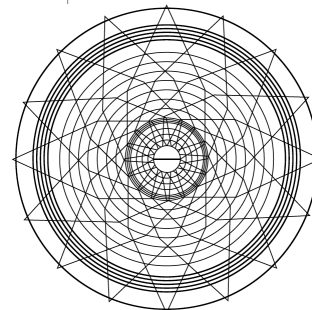


[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики



РОБОТИЗАЦИЯ НОВОСТНОЙ ЖУРНАЛИСТИКИ В РОССИИ: НОВЫЕ ПРИНЦИПЫ РАБОТЫ РЕДАКЦИИ И ТРАНСФОРМАЦИЯ ЖУРНАЛИСТСКОЙ ЭТИКИ

Исаев Е.М.

Преподаватель департамента медиа НИУ ВШЭ
(Москва, Россия)
emisaev@hse.ru

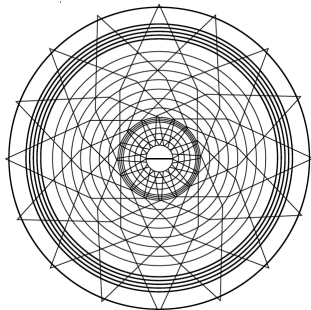
Кокорева М.В.

Корреспондент РБК (Москва, Россия)
man-kokoreva@yandex.ru

Аннотация:

Данное исследование посвящено описанию процессов трансформации журналистской этики (в частности проблемы факт-чекинга и ответственности за материалы) в процессе роботизации профессии в России при помощи анализа редакционной политики изданий, использующих алгоритмы для создания контента СМИ, а также ряда экспертных интервью с респондентами, которые столкнулись или сталкиваются с процессами роботизации СМИ в рамках своей сферы деятельности. Несмотря на потенциальные выгоды медиаиндустрии от внедрения алгоритмов в работу редакции, а также неминуемый пересмотр стандартов журналистской этики, российские медиакомпании с недоверием и скептицизмом относятся к данному направлению, а ответственность остается на плечах журналиста не только из-за уязвимости в российских реалиях алгоритма, который разработчики пока не могут довести до релевантного уровня выполнения работы, и недоверия журналистов к направлению, но и из-за отношения к алгоритму как к инструменту, с помощью которого человек может упрощать свою работу.

Ключевые слова: роботизированная журналистика, журналистика данных, журналистская этика, конвергентная журналистика, прецизионная журналистика, алгоритмическая журналистика



[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики

Применение алгоритмов в работе журналистов становится повседневной практикой в редакциях по всему миру. Среди западных и российских СМИ, использующих алгоритмы для создания новостного контента, — Los Angeles Times («Лос-Анджелес Таймс»), Associated Press («Ассошиэйтед Пресс»), Forbes («Форбс»), Mittmedia («Митмедиа»), «Яндекс для медиа», ТАСС и Sports.ru («Спортс.ру»). Например, международное информационное агентство «Ассошиэйтед пресс» по состоянию на 2017 год ежеквартально выпускает более 3700 (Marconi et al. 2017) автоматизированных новостей. К списку тем, отбираемых роботами, относятся спорт, финансы, погода или состояние на дорогах — информация, широко доступная в мониторинге и относительно несложная в ее дальнейшем анализе.

Несмотря на то, что в западной научной литературе тема роботизированной журналистики и применения направления в работе журналистов раскрыта достаточно полно, российские исследователи обращаются к ней мало (см., например, (Иванов 2015, 2017)). А поскольку алгоритмическая журналистика уже применяется в российских редакциях, вопрос о трансформации журналистской этики в связи с процессом роботизации профессии представляется особенно интересным для изучения в рамках данной работы. Так, мы ставим цель выявить, происходят ли данные изменения, и если да, то как именно, на примере ответственности редакций за выпускаемый алгоритмами материал. Для этого мы в том числе поставили перед собой задачу концептуализировать понятие роботизированной журналистики с учетом особенностей направления в России, определить термин «этика роботизированной журналистики», а также описать проблемы факт-чекинга (проверки информации) и ответственности журналиста в направлении роботизированной журналистики.

