

ИНФОРМАЦИОННЫЕ МОДЕЛИ АДАПТАЦИИ ЛИЧНОСТИ В СОЦИУМЕ

ИНФОРМАЦИОННЫЕ ВОЙНЫ С ИСПОЛЬЗОВАНИЕМ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: АНАЛИЗ ПСИХОЛОГИЧЕСКИХ МЕХАНИЗМОВ ВОЗДЕЙСТВИЯ¹

© 2019 г. Т.А. Нестик*, Е.А. Михеев**

**Доктор психологических наук, профессор РАН, заведующий лабораторией социальной
и экономической психологии,
Институт психологии Российской академии наук, Москва;
E-mail: nestik@ipras.ru*

***Соискатель, Институт психологии Российской академии наук, Москва,
E-mail: mih-news@mail.ru*

В статье рассматривается проблема механизмов психологического воздействия в ходе информационных войн с использованием технологий искусственного интеллекта. Обращается внимание на психологические эффекты групповой поляризации и подтверждения уже сложившегося мнения посредством «информационного джерримендеринга». Рассматриваются способы противодействия распространению «фейковых» новостей и дезинформации в социальных сетях. Приводятся результаты эмпирического исследования обеспокоенности использованием цифровых технологий для слежки за гражданами через интернет и манипулирования общественным мнением, а также результаты исследования предикторов, влияющих на озабоченность молодежи использованием систем искусственного интеллекта для манипулирования общественным мнением. Делается вывод о том, что использование только технических решений для защиты от дезинформации в условиях низкого социального доверия может приводить к обратным эффектам: способствовать развитию зависимости от цифровых технологий и избеганию самостоятельной оценки полученной информации, усилению подозрения

¹ Исследования, представленные в статье, проведены при поддержке грантов РФФИ №18-29-22046 мк и РФФИ №18-18-00439.

пользователей относительно возможности манипулирования общественным мнением с помощью цифровых технологий.

Ключевые слова: информационные войны, искусственный интеллект, психологический механизм, дезинформация, астротурфинг, групповая поляризация, информационный джерримендеринг, групповая идентификация, эхо-камеры.

«Фейковые» новости и дезинформация влияют на деятельность государственных организаций, в том числе обеспечивающих внутреннюю безопасность и общественный порядок. Так негативное информационно-психологическое воздействие может оказываться на руководителей полицейских структур, работающих в экстремальной обстановке (массовые беспорядки, террористические акты, межэтнические и межрелигиозные конфликты, наводнения, землетрясения) с целью искажения и дестабилизации профессиональных ценностей и мотивов, изменения профессиональной направленности и поведения сотрудников (Черненилов, 2011). Негативное информационно-психологическое воздействие влияет на общественное мнение граждан о работе правоохранительных органов и эффективность оценки деятельности полицейского ведомства населением. Источниками распространения лжи при этом могут выступать общественные организации, СМИ (традиционные и новые), незаконные вооруженные формирования, экстремистские и террористические организации, криминальные структуры и отдельные лица (Михеев, 2015).

«Фейковые» новости и дезинформация оказывают влияние не международные политические процессы. Например, они могут использоваться в рамках стратегических информационных операций для достижения геополитических целей. В частности, ложь используется для формирования в массовом сознании так называемых мифов и фантомов, которые под давлением социального конформизма могут некритично приниматься на веру и даже в дальнейшем использоваться, как аргументы для принятия тех или иных политических решений, с целью легитимизации в глазах общества какого-либо

политического проекта, направленного на достижение или ослабление «страны-мишени» (Николайчук, 2018).

Как было показано в ряде исследований, СМИ и общественное мнение, влияющие на позицию политиков, тесно связаны с готовностью к применению ядерного оружия. Возможность манипуляций с массовым сознанием способствует созданию такой ситуации, при которой в борьбе за власть оппозиционные или политические лидеры могут представить нанесение ядерного удара как вынужденной или относительно безопасной для населения страны выход из сложившегося положения (Журавлев, Нестик, Соснин, 2016). По мнению зарубежных исследователей, вполне вероятен сценарий, при котором на принятие решение политика в кризисных ситуациях могут оказать влияние дезинформационные компании в социальных сетях, например Twitter. В условиях стресса, ограниченного лимита времени, ограниченности доступа к альтернативным источникам и СМИ, актуализации инстинктивного и эмоционального мышления, информация из социальных медиа может способствовать созданию искаженных представлений о противнике и в конечном итоге принятию решения на использование ядерного оружия (Сопп, 2017). В качестве примера можно привести ситуацию, имевшую место в 2018 году в США, и связанную с сообщением о летящей в сторону Америки баллистической ракете, которое разместила в Интернет конгрессмен штата Гавайи Талси Габбард. В сообщении интернет-пользователей призывали как можно быстрее найти убежище и заверяли о том, что это не учебная тревога. Позднее конгрессмен опровергла информацию, а Тихоокеанское командование Соединенных Штатов сообщило о том, что сообщение было отправлено по ошибке (Никитина, 2018).

«Фейковые» новости и дезинформацию в своей деятельности нередко используют и террористические организации. Причем каналами распространения выступают, как созданные ими СМИ и Интернет-ресурсы, так и обычные СМИ и даже обычные пользователи социальных сетей. В частности, организация джихадистского толка «Исламское государство Ирака и Леванта» осуществляет информационно-

психологическое воздействие через созданную пропагандистскую структуру — медиа-агентство «Аль-Фуркан» (al-Furqan Institute for Media Production) (Сундиев, 2014). С помощью манипулятивных и суггестивных технологий осуществляются привлечение сторонников и вербовка новых членов террористических организаций. Эффективность применения таких методов в социальных сетях связывают с высокой степенью доверия населения к неофициальным источникам информации; привлечением аудитории путем нереального обещания решения любых проблем; формированием сетевых сообществ на основе эмпатии. Для молодежи фейковая информация может подаваться в популярных и понятных форматах, в виде демотиваторов, мемов, подкастов, коллажей, фотографий, генераторов изображений (Быкадорова, 2016).

Наконец, дезинформация в социальных сетях может наносить прямой ущерб конкретным людям и даже быть смертельно опасной. Примером может быть убийство, совершенное толпой в Индии, после того как жертву ошибочно обвинили через WhatsApp в похищении детей с целью продажи их органов (Phartiyal, 2018).

