестве места учебы ПГНИУ. Поиск был выполнен дважды, с помощью системы набора инструментов для поиска целевой аудитории VK.BARKOV.NET и поисковой системы CEUC.

Система VK.BARKOV.NET представляет собой набор инструментов для поиска целевой аудитории ВКонтакте (Официальная страница сайта. URL: <https://vk.barkov.net/> дата посещения 14.12.2023 г.)

С помощью данных систем анализировались пользователи, у которых Пермский государственный национальный исследовательский университет указан в качестве места учебы в настоящее время.

Использование двух систем объясняется тем, что разные системы могли дать различные результаты, поэтому результаты работы обеих систем были объединены.

Общее количество таких пользователей составило более 80 000.

На следующем этапе для каждого из общего числа отобранных пользователей социальной сети «ВКонтакте» были собраны анкетные данные, которые пользователь указал о себе в своем профиле.

Из общего количества были отобраны 2754 профиля пользователей. Отбор осуществлялся на основе общего объема информации, содержащейся в профилях пользователей. Задача состояла в том, чтобы отобрать тех пользователей, кто указал о себе как можно больше информации.

Для каждого из 2754 пользователей была собрана информация о том, в какие группы они входят, проанализировано содержание их профилей и опубликованных ими документов. Сбор данной информации был выполнен через API VK с помощью программы, специально написанной для этой цели.

Далее из полученных профилей методом случайного отбора была сформирована группа из 29 профилей, что соответствует типичной численности одной учебной группы студентов.

Такой подход позволяет смоделировать ситуацию комплексного изучения группы как целостного коллектива и составляющих его индивидов, а также получить репрезентативный объем данных для анализа индивидуальных различий между испытуемыми [13]. Количество профилей в объеме учебной группы дает достаточный массив информации для апробации методов контент-анализа с использованием технологий искусственного интеллекта. При этом выборка не является избыточной, что позволяет провести пилотное исследование с приемлемыми временными и вычислительными затратами. Таким образом, объем выборочной совокупности в 29 профилей оценивается как оптимальный для заявленных целей исследования.

Для каждого из 29 пользователей с помощью поисковой системы СЕУС были собраны все находящиеся в открытом доступе публикации, которые авторы сделали с момента регистрации аккаунта.

На основе собранных данных было сформировано цифровое досье на каждого пользователя, которое включало: анкетные данные профиля пользователя; данные о группах, в которых состоит пользователь; контент, опубликованный каждым пользователем.

Анализ проводился в автоматическом режиме с использованием методов обработки естественного языка с помощью большой языковой модели Antropic Claude. В связи с ограничениями технических возможностей данной модели, были использованы следующие ограничения:

- количество групп, в которых состоит пользователь, было не более 100;
- тексты публикаций обрезались по количеству строк, которые может принять бесплатная версия модели.

В ходе формирования базы данных в анкетных данных пользователей анализировались следующие показатели:

- *Возраст*. Оценивается соответствие возрастным особенностям анкетным данным.
- *Статус профиля*. Анализируется на предмет проявлений деструктивных черт личности.
- *Количество друзей, подписчиков, записей*. Оценивается активность и включенность индивида в интернет-сообщества.

– *Перечисленные интересы*. Анализируются на предмет разнообразия и ориентации на саморазвитие.

– *Любимая музыка, фильмы, книги*. Оценивается глубина предпочтений.

– *Цитаты*. Анализируются на позитивность/негативность.

– *Текст «О себе»*. Выявляются особенности Я-концепции и самооценки.

– *Политические предпочтения и религия*. Оценивается толерантность.

– *Приоритетные ориентации*. Определяется направленность личности.

– *Отношение к вредным привычкам*. Анализируется с точки зрения самоконтроля.

В данных о сообществах, в которые входит пользователь, оценивались следующие показатели:

– *Тематика сообществ*. Оцениваются интересы и ценности индивида;

– *Описания сообществ*. Анализируются на позитивный/деструктивный характер контента.