Автоматизированное создание контента и история технологии в работе журналиста

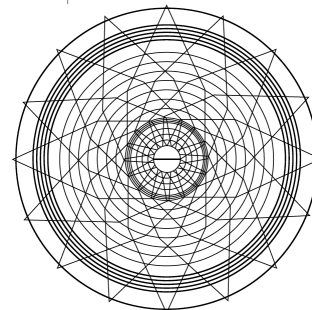
В работе «Законы медиа» (McLuhan et al. 1992) приведена достаточно широкая трактовка термина «медиа» Маршалла Маклюэна — это все, что создано человеком, даже идеи (например, законы физики), и все то, где прослеживается взаимодействие человека. В том числе медиа М. Маклюэн предлагает рассматривать (McLuhan 2002) как инструмент — продолжение человека в пространстве, усилитель человеческих возможностей, который приносит глубокие и длительные изменения в жизни людей.

Одно из самых спорных утверждений ученого — «средство коммуникации есть сообщение» — базируется на представлении о том, что именно тип коммуникации меняет среду, в которой живут люди, и их сознание, а не суть сообщения, которое становится второстепенным аспектом, и, в некотором смысле, определяется формой. С другой стороны, при создании идеи, новаторы не обращают внимание на контекст и лишь со временем осознают последствия нововведений (Federman 2004).

[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики

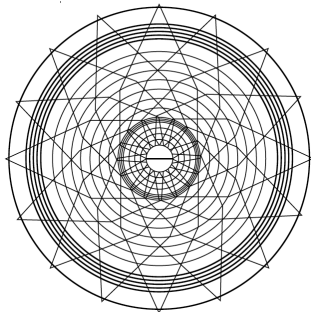


Таким образом, при неизбежном изменении технологии, в том числе и способов коммуникации в медиа, будет меняться не только содержание сообщений, но и, возможно, этические координаты в рамках профессиональной деятельности журналиста. Мы приходим к тому, что алгоритм, как и средства коммуникации, создает смысл. Человек должен обращать внимание на этот феномен, контролировать его, несмотря на то, что во многих странах создание контента и содержания новыми инструментами цифровой журналистики находится в зачаточном состоянии.

Для настоящего исследования мы изучали поздний период развития технологий: начиная с 1970-х годов, с появления инновационной платформы и «конвергентной журналистики», которая, в свою очередь, стала толчком для появления роботизированной журналистики. Как пояснял автор термина, профессор Массачусетского технологического института Итиэль де Сола Пул (de Sola Pool 1987), «конвергенция» — это стирание границ между медиа. Причем стирание таких границ происходит не только при выходе сообщения с помощью средств коммуникации, но и при сборе информации и производстве медиа (Кирия 2010). Если раньше СМИ содержали в себе один способ коммуникации (например, газета — визуальный), то с появлением телевидения стали возможными их комбинации, а с развитием Интернета и компьютерных технологий появилась возможность синтезировать типы коммуникаций (например, видео в материале сопровождает текст, поддерживается фотогалереей или предоставляет новые детали в истории, являясь логическим продолжением текста).

С 1970-х появляется «компьютерная отчетность» (computer-assisted reporting — CAR). Как писала исследователь Кейтлин Петре (Petre 2013), в 70-ые журналисты стали больше работать с аналитикой, предоставляя таким образом «более открытую и объективную» информацию для читателей. В 1973 году одним из первых выделивший особенность компьютеров в работе журналистов исследователь Филипп Мейер предложил (Meyer 2002) термин «прецизионная журналистика» (Precision Journalism). Прецизионная журналистика использует социологические методы (как правило, опросы и контент-анализ) для получения более точных ответов на журналистские вопросы, обработка таких ответов осуществлялась с помощью компьютерных технологий. Пик ее развития пришелся на конец 1980-х — начало 1990-х годов, в начале 2000-х годов «компьютерная отчетность» стала одним из вспомогательных элементов в журналистских расследованиях, задав курс на развитие сбора данных и статистического анализа — дата-журналистики (Data Journalism).