ПСИХОЛОГИЧЕСКИЕ МЕХАНИЗМЫ, ИСПОЛЬЗУЕМЫЕ В ИНФОРМАЦИОННЫХ ВОЙНАХ С ПРИМЕНЕНИЕМ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Воздействие дезинформации связано с механизмами нормативного влияния и конформности, воспринимаемого межличностного сходства, эмоционального заражения, а также эффектами каскадов доступности информации, групповой поляризации и подтверждения уже сложившегося мнения (Нестик, Журавлев, 2018; Михеев, Нестик, 2018). Еще одним механизмом воздействия является эффект множественности источников, который заключается в том, что различные аргументы от разных пользователей воспринимаются личностью как более весомые, чем один и тот же аргумент от разных источников или разные аргументы от одного пользователя (Bessi, 2015). В основе эффекта множественности источников лежат социально-психологические предпосылки: во-первых, большее количество материалов повышают

интерес у целевой аудитории и вероятность привлечения других потенциальных аудиторий; во-вторых, при необходимости большее количество сообщений может затмить другую информацию; в-третьих, множественность каналов распространения увеличивает эффективность усваивания информации; в-четвертых, получение информации из множества источников повышают уровень доверия к информации, особенно если адресант идентифицирует себя с источником информации (Carpenter, 2013).

Ряд психологических приемов и стратегий целенаправленного информационного воздействия на интернет-пользователей можно считать особенно распространенными. Во-первых, к таким приемам относится искусственная поляризация общественного мнения (Groenendyk, 2018; Iyengar, Krupenkin, 2018; Melki, Pickering, 2014; Prior, 2013). Во-вторых, использованием ложных аккаунтов от имени известных людей и организаций (Goga et al., 2015; Jung, 2011; Reznik, 2013). В-третьих, использование провокационного эмоционального контента, легко запоминающегося и воспроизводимого аудиторией (Aday, 2010; Gross, D'Ambrosio, 2004; Konijn, 2013; Zollo et al., 2015). В-четвертых, распространение конспирологических теорий, убеждающих интернет-пользователей в том, что ход событий контролируется немногочисленной группой злоумышленников (Jolley, Douglas, 2017; Lewandowsky et al., 2013; van der Linden, 2015). Наконец, широко используется защита от критики посредством дефамации и делегитимации оппонентов (Rinnawi, 2007; Lischka, 2017), а также с помощью целенаправленного провоцирования их травли в сети (Griffiths, 2014; McCosker, 2014; Thacker, Griffiths, 2012).

Еще одним механизмом социального воздействия в информационных войнах является так называемый «информационный джерримендеринг» (Stewart et al., 2019), феномен в социальных сетях, получивший название по аналогии с известным явлением в избирательных компаниях, когда избирательные округа определяются таким образом, чтобы одна из политических сил получила больше голосов. В одном из исследований на

математической модели и эксперименте с участием 2250 пользователей социальных сетей было показано, что даже немногочисленная группа сторонников той или иной точки зрения может так изменить каналы распространения информации в социальных медиа, что их взгляды будут казаться позицией большинства. При этом после того, как в подобное манипулирование общественным мнением вовлекаются множество других партий, эффект перестает действовать и воспользовавшаяся им группа пользователей теряет контроль над ситуацией.

Описанные выше механизмы делают пользователей более восприимчивыми к ложным сообщениям, поддерживающим их позитивную самооценку, в том числе через межгрупповое сравнение и аутгрупповую дискриминацию. Сложная информация в социальной сети нередко формирует желание минимизировать состояние психического дискомфорта, вызванное столкновением в сознании конфликтующих представлений («когнитивного диссонанса»). В результате пользователь перестает объективно воспринимать и обрабатывать информацию (Dietram, 2018). Усвоению аудиторией ложной информации также способствует таргетирование, то есть подготовка информации с учетом социально-психологических характеристик целевых аудиторий. В этих случаях используется эффект подтверждения, когда информация собирается и интерпретируется для защиты уже сложившегося убеждения (Kunda, 1990)

Распространению «фейковых новостей» и дезинформации в социальных сетях, в частности Твиттере способствуют боты, тролли и киборги (Hindman, 2018). Боты условно принято делить на «хороших» и «плохих», однако граница между ними расплывчата. К «хорошим» относятся чат-боты в популярных Интернет-мессенджерах (Skype, Viber, Telegram, WhatsApp и т.д.), которые занимаются поиском и предоставлением информации об услугах, товарах, местах и т.д. «Плохие» боты делятся на «простые» боты, распространяющие спам, ставящие лайки и делающие репосты, и «сложные» боты, которые используют Java-скрипты и внешние ресурсы, умеют автоматически искать необходимую информацию в Интернете, маскируют свое

происхождение, подвергая опасности сотни тысяч IP-адресов. Применительно к сфере СМИ боты делятся на «новостные» и «операционные» боты. Первый тип ботов способствует автоматическому информированию людей о событиях, второй тип используется для автоматизации регулярных организационных процессов. Существуют боты, сделанные на основе технологий искусственного интеллекта. Например, в рамках японского проекта «JX Press Corp» созданы боты, которые самостоятельно занимаются поиском, подготовкой и публикацией материалов по заданным параметрам на основе информации из Интернет (Webb, 2019).

При использовании ботов ручное управление человеком сочетается с автоматизацией процесса распространения информации с помощью технологий ИИ и машинного обучения. Подобные технологии активно применяются для создания так называемых «False Amplifiers» в ходе информационных операций в Facebook. Усиление лжи (False amplification) осуществляется по нескольким направлениям. Первое включает создание большого количества аккаунтов с небольшим количеством участников для размещения больших объемов информации. Второе направление включает создание небольшого числа аккаунтов, которые должны демонстрировать свою активность таким образом, чтобы не возникало сомнений в реалистичности пользователя самого аккаунта. Все эти аккаунты создают так называемую систему «фейковых» социальных взаимодействий и способствуют формированию искаженного общественного мнения (Weedon, 2017).

Информационные войны, основанные на использовании автоматизированных алгоритмов и систем искусственного интеллекта, опираются на три ключевых тактики (Memes That Kill..., 2018).

Во-первых, дипломатия и репутационные манипуляции: использование передовых технологий цифровой дезинформации в целях склонения противника к выгодным дипломатическим или военным решениям; или для того, чтобы выдавать себя за лидера или влиятельного лица, совершать от их лица незаконные действия.

Во-вторых, автоматизированный фишинг: таргетинговое применение вредоносных программ на основе искусственного интеллекта для подделки действий объектов, которые они бы на самом деле не совершали, в том числе действия по разглашению секретной информации.

В-третьих, компьютерная пропаганда: использование социальных медиа, человеческой психологии, слухов, сплетен и алгоритмов искусственного интеллекта для манипуляции общественным мнением.

Противодействие распространению «фейковых новостей» и дезинформации информации в сетевых сообществах в настоящее время опирается в основном на технические, а не психологические решения.