Анализ сообществ на предмет деструктивной направленности выполнялся по названию и описанию:

– экстремистские группы;

– закрытые секты;

– группы, пропагандирующие насилие и жестокость и т. д.

Тексты публикаций пользователя анализировались по следующим показателям:

– оценивается эмоциональный фон, преобладающая модальность текстов;

– анализируются упоминаемые темы на соответствие перечисленным интересам;

– выявляются проявления деструктивных черт личности.

Анализ публикаций индивида на предмет наличия текстов деструктивной направленности выполнялся по следующим признакам:

– эмоциональная окраска текстов (негативизм, агрессия);

– наличие призывов и оправдания насилия или деструктивных идей;

– упоминание фактов участия в деструктивной деятельности.

После того, как была сформирована база данных, с помощью системы искусственного интеллекта осуществлялся поиск информации о вовлеченности пользователей (студентов ПГНИУ) в интернет-сообщества, характеризующиеся деструктивными формами поведения на основе цифрового следа в социальных сетях.

Для численной оценки деструктивного поведения в социальных сетях была использована следующая шкала:

1 – для респондента характерны тяжелые формы деструктивного и саморазрушительного поведения;

2 – есть явные признаки склонности к деструктивному поведению средней тяжести;

3 – нейтральное отношение, отсутствие очевидных признаков склонности к деструкции;

4 – есть явное позитивное отношение к конструктивному поведению;

5 – респондент демонстрирует высокий уровень вовлеченности в конструктивные и созидательные группы.

На основе сводных оценок по каждому типу данных числовой оценки формировался интегральный показатель деструктивности и предлагались рекомендации по возможности предотвращения вовлечения индивида в деструктивные группы и сообщества.

На следующем этапе работы осуществлялось сравнение результатов автоматического анализа, выполненного нейросетью, с результатами анализа маркеров деструктивного поведения в социальных сетях, выделенных психологами Центра психолого-педагогической помощи ПГНИУ.

Маркеры деструктивного поведения в социальных сетях были выделены на основе экспертной оценки. В ней приняли участие 5 человек. В качестве экспертов выступили 5 психологов – сотрудников Центра психолого-педагогической помощи (1 доктор психологических наук, 1 кандидат психологических наук, 3 магистра), имеющих опыт работы с контент-анализом, а также опыт работы с проявлениями деструктивных форм поведения у обучающихся, в том числе в социальных сетях. Степень согласия экспертов составила 100 %.

Маркеры деструктивного поведения в социальных сетях были структурированы и объединены экспертами в четыре блока.

Первый блок включает контент о самоповреждениях, причинении вреда себе или окружающим. В данный блок вошли высказывания, которые включают:

- содержание угроз, свидетельствующих о причинении вреда себе и другим;
- посты, репосты, изображения, видео и подписки, содержащие семантические элементы вербального или визуального характера: самоповреждения (порезы, ушибы, раны); разного рода деформации и повреждения тела; татуировки, пирсинг;
- перебинтованные руки, множественные шрамы;
- публикация статусов с отсчетом дней, прощальных записок, угроз совершения самоубийства, сообщений, посвященных «ближайшей встрече с умершими людьми»;
- посты, свидетельствующие о склонности к самоубийству;
- прощание, например, пользователь пишет: «Мне с Вами было хорошо», просит прощения у окружающих, одноклассников.

Второй блок включает агрессивный контент. Данный блок содержит следующие маркеры, свидетельствующие о наличии признаков агрессии:

- на странице в социальных сетях размещены фотографии в военной форме, автомат, ружье, нацистские призывы и т. д.;
- посты, репосты, изображения, видео и подписки, содержащие следующие семантические элементы вербального или визуального характера (агрессия, оружие, кровь, насилие и др.);
- подписки на сообщества, романтизирующие криминальный образ жизни и тюремную культуру; статусы, содержащие тюремные жаргонизмы;
- публикация статусов, критикующих и оскорбляющих других людей по признаку национальности, религии, социального статуса;
- подписки на сообщества, содержащие упоминания следующих терминов: «ультра», «ультраправые», «белая раса», «белая сила», «многонационал», неонацизм и др.;
- использование рунического письма.