Дата-журналистика появилась в начале 2000-х годов (Coddington 2014), однако некоторые исследователи выделяют (Gray et al. 2012) небольшой временной промежуток между появлением прецизионной журналистики и дата-журналистики, отмечая начало последней в районе 2010-х. По словам теоретика Рика Гордона (Gordon 2013), такая журналистика требует три ключевых навыка: поиск информационного повода и трендов в полученной информации, навык



[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики

программирования и визуализации информации (посредством инфографики, таблиц, рисунков). Саму дата-журналистику можно определить так: «Это сбор, фильтрация и визуализация данных с целью более глубокого понимания того, что происходит вокруг нас и как оно может повлиять на общество», — то есть ее цель не рассказывать историю быстрее всех, а объяснять, что означает ход тех или иных событий. Важно отметить, что это фильтрация информации и анализ большого объема данных. В первую очередь, дата-журналистика помогает экономить журналисту время.

В 2010-х появилась «компьютерная/алгоритмическая журналистика» (computational/algorithmic journalism), объединившая в себе черты прецизионной и дата-журналистики, то есть такой формат можно рассматривать как комбинацию алгоритмов, объема данных, опыта социологической науки и методологии для развития «подотчетности» такой журналистики. Исследователь Николас Дикаполус рассматривает (Diakopoulos 2014) алгоритмическую журналистику как инструментарий, способный не только собирать и структурировать информацию, но и выделять инфоповод, генерировать текст и выпускать его на информационные ленты/сайт. С таким определением направления согласился и исследователь Марк Коддингтон (Coddington 2014), подчеркивая, что алгоритмическая журналистика сфокусирована на применении вычислительного мышления к практике сбора информации, поиска смыслов и предоставления контента.

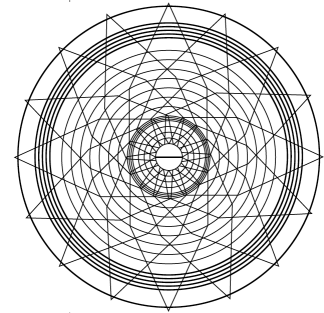
М. Коддингтон выделил несколько ключевых критериев для разграничения этих трех терминов: профессиональные навыки и знания, открытость (непрозрачность и прозрачность), теория познания (построение выборки и большой объем данных) и тип аудитории (пассивная и активная). Так, профессиональных навыков и активного участия в сборе информации требует вычислительная журналистика; дата-журналистика уже не требует традиционных методов сбора информации (интервью и другие профессиональные журналистские практики), она опирается на обработанные алгоритмом данные, которые затем используются для рассказа нарративной истории. «Прозрачность» характерна для дата-журналистики, когда читатель может проверить источники информации и результаты, тогда как в вычислительной и компьютерной журналистике используемые методы и анализ данных «гораздо сложнее». Вычислительная журналистика использует для анализа данных социологические методы, в том числе выборку, дата-журналистика и алгоритмическая журналистика анализируют весь объем данных. У вычислительной журналистики пассивный читатель, у двух других — активная онлайн-аудитория, имеющая возможность более энергичного и интерактивного участия и проверки журналистского материала.

Так, компьютерная журналистика предоставляет больше инструментария и вариативности не только для написания материалов, но и регулирования работы редакции, в том числе анализ отклика аудитории на полученный контент.