Типичная психологическая операция с использованием социальных сетей начинается с целенаправленного сбора информации, в том числе с применением фишинговых сайтов, взлома и перехвата аккаунтов, кражи личных данных. На втором этапе генерируется контент: распространяются истории для прессы, мемы, создаются фейковые аккаунты и боты. На третьем этапе происходит амплификация лжи, т.е. собственно астротурфинг: через созданные фейковые аккаунты и боты распространяются мемы, создаются вымышленные группы, осуществляется спам комментариев (Weedon et al., 2017).

Показательная в этой связи борьба с ложными сообщениями в Китае. Только в 2018 году в одной из самых крупных социальной сетей Поднебесной «WeChat» было заблокировано более 500 миллионов фейковых новостей. В борьбе за правду в стране задействовано 800 различных китайских ведомств. В 2015 году США в целях повышения общего уровня действий в сфере информационного противоборства были даже проведены совместные с ССО США (PsyOP) и ICONS специальные учения, направленные на пробацию в реальных условиях методов контрпропаганды ИГИЛ; интеграции последних нейрокогнитивных и социальных концепций для более точного выявления искомого контента с учётом современных информационно-коммуникационных

возможностей; анализа оперативной обстановки на поле боя с классификацией действующих там субъектов информационно-психологического противоборства как «нейтральных», «дружественных» и «вражеских» (Виловатых, 2018).

Администрация социальной сети Facebook заявила о том, что будет стремиться предлагать пользователям информацию на основе их предпочтений («осмысленных связей»), расставлять приоритеты в отображении содержимого, которым делятся друзья пользователей и члены их семей, а также уменьшит количество публичного контента в форме новостей и рекламы. В результате, по мнению ряда исследователей, подборка контента таким образом способствовала усилению собственных убеждений (Isaac, 2018), а в странах, где Facebook ввело эти новые функциональные возможности, это даже способствовало распространению «фейковых новостей» (Frankel, 2018).

Для выявления «дипфейка», то есть применения систем искусственного интеллекта для генерации ложного видеоконтента, неотличимого от реальности, используются все более сложные технические решения, основанные на глубоком обучении и нейросетях, требующих все больших массивов данных для обучающих выборок (Agarwal et al., 2019).

Чисто технические решения недостаточно эффективны, так как основные причины распространения ложной информации имеют психологический характер. На это указывает исследование, проведенное командой исследователей из Массачусетского университета, которая проанализировала 120 000 новостей, которыми 4,5 млн раз поделились 3 млн. пользователей Twitter в 2006 – 2017 гг. (Vosoughi et al., 2018). Сообщения были разделены на правдивые и ложные с опорой на 6 независимых фактчекинговых организаций. После выявления и устранения сообщений, сгенерированных ботами, оказалось, что ложные новости отличаются большей новизной, распространяются в 6 быстрее и с в 10 раз большим охватом, чем правда. Ложные новости были связаны с лингвистическими маркерами удивления, страха и отвращения в перепостах, а правдивые новости - с предвосхищением, досадой, радостью и доверием.

Исследование показало, что в основе распространения дезинформации в сети лежат психологические закономерности массового сознания, а не только злой умысел «фабрик троллей». Кроме того, воздействие дезинформации оставляет устойчивый след: сформированное с ее помощью убеждение трудно исправить последующим опровержением (Nyhan, Reifler, 2010; Lewandowsky et al., 2012). Иными словами, факт-чекинг и машинное научение недостаточны для противодействия информационным операциям в сети.

Есть примеры и другого подхода к борьбе с ложной информацией, основанного на использовании гуманитарных технологий, прежде всего механизмов социального научения. Так, экспериментально доказана эффективность специально разработанных онлайн-игр, повышающих сопротивляемость пользователей в отношении дезинформации. Разработчики игры предложили пользователям выступить в роли производителей ложных новостей и воспользоваться шестью типичными стратегиями манипулирования общественным мнением: сталкиванием мнений, использованием эмоционального контента, распространением конспирологических теорий, троллингом, использованием «козлов отпущения», а также использованием ложных аккаунтов от имени известных людей и организаций. При этом игроки получают информацию о том, почему эти стратегии действенны. Тестирование игры на 15 тысячах интернет-пользователях показало, что такого рода «прививка» повышает способность распознавать манипуляции в сети независимо от социально-демографических характеристик и идеологических убеждений игроков (Roosenbeek, Linden, 2019).

Общий уровень бдительности личности в отношении возможных манипулятивных воздействий в сети в значительной степени определяется осознанием рисков, связанных с информационными войнами, которые ведутся с применением автоматизированных систем и искусственного интеллекта.

СОЦИАЛЬНО-ПСИХОЛОГИЧЕСКИЕ ПРЕДПОСЫЛКИ ОЗАБОЧЕННОСТЬЮ ЛИЧНОСТИ ИСПОЛЬЗОВАНИЕМ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ИНФОРМАЦИОННЫХ ВОЙНАХ

Чтобы уточнить социально-психологические предпосылки озабоченности личности манипулированием общественным мнением с помощью цифровых технологий нами были проведены два исследования.

Первое из них было реализовано в рамках опроса, проведенного Институтом психологии РАН совместно с Исследовательской группой «ЦИРКОН» в сентябре 2019 г. (N=1600, 46% — мужчины; репрезентативная общероссийская выборка; средний возраст 44,6 лет, SD=16,9). За счет средств гранта РФФИ в анкету был включен вопрос о том, насколько респонденты озабочены использованием цифровых технологий для слежки за гражданами через интернет и манипулирования общественным мнением.

Таблица 1. Предпосылки озабоченности россиян использованием цифровых технологий для манипулирования общественным мнением (N=1600)

Предикторы	B	S.E.	β	t	Знач.
(Константа)	0,815	0,181		4,494	<0,001
Я чувствую глубокую связь между собой и всем остальным человечеством	0,058	0,024	0,071	2,436	0,015
Я горжусь тем, что являюсь гражданином России	-0,072	0,029	-0,079	-2,509	0,012
Я связываю свое будущее с судьбой своего города	-0,051	0,027	-0,059	-1,893	0,059
Я часто задумываюсь о будущем нашей страны	0,08	0,027	0,085	2,93	0,003
Сколько полных лет Вам исполнилось?	-0,008	0,002	-0,125	-4,731	<0,001
Озабоченность угрозой вооруженных конфликтов	0,155	0,056	0,11	2,775	0,006
Озабоченность экономическим упадком и криминализацией	0,18	0,066	0,105	2,726	0,007
Озабоченность угрозой вымирания нации и репрессий	0,335	0,042	0,267	7,989	<0,001
Озабоченность экологическими и природными рисками	0,126	0,045	0,089	2,79	0,005

Как показал линейный регрессионный анализ ($R=0,525$; $R^2=0,276$; $F=47,371$; $p<0,001$), беспокойство по поводу дезинформации в интернете тем выше, чем ниже возраст респондентов ($\beta=-0,125$), чем меньше они испытывают гордости за свою страну ($\beta=-0,079$), чем больше их тревожит будущее России ($\beta=0,079$), а также чем больше они озабочены другими рисками — вероятностью вооруженных конфликтов ($\beta=0,110$), экономическим упадком и криминализацией ($\beta=0,105$), вымиранием нации и репрессиями ($\beta=0,267$), экологическими и природными рисками ($\beta=0,089$).