Всего в данном блоке было выделено 106 высказываний.

Третий блок содержит маркеры, включающие депрессивный контент. К этому блоку были отнесены:

- посты, репосты, изображения, видео и подписки, содержащие семантические элементы вербального или визуального характера, свидетельствующие о страдании, депрессивных переживаниях;
- публикации депрессивных статусов;
- в аудиозаписях преобладание музыки, посвященной депрессивным темам (потеря друзей, бессмысленность жизни и др.).

В этом блоке было выделено 15 высказываний, например: «я не вижу смысла в жизни», «у меня больше нет сил», «мне невыносимо», «я устал», «я ничего не чувствую», «я совсем один».

Четвертый блок содержит маркеры, свидетельствующие о стремлении к обособленному поведению:

- состоит в закрытых или открытых сообществах, содержание которых является деструктивным;
- закрытый аккаунт;
- не отвечает старостам, тьюторам, не выходит на связь;
- удаляется из социальных сетей, чатов.

Поиск маркеров деструктивного поведения осуществлялся на основе открытых данных в социальной сети ВКонтакте с помощью методов компьютерной лингвистики.

Результаты

Для выполнения автоматической оценки уровня вовлеченности индивида в деструктивные группы в социальных сетях, согласно представленной выше методике, использовалась большая языковая модель Antropic Claude. Каждый вид данных (анкетные данные профиля, данные групп и содержание текстов публикаций) каждого из 29 пользователей предлагался этой модели для получения ответа в виде численной оценки. Затем оценки заносились в общую таблицу, после чего формировалась сводная таблица интегрированной оценки по всем отобранным пользователям.

Результаты автоматической оценки уровня вовлеченности пользователей в деструктивные интернет-сообщества представлены в табл. 1.

Таблица 1

Результаты автоматической оценки уровня вовлеченности пользователей в деструктивные интернет-сообщества, полученные на основе искусственного интеллекта

№ респондента	Группы	Тексты	Итого (ср)
1	3	3	3
2	3	текстов нет	3
3	2	4	3
4	3	3	3
5	5	3	4
6	5	3	4
7	3	4	4
8	2	3	3
9	3	3	3
10	3	3	3
11	1	3	2
12	5	3	4
13	5	3	4
14	3	3	3
15	3	3	3
16	3	3	3
17	3	3	3
18	3	3	3
19	3	3	3
20	3	3	3
21	3	3	3
22	3	4	4
23	3	3	3
24	3	3	3
25	2	4	3
26	3	3	3
27	3	3	3
28	3	3	3
29	3	3	3
Среднее значение	3,1	3,1	3

На основании полученных результатов можно сделать вывод о том, что в группе, состоящей из 29 пользователей, случайно отобранных из 2754 аккаунтов, выявлен один пользователь, для кото-

рого характерны *тяжелые формы деструктивного и саморазрушительного поведения*, что было установлено системой искусственного интеллекта на основе анализа данных профиля, содержания публикаций и групп, в которых состоит данный пользователь.

Кроме того, было выявлено 3 пользователя (3, 8, 25), у которых есть явные признаки склонности к деструктивному поведению средней тяжести. Данные пользователи были выявлены нейросетью на основе анализа текстов пользователей.

Всего нейросеть выявила 4 пользователей из 29, характеризующихся деструктивным поведением в социальных сетях.

Для того чтобы убедиться в корректности представленной методики и результатов оценки, полученных на основе анализа нейросети, было проведено сравнение результатов автоматического анализа, выполненного с помощью искусственного интеллекта, и анализа той же информации, выполненного на основе маркеров, предоставленных специалистами-психологами.

Для анализа использовались маркеры Блоков 1-3: причинение вреда, агрессия и депрессия. Поскольку данный этап является пилотным, для анализа использовались не все маркеры, указанные в этих блоках, а только текстовые высказывания и изображения свастики в описании профиля, в группах и постах на странице пользователя. Результаты анализа представлены в табл. 2.