[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики

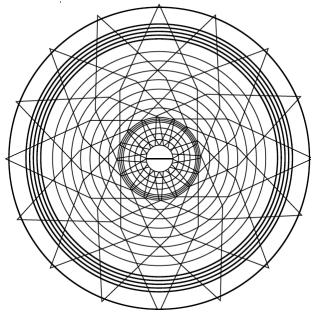


Автоматизированный контент в современной журналистике: профессиональные вызовы и этические аспекты

В 2016 году в «Яндекс для медиа» было запущено собственное информационное агентство. Специалисты отдела создали алгоритмы, способные не только писать шаблонные сводки о погоде или описывать дорожный трафик, но и анализировать интересы пользователей по результатам поисковых запросов, в том числе их предпочтения в кино и музыке, и формулировать это в краткие информационные сообщения, рассылая их затем своим подписчикам-СМИ. В США исследователи пошли (Caswell, Dörr 2018) дальше: с опорой на платформы генерации естественного языка NLG (например, WorldSmith) тестируется создание углубленных, нелинейных историй (event-by-event narrative). Исследователи выступают против исключительно «описательных» заметок алгоритмов, перейдя к созданию новостей, например, о продвижении законопроекта или ежедневных собраниях в местной администрации города. Несмотря на сложность кодирования такой информации, уже существует два варианта написания подобных многоуровневых текстов роботами. Реализация одного из них — вопрос времени.

Распространение возможности создания автоматизированного контента приводит к опасениям, что вскоре журналисты могут потерять рабочее место из-за профнепригодности. Исследователи расходятся во мнении, сможет ли алгоритм заменить профессионала. Так, теоретик Кристер Клерволл провел эксперимент (Clerwall 2014), результатами которого стали оценки новостей (написанных роботом и журналистом) участниками тестирования. Изначально респонденты не знали, кто является автором заметки. Тестирование показало, что автоматизированный контент не только на техническом уровне не уступает человеческому, но и по многим показателям опережает (скорость, информативность, объективность, точность). Более того, погрешность группы в определении авторства новости составила около 45% при том, что участники эксперимента — студенты факультета журналистики. Несмотря на то, что эксперимент критикуется (Caswell, Dörr 2018) как несколько упрощенный и несовершенный для оценки качества текстов алгоритмов (для более точного результата ученые предлагают оценивать новости алгоритмов не только по результатам перцепции читателя, но и проводить экспертный анализ сложности языка, информационной наполненности и наличия деталей), можно предположить, что алгоритмы, способные создавать новости на высоком техническом уровне, действительно требуют отдельного изучения как развивающийся феномен журналистики.

В связи с тем, что расходы на создание автоматизированного контента (важно отметить, что речь идет об использовании готового алгоритма, а не о создании таких технологий и их внедрении в редакцию) дешевле оплаты человеческого труда, роботизированная журналистика может быть выгодна для развития отдельных областей медиарынка и привлечения новой аудитории.



[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики

«Когда Associated Press делал раздел алгоритмов у себя, то большое количество редакторов, выполняющих ту же работу, что и роботы — [пишущих] шаблонные простые заметки — уволили. Associated Press решил сэкономить, но при этом зарабатывать. <...> Дорогое медиа — это контент; чем дешевле создавать, тем больше ты получаешь, тем больше рекламы, тем больше монетизация. Они [Associated Press] увеличили количество публикаций, где могут размещать рекламу без увеличения штата, без людей», — рассказала в нашем экспертном интервью глава «Яндекс. Медиа» Мария Петрова (на момент проведения интервью департамент компании имел это название).

С институционализацией алгоритмов как направления журналистики, профессия стала переживать трансформации и в этическом аспекте. Некоторые исследователи утверждают (Hollnbuchner, Dörr 2016), что ответственность за такие тексты (например, в случае ошибок) может лежать не только на журналисте и программисте, но и самом коде — алгоритмы способны создавать информационные сообщения и за счет машинного обучения принимать решения. Человек не может предсказывать с 100% точностью последующее принятие алгоритмом решений, поэтому часть ответственности, по мнению экспертов, может переходить к автоматизированной системе. С другой стороны, это звучит и парадоксально, ведь алгоритм на основе самообучаемости обрабатывает лишь ту информацию, которую в него закладывают члены редакции и разработчик.