Тот факт, что значимость угрозы использования цифровых технологий в информационных войнах характера для молодежи и связана с низкой положительной гражданской идентификацией, может указывать на два типа уязвимости личности в отношении информационных операций. Первый из них состоит в том, что внимание к рискам дезинформации снижается при сильной групповой идентификации. Второй связан с возможностью маскировать информационное воздействие или снижать сомнения в правдивости информации за счет использования негативной идентичности, ссылок на то, что это якобы «правда, противостоящая патриотической пропаганде». Первый тип уязвимости более характерен для людей пожилого возраста, идентифицирующих себя со своей страной и своим городом. Второй тип может быть более характерен для российской молодежи.

Второе исследование было проведено для того, чтобы выявить социально-психологические предикторы, влияющие на озабоченность молодежи использованием систем искусственного интеллекта для манипулирования общественным мнением. Участниками исследования стали студенты гуманитарных и технических ВУЗов гг. Москвы и Курска ($N=355$; 51% — мужчины; средний возраст — 21,9 лет). Для выявления опасений по поводу воздействия технологий ИИ на общество использовалась авторская анкета. Для измерения доверия использовались шкалы из World Values Survey, измеряющие социальное доверие ($N=1465$; 6 пунктов; $\alpha=0,719$; $M=2,94$; $SD=0,683$) и институциональное доверие ($N=1459$; 4 пункта; $\alpha=0,800$; $M=2,16$;

SD=0,821). Для измерения ценностных ориентаций нами использовались краткий «Портретный опросник ценностных ориентаций» Ш. Шварца (Ten-Item Values Inventory, TIVI; Sandy et al., 2016), а также «Опросник моральных оснований» (the Moral Foundations Questionnaire, MFQ) Дж. Хайдта в адаптации О.А. Сычева. Для измерения характеристик образа мира мы использовали опросник «Социальные аксиомы» (the Social Axioms Survey, SAS) М. Бонда и К. Леонга в адаптации А.Н. Татарко и Н.М. Лебедевой, а также шкалу «Общая вера в справедливость мира» К. Далберт в адаптации С.К.Нартовой-Бочавер.

Как показал линейный регрессионный анализ ($R=0,344$; $R^2=0,119$; $F=9,384$; $p<0,001$), беспокойство по поводу использования ИИ для создания и распространения ложной и порочащей информации в сети обратно связано с доверием к другим социальным группам ($\beta=-0,121$) и прямо связано с социальным цинизмом, то есть низким доверием к социальным институтам ($\beta=0,189$). Его повышают ориентация на гармонию в отношениях между людьми ($\beta=0,227$) и заботу о людях в качестве основания моральной оценки социальных явлений ($\beta=0,124$).

Таблица 2. Предпосылки озабоченности студентов использованием ИИ для создания и распространения ложной и порочащей информации в сети (N=355)

Предикторы	B	S.E.	β	t	Знач.
(Константа)	1,504	0,49		3,067	0,002
Аутгрупповое доверие (WVS)	-0,206	0,087	-0,121	-2,357	0,019
Социальный цинизм (SAS)	0,313	0,086	0,189	3,629	<0,001
Зависимость от судьбы (SAS)	-0,207	0,081	-0,14	-2,546	0,011
Шкала заботы (MFQ)	0,171	0,075	0,124	2,287	0,023
Ориентация на гармонию в отношениях между людьми (SAS)	0,365	0,09	0,227	4,064	<0,001

Низкое социальное доверие является предиктором опасений использования ИИ в информационных войнах. Проведенный нами линейный регрессионный анализ ($R=0,406$; $R^2=0,164$; $F=14,102$; $p<0,001$) показал, что беспокойство по поводу использования ИИ для манипуляции сознанием людей прямо связано с социальным

цинизмом ($\beta=0,176$) и обратно связано с технологическим оптимизмом ($\beta=-0,212$). Его повышают ориентация на гармонию в отношениях между людьми ($\beta=0,097$), доброжелательность ($\beta=0,125$), эмоциональная стабильность ($\beta=0,106$) и открытость новому опыту ($\beta=0,093$), а также чистота и святость в качестве основания моральной оценки социальных явлений ($\beta=0,158$).

Таблица 3. Предпосылки озабоченности студентов использованием ИИ для манипуляции сознанием людей (N=355)

Предикторы	B	S.E.	β	t	Знач.
(Константа)	1,55	0,748		2,072	0,039
Социальный цинизм (SAS)	0,301	0,085	0,176	3,53	<0,001
Шкала чистоты и святости (MFQ)	0,198	0,067	0,158	2,936	0,004
Открытость к новому опыту	0,1	0,054	0,093	1,854	0,065
Доброжелательность	0,147	0,061	0,125	2,434	0,015
Эмоциональная стабильность	0,094	0,044	0,106	2,127	0,034
Технооптимизм	-0,506	0,122	-0,212	-4,143	<0,001
Ориентация на гармонию в отношениях между людьми (SAS)	0,162	0,09	0,097	1,785	0,075

Использование линейного регрессионного анализа методом обратных шагов ($R=0,437$; $R^2=0,191$; $F=11,729$; $p<0,001$) позволило выделить социально-психологические предикторы поддержки использования ИИ для выявления «фейков» и блокирования распространения ложной информации в сети. Она более характерна для женщин, чем для мужчин ($\beta=0,155$) и тем выше, чем больше ориентация респондентов на уважение к авторитетам ($\beta=0,158$) и вера в справедливость мира ($\beta=0,239$). Другими ее предпосылками являются информированность о технологиях ИИ ($\beta=0,133$) и готовность их использовать ($\beta=0,223$). При этом поддержка применения систем ИИ для борьбы с дезинформацией обратно связана с аутгрупповым доверием ($\beta=-0,096$) и ориентацией на достижения ($\beta=-0,101$).

Полученные данные показывают, что недоверие к технологиям искусственного интеллекта тесно связано с недоверием социальным.