Таблица 2

Результаты исследования деструктивного поведения пользователей в социальной сети ВКонтакте, полученные на основе маркеров деструктивного поведения, выделенных в ходе экспертной оценки

№ респондента	Причинение вреда	Агрессия	Депрессия	Сумма
1	0	2	0	2
3	0	2	0	2
4	0	2	0	2
5	0	1	0	1
6	0	1	1	2
7	2	3	0	5
9	0	2	0	2
10	3	2	0	5
13	2	1	0	3
14	0	1	0	1
15	0	1	0	1
16	0	1	0	1
17	0	4	1	5
18	2	12	0	14
19	2	2	0	4
21	0	1	0	1
22	0	3	0	3
23	0	2	0	2
24	1	3	0	4
25	0	2	0	2
27	0	2	0	2
28	0	8	0	8

Примечание: пользователи, имеющие нулевые значения по всем трем блокам, в таблице не указаны.

Согласно представленным в таблице результатам, один пользователь имеет высокие показатели по шкале агрессия, один – существенные и 3 человека – показатели выше среднего. Таким образом, на основе маркеров деструктивного поведения, выделенных в ходе экспертной оценки, было выявлено пять пользователей, характеризующихся деструктивным, аутодеструктивным контентом и нуждающихся в особом внимании со стороны специалистов.

Сопоставление результатов, полученных с помощью искусственного интеллекта на основе анализа анкет и контента, представленного пользователями в социальных сетях, а также данных, полученных на основе маркеров, выделенных в ходе экспертной оценки психологами Центра психолого-педагогической помощи ПГНИУ, свидетельствует о различии полученных результатов. Это проявляется в том, что в ходе анализа анкет и контента нейросеть выявила 4-х пользователей (3, 8, 11, 25), а методика на основе маркеров, выделенных на основе экспертной оценки психологов Центра психолого-педагогической помощи ПГНИУ, – 5 пользователей, требующих особого внимания специалистов (7, 10, 17, 18, 28). Различие в результатах объясняется тем, что методика, используемая искусственным интеллектом, акцентировала внимание на деструктивный контент в социальных сетях, в то время как методика психологов Центра психолого-педагогической помощи помимо контента, свидетельствующего о склонности к самоповреждениям, причинению угрозы жизни себе и другим, включала также маркеры, связанные с психическим состоянием, например, во внимание принимались показатели, свидетельствующие о наличии признаков агрессии, депрессии.

Поскольку методика, на которой обучалась нейросеть, и маркеры, представленные Центром психолого-педагогической помощи, существенно отличаются друг от друга, пользователь, выявленный методикой на основе искусственного интеллекта, не мог быть выявлен методикой специалистов Центра психолого-педагогической помощи, ввиду отсутствия соответствующих ключевых слов или выражений в составе предоставленных маркеров. Очевидно, что в случае, если бы необходимые маркеры (вовлечение индивида в деструктивные интернет-сообщества в социальных сетях) присутствовали, то степень совпадения результатов была бы выше.

Выводы

С целью изучения возможности формирования отчета о вовлеченности студентов в деструктивные группы на основе данных, получаемых из открытых источников социальных сетей с применением искусственного интеллекта, нами было реализовано решение следующих задач:

1. В ходе исследования была сформирована база данных пользователей из пространства социальной сети «ВКонтакте». Базу данных составили пользователи, у которых Пермский государственный национальный исследовательский университет указан в качестве места учебы в настоящее время. Поиск осуществлялся с помощью технических программных средств VK.BARKOV.NET и поисковой системы СЕУС.

2. Были проанализированы анкетные данные профиля пользователя; данные о группах, в которых состоит пользователь; контент, опубликованный каждым пользователем. Анализ проводился в автоматическом режиме с использованием методов обработки естественного языка с помощью большой языковой модели Antropic Claude, а также на основе специальных маркеров.

