

Исследование выполнено при частичной финансовой поддержке РФФИ в рамках научного проекта № 18-29-22104

Литература:

1. *Kulagina I., Iskhakov A.* Problems of Automation of the Aggression Analysis in Socio-Cyberphysical Environment / Proceedings of the 8th Scientific Conference on Information Technologies for Intelligent Decision Making Support (ITIDS 2020). – Advances in Intelligent Systems Research. – 2020. – Volume 174. – P. 35-40.

2. *Охапкин В.П., Охапкина Е.П., Исхакова А.О., Исхаков А.Ю.* Деструктивное информационно-психологическое воздействие в социальных сетях // Моделирование, оптимизация и информационные технологии. – 2020. – №8(1). – С. 1-14.

3. *Новиков Д.А.* Теория управления организационными системами. – М.: Издательство физико-математической литературы, 2012. – 604 с.

Боресков Г.К.

Этические аспекты применения инструментов искусственного интеллекта для обеспечения пространства доверия в электронных СМИ

Аннотация: В работе рассматриваются специфические риски этического характера, связанные с применением инструментов искусственного интеллекта для обеспечения безопасного коммуникационного пространства доверия на площадках электронных СМИ.

Ключевые слова: социальные сети, электронные СМИ, инструменты искусственного интеллекта, безопасность, доверие

За последнее десятилетие рынок телевизионного вещания в Российской Федерации претерпел значительные изменения. Телевизионные информационные программы теряют привлекательность для современной аудитории, в то же время существенно возрастает популярность разного рода интернет-сервисов. Эту тенденцию отражают, например, результаты

исследования фонда Общественное мнение «Допускает ли ТВ-аудитория переключение на другие источники информации?» [1].

В ситуации стремительного развития социальных сетей, миграции аудитории от телевидения в сторону онлайн потребления контента «по запросу» и технического проникновения широкополосного интернета во многие отдаленные регионы страны одним из наиболее перспективных направлений развития вещания становится формирование трансмедийных суперплатформ (далее – ТМСП), способных охватить как все доступные среды современной медиакommunikации, так и большинство социально-демографических слоев населения и доставить контент до пользователя, где бы он не предпочитал получать его в эпоху гибридного телесмотрения. Эффективность такого подхода подчеркивается в работе Дженкинса, Форда и Грина «Spreadable media: Creating value and meaning in a networked culture» [2].

В силу высокой конкуренции поставщиков контента как в сфере телевидения, так и в сферах социальных сетей и мобильных платформ, вещание ТМСП должно максимально соответствовать явным и неявным запросам и предпочтениям аудитории, не только оперативно адаптироваться к возникающим информационным трендам, но и целенаправленно порождать их, не искать свою зрительскую нишу, а постоянно расширять ее, одновременно привлекая новых зрителей и удерживая имеющихся.

Обеспечить это возможно с использованием модулей искусственного интеллекта, функционал которых предусматривает:

- выявление нарождающихся информационных трендов, подготовку рекомендаций по формированию контента, направленного на расширение аудитории на волне тренда;
- анализ представительства аудитории в социальных сетях, сегментирование и кластеризацию, выявление лидеров мнений и горизонтальных связей;
- выявление факторов, определяющих реакцию аудитории на демонстрируемый контент, дающее возможность существенно повысить востребованность контента, проектируя его в соответствии с этими факторами, отражающими предпочтения аудитории;

- выявление связей между характеристиками просматриваемого контента и динамикой мнений и предпочтений зрителей.

При этом адресное донесение до пользователя ТМСП востребованного им информационного контента будет эффективным только при условии доверия пользователя к этой платформе, обеспечение которого является еще одним важным направлением применения инструментов искусственного интеллекта в сфере трансмедийного вещания. Традиционно используемые для этого инструменты, основанные на объединении технологий модерации и краудсорсинга исследуются в книге Тарлтона Гиллеспи «Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media» [3]. Эти инструменты имеют ряд существенных ограничений, включая:

- риски предвзятых, либо тенденциозных действий модераторов, разрушающих доверие участников к коммуникационному пространству – эта тенденция стала основной темой статьи Джоан Донаван «Why social media can't keep moderating content in the shadows» [4];

- высокую сложность идентификации целенаправленно спланированных масштабных манипулятивных информационных воздействий среди множества случайных «вбросов» обычных пользователей без применения специального инструментария, функционирующего на постоянной основе в автоматическом режиме;

- невозможность без такого инструментария определить потенциал распространения и конечный эффект манипулятивных воздействий.