Таблица 4. Предпосылки желательности использования ИИ для выявления «фейков» и блокирования распространения ложной информации в сети (N=355)

Предикторы	B	S.E.	β	t	Знач.
(Константа)	0,557	0,711		0,783	0,434
Аутгрупповое доверие (WVS)	-0,26	0,135	-0,096	-1,932	0,054
Шкала уважения к авторитетам (MFQ)	0,28	0,092	0,158	3,046	0,002
Достижения	-0,191	0,092	-0,101	-2,068	0,039
Ваш пол	0,58	0,19	0,155	3,048	0,002
Информированность о развитии технических систем на основе искусственного интеллекта	0,132	0,065	0,103	2,046	0,042
Общая вера в справедливый мир	0,418	0,093	0,239	4,505	0
Готовность использовать технологии ИИ	0,597	0,133	0,223	4,472	0

Тот факт, что поддержка технологий искусственного интеллекта, призванных оградить интернет-пользователей от ложной информации, прямо связана с уважением к авторитетам, то есть с конформностью личности, может указывать на потенциальные риски применения таких технологий: сочетание низкой ориентации на достижения, высокой конформности и веры в справедливость мира будут приводить к эффектам избыточного доверия цифровым технологиям, на которые будет перекладываться ответственность за решения в сфере информационной безопасности.

Можно предположить, что в условиях краха доверия к социальным институтам использование технических решений для защиты от дезинформации будет приводить к обратным эффектам. У одних пользователей оно будет развивать зависимость от цифровых технологий и избегание самостоятельной оценки полученной информации. У других пользователей вместо повышения доверия к информации, оно будет усиливать подозрения в возможности манипулирования общественным мнением с помощью цифровых технологий.

В целом проведенный нами анализ научной литературы и данные наших собственных эмпирических исследований говорят о необходимости разработки цифровых социально-психологических технологий, направленных на повышение уровня социального доверия и устойчивости к попыткам манипулятивного воздействия. Системы

искусственного интеллекта могут сыграть в этих решениях важную роль. Но не за счет подмены человека в задачах выявления и блокирования дезинформации, а за счет формулирования вопросов к человеку, позволяющих повысить его рефлексивность и сделать его отношение к полученной информации более критическим.

ЗАКЛЮЧЕНИЕ

Таким образом, теоретический анализ показал наличие большого числа психологических эффектов, влияющих на дезинформирование и продвижение «фейковых» новостей, реализуемых в рамках информационных войн с использованием технологий искусственного интеллекта. Эти эффекты оказывают влияние на динамику, степень и характеристики усвоения деструктивной информации, способствуют повышению уровня манипулируемости массовым сознанием.

Противодействие распространению «фейковых новостей» и дезинформации в сетевых сообществах в настоящее время опирается в основном на технические решения. На наш взгляд, этого недостаточно, так как основные причины распространения ложной информации имеют психологический характер. В частности, исследования показывают, что в основе распространения дезинформации в сети лежат психологические закономерности массового сознания, а не только злой умысел «фабрик троллей». Также, воздействие дезинформации оставляет устойчивый след: сформированное с ее помощью убеждение трудно исправить последующим опровержением.

В ходе эмпирического исследования было обнаружено два типа уязвимости к методам информационных операций. У первого типа, представленного в основном людьми пожилого возраста, идентифицирующих себя со своей страной и своим городом, обнаружено снижение внимания к рискам дезинформации при сильной групповой идентификации. Второй тип, характерный для российской молодежи, связан с возможностью маскировать информационное воздействие или снижать сомнения в правдивости информации за счет использования негативной идентичности.

Проведенное нами эмпирическое исследование показало, что использование только технических решений для защиты от дезинформации будет приводить к обратным эффектам, способствовать развитию зависимости от цифровых технологий и избеганию самостоятельной оценки полученной информации, усилению подозрения пользователей относительно возможности манипулирования общественным мнением с помощью цифровых технологий.

СПИСОК ЛИТЕРАТУРЫ

- Быкадорова А., Чурилов С, Шаповалова Е.* Экстремизация молодежной аудитории сети интернет. Казань: Казанский педагогический журнал. №3, 2016. С. 181—182 .
- Виловатых А.В.* Меметика как инструмент современного информационного противоборства // Журнал «Проблемы национальной стратегии». Москва: РИСИ. 2018. №5(50). С. 141—154.
- Журавлев А.Л., Нестик Т.А., Соснин В.А.* Социально-психологические аспекты геополитической стабильности и ядерного сдерживания в XXI веке — М. : Изд-во ИП РАН, 2016.
- Михеев Е. А., Нестик Т. А.* Дезинформация в социальных сетях: состояние и перспективы психологических исследований // Социальная психология и общество. 2018. Т. 9. № 2. С. 5—20. doi:10.17759/sps.2018090201.
- Михеев Е.А.* Информационно-психологическое воздействие в механизме формирования общественного мнения о деятельности сотрудников органов внутренних дел Российской Федерации // Журнал "Закон и право", №8 М.: Изд-во "Юнити-Дана", 2015. С.161-163.
- Нестик Т.А., Журавлев А.Л.* Психология глобальных рисков. М.: Изд-во ИП РАН, 2018.
- Николайчук И.А., Янглева М.М., Якова Т.С.* Крылья Хаоса. Масс-медия, мировая политика и безопасность государства. М.: Изд-во ИКАР, 2018. 352 с.
- Никитина О.* Гавайи предупредили о летящей в сторону штата баллистической ракете. Газета «Взгляд», 13 января, 2018 г.. URL: <https://vz.ru/news/2018/1/13/903394.html>. (Дата обращения: 23.07.2019).

- Сундиев И.Ю., Смирнов А.А.* Медиаресурсы в экстремистской итеррористической деятельности // Свободная Мысль. - 2014. - № 4 (1646). - С. 55-72.
- Черненилов В. И.* Психология управления: учебник / под ред. В. И. Черненилова. М.: АУ МВД России, 2011. - 350 с.
- Aday, S.* (2010) Leading the charge: media, elites, and the use of emotion in stimulating rally effects in wartime. *Journal of Communication* 60(3):440–465. doi: 10.1111/j.1460-2466.2010.01489.x (Accessed: 13.11.2019).
- Agarwal, S., Farid, H., Gu, Y., He M., Nagano, K., Li H.* (2019) Protecting World Leaders Against Deep Fakes. *CVPR Workshops*. URL: <https://farid.berkeley.edu/downloads/publications/cvpr19/cvpr19a.pdf> (Accessed: 13.11.2019).
- Bessi, A., Petroni, F., Del Vicario, M. et al.* (2015) Viral misinformation: The role of homophily and polarization. *Proceedings of the 24th International Conference on World Wide Web Companion* 18 May 2015. 355-356. URL: <https://arxiv.org/abs/1411.2893> (Accessed: 13.11.2019).
- Carpenter, D., Ko M.* (2013) Online astroturfing: A theoretical perspective. *Proceedings of the Nineteenth Americas Conference on Information Systems, Chicago, Illinois, August 15-17*. URL: https://www.researchgate.net/profile/Darrell_Carpenter/publication/286729041 (Accessed: 22.09.2019).
- Conn, A.* (2017) Three Tweets to Midnight: Nuclear Crisis Stability and the Information Ecosystem/The Future of life Institute. November 15, 2017 URL.: <https://futureoflife.org/2017/11/15/three-tweets-midnight-nuclear-crisis-stability-information-ecosystem/?cn-reloaded=1> (Accessed 08.10.2019).
- Frankel, S., Fandos, N.* (2018) Facebook identifies new influence operations spanning globe. *The New York Times*. 21 August, <https://www.nytimes.com/2018/08/21/technology/facebook-political-influence-midterms.html>.
- Goga, O., Venkatadri, G., Gummadi, K.P.* (2015) The Doppelgänger Bot Attack: Exploring Identity Impersonation in Online Social Networks. In: *Proceedings of the Internet Measurement Conference. ACM, New York (IMC '15)*, 201. 141–153.
- Griffiths, M.D.* (2014) Adolescent trolling in online environments: a brief overview. *Education for Health* 32(3):85–87.