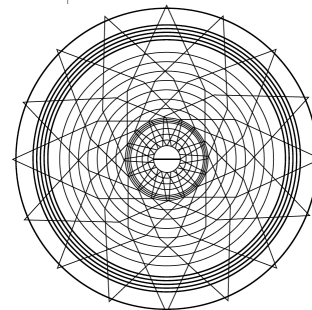
Бывший сотрудник Associated Press, специалист по этике и редакционным стандартам агентства Том Кент в 2015 году выпустил (Kent 2015) чек-лист для работы с алгоритмами. В первую очередь, Кент предложил изначально решить вопрос, кто несет ответственность в случае ошибки робота и ответить на основные вопросы: из каких источников будет брать алгоритм информацию и как проверять ее достоверность; есть ли у редакции права на пользование этой информацией; скрывать или нет от читателей автора новости.

В нашей работе для изучения трансформации этики в работе журналиста мы взяли за основу исследование Константина Дорра и Катарины Хольнбухнер, в котором утверждается (Hollnbuchner, Dörr 2016), что при развитии роботизированной журналистики происходит трансформация журналистской этики. Этика роботизированной журналистики — часть профессиональной деятельности журналиста на пересечении кибер-этики (выражение нравственных понятий (Herman 2011) — честности, ответственности и уважения к личности при программировании компьютеров, а также анализ поведения людей в Сети через указанные понятия) и медиа-этики (digital media ethic — изучение возможных конфликтов при технологическом и социальном развитии журналистики) (Hollnbuchner, Dörr 2016). Как правило, профессиональная этика журналиста находится в области медиа-этики, однако при развитии технологий алгоритмическая журналистика как часть профессиональной сдвигается к области кибер-этики, в том числе из-за перехода части ответственности за роботизированные материалы или программы с редакции на разработчиков.

[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики



Согласно выдвинутой теории Константина Дорра и Катарины Хольнбухнер, мы предлагаем под этикой роботизированной журналистики в нашем исследовании понимать нормы автоматизированного создания контента, его применения и распространения в профессиональной журналистике.

Роботизированная журналистика и особенности ее применения в российских изданиях

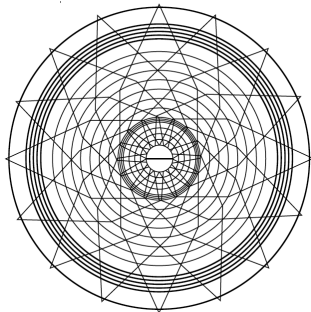
Эмпирическим материалом для нашего исследования стали экспертные интервью с шестью представителями медиаиндустрии в России. На основе описанного журналистами и разработчиками опыта, применяемого в редакциях (РБК, ТАСС, Mail.ru, «Яндекс. Медиа», Sports.ru и Socialdatahub) удалось выявить отношение индустрии к инновационному направлению, выделить его основные проблемы и обозначить перспективы развития алгоритмов в медиа, опираясь на поставленные задачи исследования.

Для отбора экспертов был выбран метод целевой выборки. Выборка составила шесть человек, критерием для отбора стал их опыт в использовании роботизации в работе, а также последующий осознанный отказ от нее. Было принято решение сделать остановку на шести экспертах, поскольку поле достигло информационного насыщения. Кроме того, мы стремились в построении выборки прийти к наибольшему разнообразию типов применения роботизированной журналистики в компаниях или редакциях для того, чтобы рассмотреть проблему с разных сторон.

На основе полученных данных был проанализирован вопрос о трансформации этики журналиста и сдвигении профессиональной ответственности журналиста за материалы. В том числе мы концептуализировали термин «роботизированной журналистики», так как понятие представляется достаточно общим.