3. Анализ содержания профилей и опубликованных пользователями документов с помощью системы искусственного интеллекта позволил выявить пользователей, вовлеченных в деструктивные социальные группы на основе цифрового следа в социальных сетях.

4. В ходе исследования с помощью системы искусственного интеллекта на основе анализа данных профиля, содержания публикаций и групп, в которых состоит данный пользователь, был выявлен один пользователь, характеризующийся *тяжелыми формами деструктивного и саморазрушительного поведения*; также было выявлено 3 пользователя, у которых есть явные признаки склонности к деструктивному поведению средней тяжести. Данные пользователи были выявлены нейросетью на основе анализа текстов пользователей.

5. Сравнительный анализ результатов, полученных на основе анализа нейросетью, выполненных с помощью искусственного интеллекта, с результатами анализа маркеров деструктивного поведения в социальных сетях, выделенных экспертами-психологами, показал, что в ходе анализа анкет и контента нейросеть выявила 4-х пользователей, а методика на основе маркеров, выделенных в ходе экспертной оценки психологами Центра психолого-педагогической помощи ПГНИУ, – 5 пользователей, требующих особого внимания специалистов. Различие в результатах объясняется тем, что методика, используемая искусственным интеллектом, акцентировала внимание на деструктивный контент, в то время как методика психологов Центра психолого-педагогической помощи помимо контента, свидетельствующего о склонности к самоповреждениям, причинению угрозы жизни себе и другим, включала также маркеры, связанные с психическим состоянием, например, во внимание принимались показатели, свидетельствующие о наличии признаков агрессии, депрессии.

Заключение

Таким образом, важным направлением в профилактике деструктивных форм поведения у обучающихся в вузе является оценка и мониторинг деструктивного поведения на основе данных, взятых из открытых источников (социальные сети, онлайн-мессенджеры, сайты, блоги и т. д.).

Результаты анализа данных, полученные на основе автоматической оценки с помощью искусственного интеллекта, о вовлеченности в деструктивные интернет-сообщества пользователей социальной сети ВКонтакте, демонстрируют способность дополнять методику на основе маркеров деструктивного поведения, выделенных в ходе экспертной оценки психологами Центра психолого-педагогической помощи ПГНИУ.

Результаты сравнения показывают, что методика оценки деструктивного поведения в социальных сетях на основе искусственного интеллекта может быть применена для усовершенствования методики анализа, осуществляемого на основе маркеров деструктивного поведения, которая является более научно обоснованной. В то же время методика, включающая в себя анализ маркеров деструктивного поведения с помощью методов компьютерной лингвистики, в отличие от методики анализа, осуществленного на основе систем искусственного интеллекта, может быть использована на локальных компьютерах, не имеет ограничений по производительности, не требует обращения к платным облачным сервисам, поэтому является более перспективной для внедрения в проект портала психологической безопасности ПГНИУ и масштабирования ее на другие учебные заведения.

Применение автоматической оценки искусственным интеллектом может существенно повысить охват аудитории учащихся университета, с другой стороны, может проводиться регулярно. Данные такой оценки могут служить источником информации для более углубленной оценки пользователей специалистами-психологами. Для практического внедрения предложенной методики необходимо в дальнейшем разработать систему деанонимизации пользователей, с целью сопоставления пользователей социальной сети и реальных людей, с которыми могли бы работать специалисты в случае выявления определенных рисков. Представленный в статье подход позволяет комплексно оценить потенциальные риски вовлечения индивида в деструктивные интернет-сообщества и социальные группы на основе его цифрового следа в социальных сетях, выстраивать воспитательную и профилактическую работу; осуществлять психологическое сопровождение студентов с целью обеспечения психологической безопасности личности и образовательного пространства вуза.