- Groenendyk, E.* (2018) Competing motives in a polarized electorate: political responsiveness, identity defensiveness, and the rise of partisan antipathy. *Political Psychology*, 39:159–171. doi: 10.1111/pops.12481.
- Gross, K., D'Ambrosio, L.* (2004) Framing emotional response. *Political psychology* 25(1):1–29.
- Jolley, D, Douglas, K.M.* (2017) Prevention is better than cure: addressing antivaccine conspiracy theories. *Journal of Applied Social Psychology*. doi: 10.1111/jasp.12453.
- Jung, A.M.* (2011) Twittering away the right of publicity: personality rights and celebrity impersonation on social networking websites. *Chicago-Kent Law Review* 86 (1):381–418.
- Konijn, E.A.* (2013) The role of emotion in media use and effects. In: Dill KE (ed.), *The Oxford Handbook of Media Psychology*. New York. London: Oxford University press, 186–211.
- Lewandowsky, S, Ecker, U.K., Seifert, C.M., Schwarz, N, Cook, J.* (2012) Misinformation and its correction: continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3):106–131. doi: 10.1177/1529100612451018.
- Lewandowsky, S, Gignac, G.E., Oberauer, K.* (2013) The role of conspiracist ideation and worldviews in predicting rejection of science. *PLOS ONE* 8(10):1–11. doi: 10.1371/journal.pone.0075637.
- Lischka, J.A. (2017) A Badge of Honor?, *Journalism Studies*, 20:2, 287-304. doi: 10.1080/1461670X.2017.1375385.
- Memes That Kill: The Future Of Information Warfare/CB Insights, 2018. URL: <https://www.cbinsights.com/research/future-of-information-warfare/> (Accessed:15.10.2019).
- McCosker, A.* (2014) Trolling as provocation: YouTube's agonistic publics. *Convergence: The International Journal of Research into New Media Technologies*, 20(2):201–217. doi: 10.1177/1354856513501413.
- Melki, M, Pickering, A.* (2014) Ideological polarization and the media. *Economic Letters*, 125 (1):36–39. doi: 10.1016/j.econlet.2014.08.008.

- Nyhan, B., Reifler, J.* (2010) When corrections fail: the persistence of political misperceptions. *Political Behavior*, 32(2):303–330. doi: 10.1007/s11109-010-9112-2.
- Phartiyal S., Patnaik S., Ingram D.* When a text can trigger a lynching: WhatsApp struggles with incendiary messages in India. Reuters UK, 25 June, 2018. <https://uk.reuters.com/article/us-facebook-india-whatsapp-fake-news/when-a-text-can-trigger-a-lynching-whatsapp-struggles-with-incendiary-messages-in-india-idUKKBN1JL0OW>.
- Prior, M.* (2013) Media and political polarization. *Annual Review of Political Science*, 16 (1):101–127. doi: 10.1146/annurev-polisci-100711-135242.
- Reznik, M.* (2013) Identity theft on social networking sites: developing issues of internet impersonation. *Touro Law Review* 29(2):455–484.
- Rinnawi, K.* (2007) De-legitimization of media mechanisms: israeli press coverage of the Al Aqsa Intifada. *International Communication Gazette*. 69(2):149–178. doi: 10.1177/1748048507074927.
- Roosenbeek, J., Linden, S.* (2019) Fake news game confers psychological resistance against online misinformation // *Palgrave Communications* 5, 65. doi:10.1057/s41599-019-0279-9.
- Stewart, A.J., Mosleh, M., Diakonova, M. et al.* (2019) Information gerrymandering and undemocratic decisions. *Nature* 573, 117–121. doi:10.1038/s41586-019-1507-6.
- Sundiev I. Yu., Smirnov A.A.* Media resources in extremist and terrorist activities: functional analysis/Website of the Autonomous non-profit organization "Center for interdisciplinary research" URL.: <http://spkurdyumov.ru/networks/mediaresursy-v-ekstremistskoj-i-terroristicheskoy-deyatelnosti/> (Accessed: 12.11.2019).
- Thacker, S., Griffiths, M.D.* (2012) An exploratory study of trolling in online videogaming. *International Journal of Cyber Behavior, Psychology and Learning*, 2(4):17–33. doi: 10.4018/ijcbpl.2012100102.
- Webb, A.* (2019) Trend Report For Journalism, Media & Technology/The Future Today Institute. 2019 p.31 URL.: <https://futuretodayinstitute.com/2019-journalism-media-tech-trends/> (Accessed 21.09.2019).

Weedon, J., Nuland, W., Stamos, A. (2017) Information Operations and Facebook[Электронный ресурс]. Version 1.0 edition. 13 p. URL: <https://ru.scribd.com/document/346741164/Facebook-and-Information-Operations-v1> (Accessed: 13.05.2019).

Van der Linden, S. (2015) The conspiracy-effect: exposure to conspiracy theories (about global warming) decreases pro-social behavior and science acceptance. *Personality and Individual Differences*. 87:171–173. doi: 10.1016/j.paid.2015.07.045.

Vosoughi, S., Roy, D., Aral, S. (2018) The spread of true and false news online. *Science*. 2018. 359(6380):1146–1151. doi: 10.1126/science.aap9559.

Zollo, F., Novak, P.K., Del Vicario, M., Bessi, A., Mozetič, I., Scala, A., Caldarelli, G., Quattrociocchi, W. (2015) Emotional dynamics in the age of misinformation. *PLoS ONE*. 10(9):e0138740. doi: 10.1371/journal.pone.0138740.

Статья поступила в редакцию: 20.11.2019. Статья опубликована: 30.12.2019.