Автоматизированная журналистика — это сложная система, включающая множество алгоритмов, направленных на решение различных задач. Существует большое количество рутинных процессов, связанных с работой журналиста (в том числе написание однотипных небольших заметок), которые можно подвергнуть автоматизации. Роботизированная журналистика, в первую очередь, — это набор инструментов для решения повседневных задач, экономии времени журналиста и ускорения его работы. «Это способ, когда робот делает грязную работу, а человек дорабатывает до конца» (глава Департамента разработки новой продукции Tribuna & Sports.ru Михаил Калашников).

Стоит выделить также ряд ключевых характеристик «роботизированной журналистики». Так, при анализе открытых источников данных она может определять тренды и некоторые аномалии. Иногда, отмечает Мария Петрова (глава департамента «Яндекс. Медиа»), алгоритм может неправильно оценить ситуацию, но команда разработчиков это видит и снова правит код, обучая машину.



[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики

Написание информационных сообщений — еще одна важная задача направления. Алгоритмы пишут тексты в разы быстрее человека. Однако программа может делать аналитические выводы только по установленным заранее правилам.

Важно добавить, что создание шаблонных новостей — это не основной подход алгоритмической журналистики. Система, когда робот подставлял в шаблон нужные данные, была освоена еще в 90-х годах. Сейчас система может генерировать сложносочинённые и сложноподчиненные предложения на основе выборки готового корпуса языка (в русском языке этим может стать национальный корпус русского языка), когда скрипт с помощью машинного обучения учится выстраивать связи между словами.

Для создания информационного сообщения робот должен уметь находить информационные поводы и определять события в кластеры по темам (как, например, это делают агрегаторы новостей или «Яндекс», объединяя новости в сюжеты). Как правило, сейчас информационным поводом становится либо некоторый «скачок» в данных, проявление аномалий (погода, трафик на дорогах, курс валют и другие финансовые показатели), либо событие (чаще всего спортивное, например, футбольный матч). «Роботизированная журналистика — явление, в котором новости, и в частности информационные поводы, создаются в большей степени роботами» (глава Отдела инноваций ТАСС Александр Маас).

До сих пор можно встретить статьи, в которых выражены опасения насчет замены человеческого труда роботами (Miller 2015), однако эксперты опровергают это, опираясь на ключевую задачу роботизированной журналистики — упрощение работы журналиста. Таким образом, учитывая все выделенные ранее особенности, под роботизированной журналистикой мы предлагаем понимать следующее: «Та часть технологически-ориентированной журналистики, сфокусированной на применении вычислительного мышления к практике сбора информации, поиска смысла и аномалий и представлении информации в виде сообщений или графических изображений, упрощении оформления контента (поиск фотографий, генерация заголовков) с целью решения повседневных задач, экономии времени журналиста и ускорения его работы».

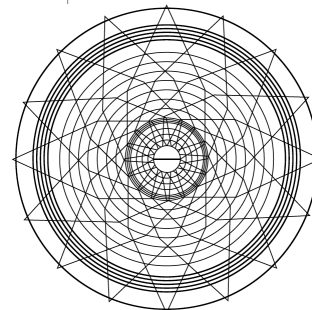
Для изучаемого направления мы выделили несколько основных задач в условиях его применения российскими редакциями. Первая задача — мониторинг. Существуют алгоритмы, которые через социальные медиа ищут необходимый контент и отправляют его в виде самой информации (однако в российских редакциях такой алгоритм, помимо «Яндекс. Медиа», еще только в разработке) и используется нотификация.

При этом на этапе мониторинга алгоритм должен проводить первичный факт-чекинг. Чтобы разобраться в этом, стоит рассмотреть следующий пример. Может быть следующий инфоповод: «Президент Владимир Путин прилетел в Крым и раздал нуждающимся по 20 000 рублей». Робот выделяет в сообщении несколько фактов: «Владимир Путин прилетел», «Крым», «раздал деньги

[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики



нуждающимся», «20 000 рублей». На следующем этапе проходит мониторинг всех доступных источников, и выясняется, что, например, в информации о деньгах «80% источников» пишут «20 000 рублей», «10%» — «5000 рублей», и алгоритм автоматически выбирает большинство.