Таким образом, ключевым инструментом обеспечения пространства доверия в информационной среде ТМСП должна стать информационная система, способная в режиме автоматизированного мониторинга средств массовой электронной коммуникации выявлять манипулятивные информационные воздействия, прогнозировать развитие линии распространения информации и определять источники и «слабые точки» таких линий, что создает возможности не только управляемого купирования негативных воздействий, но и своевременного

принятия этических и не вызывающих отторжения у участников информационной среды превентивных мер.

Эта система должна объединять в себе автоматизированные средства мониторинга соцсетей, анализа больших данных и выявления информационных воздействий на основе алгоритмов искусственного интеллекта, решая задачи:

- многовекторного анализа эмоционального контекста социальных сетей;
- выявления и типизации линий воздействия;
- предотвращения информационных угроз и реагирования на случаи их реализации.

Важно не забывать, что пространство доверия может обеспечиваться не только нивелированием случайных и целенаправленных манипулятивных воздействий, но и путем расширения доступа его участников к достоверной информации, предоставления им инструментов ее структурирования и достижения консенсуса, для чего также могут быть применены инструменты искусственного интеллекта. Этот подход предложен в статье А.Н. Райкова «Accelerating technology for self-organising networked democracy» [5]. При этом соблюдение этических норм, очевидно, становится одним из ключевых факторов для создания среды доверия в медиа-пространстве.

Это ставит ряд задач в сфере этических аспектов применения искусственного интеллекта в медиа-среде, среди которых:

- разработка методик классификации этических рисков при применении инструментов искусственного интеллекта в медиа-среде, их идентификации и предотвращения;
- прогнозирование последствий разработки, внедрения и использования инструментов искусственного интеллекта в сфере электронных средств массовой информации и массовой коммуникации;
- определение базовых ценностей и этических ориентиров для различных акторов искусственного интеллекта в медиа-среде.

В качестве одного из базовых ориентиров для этой работы может быть взято «Предварительное исследование возможности подготовки нормативного акта по вопросам этики применения искусственного интеллекта», подготовленное Рабочей группой КОМЕСТ [6] – всемирная комиссия по этике научных знаний и

технологий КОМЕСТ является консультативным органом и форумом для размышлений, учрежденным ЮНЕСКО в 1998 г. В документе подчеркивается: «Искусственный интеллект играет все более важную роль в обработке, структурировании и предоставлении информации. Автоматизированная журналистика и алгоритмические новостные ленты в социальных сетях – лишь некоторые из примеров этой тенденции, в связи с которой возникают вопросы доступа к информации, дезинформации, дискриминации, свободы выражения мнений, неприкосновенности частной жизни, а также медийной и информационной грамотности».

Мы можем выделить ряд этических рисков, специфических именно для сферы применения инструментов искусственного интеллекта в СМИ:

- усугубление поляризации мнений в результате формирования алгоритмами медиа-среды так называемых «фильтрующих пузырей» и «эхо-камер». Пользователи, попадающие в один и тот же «пузырь», могут не только подвергаться воздействию потока единообразным образом отфильтрованной информации, но также коммуницировать в основном друг с другом. Это приводит к заполнению медиа-среды все более однородными группами сторонников того или иного мнения, занимающими при этом все более полярные и радикальные позиции по отношению друг к другу. В ряде случаев развитие подобных тенденций может привести даже к серьезной дестабилизации общества, например, подорвав доверие к обнаруженным результатам выборов у пользователей, необоснованно экстраполировавших предпочтения своих «друзей» в социальных сетях (в большинстве, близких им по политическим взглядам) на всех избирателей. Вопросы политической поляризации, в том числе, как результат взаимодействий в социальных сетях подробно рассматриваются в работе «Political sectarianism in America» [7];