INFORMATION WARS USING ARTIFICIAL INTELLIGENCE SYSTEMS: ANALYSIS OF PSYCHOLOGICAL MECHANISMS OF INFLUENCE²

T. A. Nestik *, E. A. Mikheev**

* Doctor of psychology, Professor of the Russian Academy of Sciences, head of the
laboratory of social and economic psychology,
Institute of psychology of the Russian Academy of Sciences, Moscow;
E-mail: nestik@ipras.ru

² The article was prepared with the support of the grant of the Russian Foundation for basic research No. 18-29-22046 MK and by a grant of the RSF No. 18-18-00439.

** Postgraduate student, Institute of psychology of the Russian Academy of Sciences,
Moscow,
E-mail: mih-news@mail.ru

The article deals with the problem of psychological mechanisms of influence in the course of information wars using artificial intelligence technologies. Attention is drawn to the psychological effects of normative influence, interpersonal attraction on the basis of perceived similarity, the creation of homogeneous communities and group polarization ("echo chambers"). Theoretical reasoning is supported by the results of domestic and foreign studies. The ways of countering the spread of "fake" news and misinformation in social networks are considered. The results of the empirical study concerns the use of digital technologies to spy on citizens via the Internet and manipulate public opinion, as well as the results of a study of predictors influencing the concern of youth use of artificial intelligence systems for the manipulation of public opinion.

Keywords: information warfare, artificial intelligence, psychological mechanism, disinformation, astroturfing, group polarization, information jerrymandering, group identification, echo chambers.

REFERENCES

- Bykadorova, A., Churilov, S., Shapovalova, E. (2016)* Ekstremizatsiya molodezhnoi auditorii seti Internet [Extremism of the Internet youth audience]. Kazan: Kazan pedagogical journal, №3. 181—182.
- Vilovatykh, A.V. (2018)* Memetika kak instrument sovremennogo informatsionnogo protivoborstva [Memetics as a tool of modern information warfare]. Problemy natsional'noi strategii. Moscow: Publ. RISI. 141—154.
- Zhuravlev, A.L. (2016)* Socio-psychological aspects of geopolitical stability and nuclear deterrence in the 21st century/Zhuravlev A.L., Nestik T.A., Sosnin V.A. Moscow, Publ Institut psikhologii RAN.
- Kotlyarova, V.V., Shubima, M.M., Sisoeva, O.N. (2016)* Youth extremism in social networks: specificity and theoretical understanding. Moscow: Alma mater (Vestnik vishey shkoli). № 5. URL.: http://www.kstu.kz/wp-content/uploads/docs/restricted/lib/periodic/Alma%20mater_2016_5_95.pdf(Accessed: 15.11.2019).
- Mikheev, E.A., Nestik, T.A. (2018)* Dezinformatsiya v sotsial'nykh setyakh: sostoyanie i perspektivy psikhologicheskikh issledovaniy [Disinformation in social networks: current

state and perspective research directions]. *Sotsial'naya psikhologiya i obshchestvo*. 9 (2). Moscow: Publ. «MGPPU», 5—20. doi: 10.17759/sps.2018090201. URL: http://psyjournals.ru/files/93632/sps_2018_n2_Miheeva_Nestik.pdf (Accessed: 23.07.2019).

Mikheev, E.A. (2015) Informatsionno-psikhologicheskoe vozdeistvie v mekhanizme formirovaniya obshchestvennogo mneniya o deyatel'nosti sotrudnikov organov vnutrennikh del Rossiiskoi Federatsii [Information and psychological impact in the mechanism of formation of public opinion on the activities of employees of the Russian interior Ministry]. *Zhurnal "Zakon i pravo"*, №8 М.: Publ. "Yuniti-Dana". 161-163.

Nestik, T.A., Zhuravlev, A.L. (2018) Psychology of global risks. Moscow, Publ Institut psikhologii RAN.

Nikolaichuk, I.A., Yanglyaeva, M.M., Yakova, T.S. (2018) Kryl'ya Khaosa. Mass-mediya, mirovaya politika i bezopasnost' gosudarstva [Mass media, world politics and state security]. Moscow, Publ. IKAR.

Nikitina, O. Hawaii warned about flying in the direction of the state of ballistic missile. *Vzglyad Newspaper*, 13 January 2018. URL: <https://vz.ru/news/2018/1/13/903394.html>. (Accessed: 23.07.2019).

Sundiev, I. Yu., Smirnov, A.A. (2014) Mediaresursy v jekstremistskoj iterroristicheskoy dejatel'nosti [Media Resources in Extremist and Terrorist Activity. Functional analysis] *Svobodnaja Mysl' [Free thought]* 4(1646). 55-72.

Chernenilov, V.I. (Ed.) (2011) Psikhologiya upravleniya: uchebnik [Management psychology: textbook]. Moscow, Publ AU MVD Rossii.

Aday, S. (2010) Leading the charge: media, elites, and the use of emotion in stimulating rally effects in wartime. *Journal of Communication* 60(3):440—465. doi: 10.1111/j.1460-2466.2010.01489.x (Accessed: 13.11.2019).

Agarwal, S., Farid, H., Gu, Y., He M., Nagano, K., Li H. (2019) Protecting World Leaders Against Deep Fakes. CVPR Workshops. URL: <https://farid.berkeley.edu/downloads/publications/cvpr19/cvpr19a.pdf> (Accessed: 13.11.2019).

Bessi, A., Petroni, F., Del Vicario, M. et al. (2015) Viral misinformation: The role of homophily and polarization. Proceedings of the 24th International Conference on World

Wide Web Companion 18 May 2015. 355-356. URL:
<https://arxiv.org/abs/1411.2893> (Accessed: 13.11.2019).

Carpenter, D., Ko M. (2013) Online astroturfing: A theoretical perspective. Proceedings of the Nineteenth Americas Conference on Information Systems, Chicago, Illinois, August 15-17. URL: https://www.researchgate.net/profile/Darrell_Carpenter/publication/286729041 (Accessed: 22.09.2019).

Conn, A. (2017) Three Tweets to Midnight: Nuclear Crisis Stability and the Information Ecosystem/The Future of life Institute. November 15, 2017 URL.: <https://futureoflife.org/2017/11/15/three-tweets-midnight-nuclear-crisis-stability-information-ecosystem/?cn-reloaded=1>(Accessed 08.10.2019).

Frankel, S., Fandos, N. (2018) Facebook identifies new influence operations spanning globe. The New York Times. 21 August, <https://www.nytimes.com/2018/08/21/technology/facebook-political-influence-midterms.html>.