По словам Александра Мааса (ТАСС), несмотря на недоверие информационных агентств к алгоритмам в проверке информации, разработка может показывать высокий результат: при анализе через социальные медиа аккаунта ньюсмейкера (определяя первоисточник, дату опубликования, место), алгоритм может оценить практически с 95% вероятностью оригинальность новости.

Однако среди редакций отношение к автоматическому факт-чекингу разнится: в одних изданиях его считают важным этапом в создании алгоритмом контента (ТАСС, SocialDataHub), но относятся с недоверием (ТАСС), в других он даже не предусмотрен («Яндекс», Sports.ru).

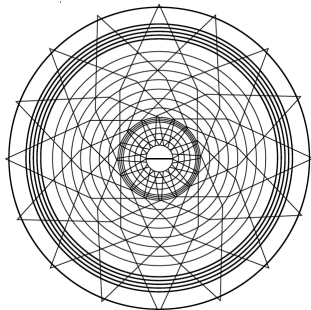
Например, факт-чекинга нет в Терминале «Яндекса» (информационная лента для подписчиков), так как информация из системы компании «одинаково интерпретируема» и при генерации сообщений «сложно соврать». «Мы же не публикуем слухи, данные о чьей-то зарплате, наша информация основана на измерении данных», — отмечает Мария Петрова («Яндекс»). Также факт-чекинг не был предусмотрен в генерации спортивных новостей в редакции Sports.ru по аналогичной с «Яндексом» причине, и, как отметил Михаил Калашников, за три года службы алгоритм не совершал фактических ошибок.

Вторая задача направления — упрощение ежедневной работы журналиста, в том числе создание текстовых сообщений. Например, в редакции Sports.ru алгоритм создавал новостные заметки о матчах, используя открытые сводки данных о составах команд и пр. Кроме того, алгоритм генерировал заголовки.

Согласно ответам экспертов «Яндекс. Медиа» и ТАСС, можно выделить еще две весьма неочевидные задачи алгоритмов. Первая — формирование информационного повода. Полностью на автоматическом и самостоятельном создании инфоповода работает Терминал «Яндекса», но, как показало исследование, больше никто в редакциях не доверяет машине полный контроль и мониторинг информационного поля. Тестирует анализ информационного поля, формирование инфоповода и в дальнейшем формирование сводок (как единый процесс) и ТАСС.

Стоит отметить распознавание изображений, видео и звука посредством новых технологий, которое также тестируется в ТАСС. Например, алгоритм уже способен определять лица, находить повторы, логотипы компаний, распознавать объекты на снимке — агентство использует собственную разработку для анализа информационного поля. Также ТАСС тестирует распознавание видео — алгоритм учится понимать, кто выступает на видео, транскрибировать запись.

Однако разработки в роботизированной журналистике находятся на начальном этапе, в алгоритмах нередко происходят сбои, что в российских СМИ



[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики

неизбежно вызывает скептицизм и даже отказ от алгоритмов после неудавшегося опыта.

Безусловно, одна из самых больших проблем — это морфология русского языка. Для английского языка есть множество корпусов текста для автоматических генераций, более того, такие корпуса текстов доступны и публичны. Однако мы выяснили, что в России такие разработки практически не проводятся, нет и публичных инструментов с готовыми корпусами. Более того, само проведение таких разработок является довольно дорогим (Мария Петрова, «Яндекс. Медиа»).

Сложность русского языка отмечали в первую очередь все опрошенные специалисты. Основные проблемы — это окончания, склонения и падежи. Более того, компьютер не в состоянии распознавать сарказм. Еще одна сложность — произвольное построение предложений и связь между предложениями. В рамках проблемы морфологии важно отметить контекст: так, в микроблоге Twitter робот не сможет связать разрозненные сообщения в один кластер по смыслу.