- тенденциозность и предвзятость при модерации пользовательского контента, приоритизации результатов поисковой выдачи и формировании подборок новостей, что может повлечь обвинения в автоматизированной цензуре и неправомерных ограничениях свободы слова;

- превращение алгоритмов, производящих отбор и первичную обработку информационных материалов для редакции современного СМИ, из инструмента, кардинально облегчающего и ускоряющего выполнение рутинных операций, в своего рода фильтр, искажающий реальную картину мира для самих сотрудников редакции. Это можно рассматривать как более радикальный вариант «фильтрующего пузыря»;

- расфокусировка редакционной политики и деградация имиджа СМИ в результате «гиперадаптации» контента к актуальным информационным трендам и предпочтениям аудитории, выявленным с помощью инструментов искусственного интеллекта.

Речь идет о том, что стремление редакции любого СМИ к расширению аудитории может быть существенно усилено рекомендациями алгоритмов искусственного интеллекта.

Литература:

1. Фонд «Общественное мнение». Допускает ли ТВ-аудитория переключение на другие источники информации? 06 Февраля 2021 – URL: <https://fom.ru/SMI-i-internet/14536> (дата обращения 15.10.2021).

2. *Jenkins H., Ford S., Green J.* Spreadable media: Creating value and meaning in a networked culture. – N.Y.: New York University Press, 2013. – 352 p.

3. *Tarleton Gillespie.* Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media. – New Heaven, CT: Yale University, 2018. – 298 p.

4. *Joan Donovan.* Why social media can't keep moderating content in the shadows // MIT Technology Review. – November 6, 2020.

5. *Raikov A.* Accelerating technology for self-organising networked democracy // Futures. – October 2018. – Volume 103. – P. 17-26. – URL: <https://www.sciencedirect.com/science/article/abs/pii/S0016328717302380> (дата обращения 14.10.2021).

6. COMEST Extended Working Group on Ethics of Artificial Intelligence. Preliminary study on the Ethics of Artificial Intelligence (COMEST) (26 February 2019 Paris). – URL: <https://unesdoc.unesco.org/ark:/48223/pf0000367823> (дата обращения 15.10.2021).

7. *Eli J. Finkel, Christopher A. Bail, Mina Cikara, Peter H. Ditto, Shanto Iyengar, Samara Klar, Lilliana Mason, Mary C. McGrath, Brendan Nyhan, David G. Rand, Linda J. Skitka, Joshua A. Tucker, Jay J. Van Bavel, Cynthia S. Wang, James N. Druckman.* Political sectarianism in America // *Science*. – 30 Oct 2020. – Vol. 370. Issue 6516. – P. 533-536.

Охалкина Е.П.

Разработка динамической системы функционирования сообществ социальной сети

Аннотация: Протекающие в сообществе социальной сети процессы показаны в виде системы дифференциальных уравнений первого порядка. Рассмотрена имитационная модель построенной исследованной динамической системы при различных начальных условиях имитации.

Ключевые слова: дифференциальные уравнения, социальные сетевые сервисы, системная динамика, информационно-коммуникационные технологии, имитационное моделирование

Примем, что в рамках моделирования предметной области для системы сообществ социальной сети (далее ССС) определены три обязательных параметра:

1. L – общее количество отметок «нравится» (так называемых «лайков» ('like')) на внутреннем сленге аудитории сообщества социальной сети) за анализируемый период времени t .
2. P – общее количество подписчиков (постоянных читателей) для ССС за анализируемый период времени t .
3. S – общее количество сообщений в ССС вне зависимости от адресатов и получателей за анализируемый период времени t .

Каждый из приведенных параметров участвует в процессах системы ССС. Таким образом, для каждого параметра может быть определен закон его изменения, который в рамках математической постановки будет представлен характеристическим уравнением.

Пусть l_i – количество отметок типа 'like' для одной произвольной записи («поста») сообщества социальной сети. Тогда