Goga, O., Venkatadri, G., Gummadi, K.P. (2015) The Doppelgänger Bot Attack: Exploring Identity Impersonation in Online Social Networks. In: Proceedings of the Internet Measurement Conference. ACM, New York (IMC '15), 201. 141–153.

Griffiths, M.D. (2014) Adolescent trolling in online environments: a brief overview. *Education for Health* 32(3):85–87.

Groenendyk, E. (2018) Competing motives in a polarized electorate: political responsiveness, identity defensiveness, and the rise of partisan antipathy. *Political Psychology*, 39:159–171. doi: 10.1111/pops.12481.

Gross, K., D'Ambrosio, L. (2004) Framing emotional response. *Political psychology* 25(1):1–29.

Jolley, D, Douglas, K.M. (2017) Prevention is better than cure: addressing antivaccine conspiracy theories. *Journal of Applied Social Psychology*. doi: 10.1111/jasp.12453.

Jung, A.M. (2011) Twittering away the right of publicity: personality rights and celebrity impersonation on social networking websites. *Chicago-Kent Law Review* 86 (1):381–418.

- Konijn, E.A.* (2013) The role of emotion in media use and effects. In: Dill KE (ed.), *The Oxford Handbook of Media Psychology*. New York. London: Oxford University press, 186–211.
- Lewandowsky, S, Ecker, U.K., Seifert, C.M., Schwarz, N, Cook, J.* (2012) Misinformation and its correction: continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3):106–131. doi: 10.1177/1529100612451018.
- Lewandowsky, S, Gignac, G.E., Oberauer, K.* (2013) The role of conspiracist ideation and worldviews in predicting rejection of science. *PLOS ONE* 8(10):1–11. doi: 10.1371/journal.pone.0075637.
- Lischka, J.A.* (2017) A Badge of Honor?, *Journalism Studies*, 20:2, 287-304. doi: 10.1080/1461670X.2017.1375385.
- Memes That Kill: The Future Of Information Warfare/CB Insights, 2018. URL: <https://www.cbinsights.com/research/future-of-information-warfare/> (Accessed:15.10.2019).
- McCosker, A.* (2014) Trolling as provocation: YouTube’s agonistic publics. *Convergence: The International Journal of Research into New Media Technologies*, 20(2):201–217. doi: 10.1177/1354856513501413.
- Melki, M, Pickering, A.* (2014) Ideological polarization and the media. *Economic Letters*, 125 (1):36–39. doi: 10.1016/j.econlet.2014.08.008.
- Nyhan, B, Reifler, J.* (2010) When corrections fail: the persistence of political misperceptions. *Political Behavior*, 32(2):303–330. doi: 10.1007/s11109-010-9112-2.
- Phartiyal S., Patnaik S., Ingram D.* When a text can trigger a lynching: WhatsApp struggles with incendiary messages in India. Reuters UK, 25 June, 2018. <https://uk.reuters.com/article/us-facebook-india-whatsapp-fake-news/when-a-text-can-trigger-a-lynching-whatsapp-struggles-with-incendiary-messages-in-india-idUKKBN1JL0OW>.
- Prior, M.* (2013) Media and political polarization. *Annual Review of Political Science*, 16 (1):101–127. doi: 10.1146/annurev-polisci-100711-135242.
- Reznik, M.* (2013) Identity theft on social networking sites: developing issues of internet impersonation. *Touro Law Review* 29(2):455–484.

- Rinnawi, K.* (2007) De-legitimization of media mechanisms: israeli press coverage of the Al Aqsa Intifada. *International Communication Gazette*. 69(2):149–178. doi: 10.1177/1748048507074927.
- Roozenbeek, J., Linden, S.* (2019) Fake news game confers psychological resistance against online misinformation // *Palgrave Communications* 5, 65. doi:10.1057/s41599-019-0279-9.
- Stewart, A.J., Mosleh, M., Diakonova, M. et al.* (2019) Information gerrymandering and undemocratic decisions. *Nature* 573, 117–121. doi:10.1038/s41586-019-1507-6.
- Sundiev, I. Yu., Smirnov, A.A.* Media resources in extremist and terrorist activities: functional analysis/Website of the Autonomous non-profit organization "Center for interdisciplinary research" URL.: <http://spkurdyumov.ru/networks/mediaresursy-v-ekstremistskoj-i-terroristicheskoy-deyatelnosti/> (Accessed: 12.11.2019).
- Thacker, S., Griffiths, M.D.* (2012) An exploratory study of trolling in online videogaming. *International Journal of Cyber Behavior, Psychology and Learning*, 2(4):17–33. doi: 10.4018/ijcbpl.2012100102.
- Webb, A.* (2019) Trend Report For Journalism, Media & Technology/The Future Today Institute. 2019 p.31 URL.: <https://futuretodayinstitute.com/2019-journalism-media-tech-trends/> (Accessed 21.09.2019).
- Weedon, J., Nuland, W., Stamos, A.* (2017) Information Operations and Facebook[Электронный ресурс]. Version 1.0 edition. 13 p. URL: <https://ru.scribd.com/document/346741164/Facebook-and-Information-Operations-v1> (Accessed: 13.05.2019).
- Van der Linden, S.* (2015) The conspiracy-effect: exposure to conspiracy theories (about global warming) decreases pro-social behavior and science acceptance. *Personality and Individual Differences*. 87:171–173. doi: 10.1016/j.paid.2015.07.045.
- Vosoughi, S., Roy, D., Aral, S.* (2018) The spread of true and false news online. *Science*. 2018. 359(6380):1146–1151. doi: 10.1126/science.aap9559.
- Zollo, F., Novak, P.K., Del Vicario, M., Bessi, A., Mozetič, I., Scala, A., Caldarelli, G., Quattrocioni, W.* (2015) Emotional dynamics in the age of misinformation. *PLoS ONE*. 10(9):e0138740. doi: 10.1371/journal.pone.0138740.

Библиографическая ссылка на статью:

Нестик Т.А., Михеев Е.А. Информационные войны с использованием систем искусственного интеллекта: анализ психологических механизмов воздействия // Институт психологии Российской академии наук. Организационная психология и психология труда. 2019. Т. 4. № 4. С. 148-174.

Nestik T.A., Mikheev E.A. (2019). Informacionnyye vojny s ispol'zovaniem sistem iskusstvennogo intellekta: analiz psihologicheskikh mehanizmov vozdejstvija [Information wars using artificial intelligence systems: analysis of psychological mechanisms of influence]. *Institut Psikhologii Rossiyskoy Akademii Nauk. Organizatsionnaya Psikhologiya i Psikhologiya Truda* [Institute of Psychology of the Russian Academy of Sciences. *Organizational Psychology and Psychology of Labor*], 4 (4), 148-174.

Адрес статьи: <http://work-org-psychology.ru/engine/documents/document511.pdf>