СПИСОК ЛИТЕРАТУРЫ

1. Беликов В.А., Романов П.Ю., Леушканова О.Ю., Васева Н.С. Анализ проблемы десоциализации личности подростков под воздействием социальных сетей, содержащих контент деструктивного характера // Перспективы науки и образования. 2022. № 3 (57). С. 375.
2. Дьячек П.Э., Арцыбашев И.Г. Возможные способы защиты подростков от влияния деструктивного контента в социальных сетях // Государственная молодежная политика: вызовы и современные технологии работы с молодежью. 2022. С. 161.
3. Жданова С.Ю., Печеркина А.В. Обеспечение психологической безопасности личности в образовательной среде вуза // Комплексные исследования человека: психология: сб. науч. ст. Томск: Издательский Дом Томского государственного университета, 2017. С. 72–75.
4. Жданова С.Ю., Доронина В.Ф. Совладание с контентными рисками в социальных сетях у студентов // Вестник Пермского государственного гуманитарно-педагогического университета. 2019. № 1. С. 36–43.
5. Жданова С.Ю. Роль психологической службы в обеспечении психологической безопасности в вузе // Социальные и гуманитарные науки: теория и практика. 2022. № 1(6). С. 194–208.
6. Загоруйко Р.Е. Психолого-педагогическое предотвращение размещения подростками деструктивного и аутодеструктивного контента в социальных сетях и мессенджерах: обзорное исследование // *Advances in Science and Technology*: сб. науч. ст. М.: ООО Актуальность, 2024. С. 429–433.
7. Захарова Ю.А. Потенциальные риски психологической безопасности пользователей социальных сетей // Шаг в науку: сб. науч. ст. М.: Медиагруппа «ХАСК», 2024. С. 462–466.
8. Зинченко Ю.П., Шилко Р.С. Психологические аспекты информационной безопасности и противодействия терроризму посредством медиа // Информационная и психологическая безопасность в СМИ: сб. науч. ст. М.: Аспект Пресс, 2008. С. 199–226.
9. Игумнов С.А., Шамарова Е.Ю. Контент-анализ сайтов аутодеструктивного содержания в русскоязычном сегменте Интернет // Цифровое общество как культурно-исторический контекст развития человека. 2020. С. 148.

10. Кошкина К.С. Социально-психологическая профилактика деструктивного поведения подростков в сети Интернет // Региональная культура как компонент содержания непрерывного образования: сб. науч. ст. Липецк, 2024. 186–190.
11. Краева В.С., Жданова С.Ю. Актуальные проблемы обеспечения психологической безопасности в вузе // Современные проблемы теории и практики специальной и социальной педагогики: сб. науч. ст. Сургут, 2022. С. 500–507.
12. Комаров В.В. Психолого-педагогическая работа по профилактике девиантного поведения обучающихся в сети Интернет // Психолого-педагогические проблемы развития современного образования. 2022. С. 167.
13. Минаев В.А., Бондарь К.М., Рабчевский А.Н., Федорович В.Ю. Противодействие экстремистской идеологии в социальных медиа: математические модели и методы // Дальневост. юрид. ин-т МВД России. 2023. С. 232.
14. Новиков Д.В. Профилактика деструктивного поведения обучающихся высших учебных заведений под влиянием сети Интернет // Педагогическое образование: новые вызовы и цели: VII Международный форум по педагогическому образованию: сб. науч. ст. Казань, 2021. С. 8–19.
15. Печеркина А.В., Жданова С.Ю., Зарипова Л.З. Проблемы организации социально-психологического тестирования обучающихся в условиях высшего учебного заведения // Социальные и гуманитарные науки: теория и практика. 2020. № 1(4). С. 395–402.
16. Пучнина М.Ю., Пучнин А.В. Анализ деятельности адептов радикального идейного движения «Колумбайн» // Lex Russica (Русский закон). 2021. № 12 (181). С. 76.
17. Синицын Ю.Н., Хохлова У.В. Профилактика деструктивного поведения обучающихся в цифровой образовательной среде // Цифровизация в системе образования: передовой опыт и практика внедрения: сб. науч. ст. Краснодар: Новация, 2024. С. 98–101.
18. Тимошина Е.М. Деструктивные субкультуры несовершеннолетних как условия их виктимизации и криминализации // Ученые записки Казанского юридического института МВД России. 2021. № 2(12). С. 273.
19. Шиняев К.А. Профилактика деструктивного поведения личности несовершеннолетнего // Вестник практической психологии образования. 2018. № 1. С. 78–80.
20. Черкасова М.Н., Тактарова А.В. Деструктивная поведенческая модель поколения Z: лингвопрагматический анализ медиа-текстов // Гуманитарные и социальные науки. 2022. № 2. С. 97.
21. Юдеева Т.В. Мотивация деструктивного коммуникационного поведения подростков в социальных сетях // Международный научно-исследовательский журнал. 2022. № 2-2 (116). С. 179.

Поступила в редакцию 22.07.2024

Жданова Светлана Юрьевна, доктор психологических наук, профессор,
заведующая кафедрой психологии развития ПГНИУ,
директор центра психолого-педагогической помощи ПГНИУ
E-mail: Svetlanaur@gmail.com

Рабчевский Андрей Николаевич, кандидат технических наук, старший преподаватель
Института компьютерных наук и технологий ПГНИУ
E-mail: ran@psu.ru

ФГАОУ ВО «Пермский государственный национальный исследовательский университет»
614990, Россия, г. Пермь, ул. Букирева, 15

Ашихмин Евгений Георгиевич, магистр социологии, аспирант по направлению социология
ФГАОУ ВО «Пермский национальный исследовательский политехнический университет»
614990, Россия, г. Пермь, Комсомольский пр., 29
E-mail: e.ashikhmin@icloud.com

S.Yu. Zhdanova, A.N. Rabchevsky, E.G. Ashikhmin

ASSESSMENT AND MONITORING OF DESTRUCTIVE FORMS OF STUDENTS' BEHAVIOUR IN SOCIAL NETWORKS THROUGH AI

DOI: 10.35634/2412-9550-2025-35-1-48-59

The article considers the current problem of ensuring security in the university. The purpose of the study: to assess the involvement of students in destructive phenomena based on data obtained from open sources (social networks) using artificial intelligence.

Methods: from the space of the social network “VKontakte”, using the VK.BARKOV.NET system and the SEUS search engine, users of the VKontakte social network were selected who indicated Perm State National Research

University as their place of study (80,000 people). Then, using a neural network, information was searched for on the involvement of users in destructive social phenomena.

Results of the study: using an artificial intelligence system based on the analysis of profile data, the content of publications and groups in which this user is a member, one user was identified who is characterized by involvement in severe forms of destructive and self-destructive behavior; 3 users who have clear signs of a tendency to destructive phenomena of moderate severity. A comparative analysis of the results obtained based on the neural network analysis with the results of the analysis of markers identified by expert psychologists showed that the neural network identified 4 users, and the method based on markers selected by expert psychologists – 5 users.

Discussion of the results: the difference in the results is explained by the fact that the method used by artificial intelligence focused on involvement in destructive content, while the method of expert psychologists, in addition to destructive content, also included markers related to the mental state of students.

Conclusions: the approach described in the article allows us to assess the potential risks of an individual's involvement in destructive phenomena based on his digital footprint in social networks.

Keywords: security, destructive behavior, social networks, artificial intelligence, markers of destructive behaviour.

Received 22.07.2024

Zhdanova S.Yu., Doctor of Psychology, Professor, Head of the Department of Developmental Psychology,
Director of the Center for Psychological and Pedagogical Assistance of PSU

E-mail: Svetlanaur@gmail.com

Rabchevsky A.N., Candidate of Technical Sciences, Senior Lecturer,
Institute of Computer Science and Technology

E-mail: ran@psu.ru

Perm State University

Bukireva st., 15, Perm, Russia, 614990

Master of sociology, Postgraduate student in sociology

Perm National Research Polytechnic University

Komsomolskiy Prospekt, 29, Perm, Russia, 614990

E-mail: e.ashikhmin@icloud.com