В Sports.ru алгоритмы для написания новостей и генерации заголовков работали почти три года с 2014. Редакция решила поэкспериментировать с автоматизацией и облегчить труд авторов, но позднее от инструмента отказались: добиться того, чтобы робот в 100% случаев не допускал ошибок в применении русского было невозможно (например, алгоритм мог воспринять название футбольного клуба «Крылья Советов» как имя человека и в соответствующем ключе его просклонять). Каждая такая ошибка «наносила ущерб имиджу издания и понижала доверие читателей», в итоге в редакции Sports.ru «затрачивалось много ресурсов, в том числе временных» на исправления, но был «неясен результат»: «Сил доделывать [алгоритм] до конца нет, но мы вернемся еще, постараемся доработать» (Михаил Калашников, Sports.ru).

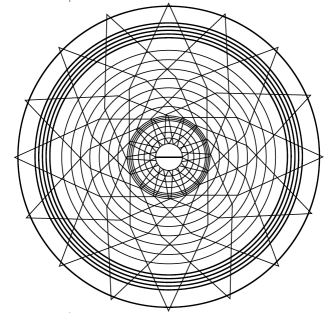
Более того, разработка таких технологий очень дорога: задача требует большого количества инженерных ресурсов, а значит — денежных. Такие технологии также крайне сложно монетизировать и найти в этом выгоду, особенно в России. По причине дороговизны разработки технологии и сложности прогнозирования доходов таких вложений редакция Mail.ru пока не использует автоматизацию в своей работе: «Мы вложимся, и? Что мы получим? Будут ли алгоритмы работать? Боюсь, что это не очевидно» (шеф-редактор медиапроектов Mail.ru Ольга Сидорова).

Еще одна проблема, которая вытекает из сложностей монетизации и вложения ресурсов редакции — стратегия развития СМИ. На данный момент алгоритм, который станет генерировать новости, мало кому из редакций необходим, СМИ ищут другие способы наполнения контентом и привлечения аудитории (например, как отмечал медиаменеджер РБК Иван Макаров, издания выкупают колонки известных авторов). Адаптация продукта к работе редакции

[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики



также неоднозначный этап в развитии направления: внедрение алгоритмов в работу журналистов требует серьезного обучения базового кадрового состава.

Более того, каждый алгоритм не универсален, и если редакция захотела писать не только о погоде, то ей придется заказывать новую разработку. «Яндекс» забывает, что технологию если делаешь, то ее нужно поддерживать. Это со стороны кажется, что вот мы сделали алгоритм и он у нас работает, потом выясняется, что алгоритм ломается <...> понадобятся доработки, время, деньги». (Михаил Калашников, Sports.ru).

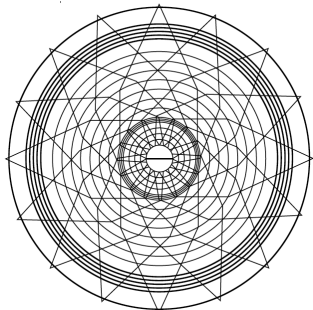
Проблемой может стать формирование информационных поводов и распределение тем по кластерам, что применяется в алгоритмах для проведения факт-чекинга. Например, 26 марта по всей России прошли антикоррупционные митинги, инициированные основателем Фонда борьбы с коррупцией (ФБК) политиком Алексеем Навальным. Произошли массовые задержания, однако в топ-5 «Яндекса» новость так и не смогла выйти (Брызгалова, Болецкая 2017), зато среди рейтинговых сюжетов оказалась ярмарка в центре столицы.

По данным наших экспертов, вес сюжетов может завышаться СМИ искусственно. Крупные игроки внутри одного холдинга имеют большое количество региональных изданий, которые ссылаются на лидеров холдинга, увеличивая тем самым их цитируемость. В самом «Яндексе» «умолчание» об акциях протеста объяснили тем, что митинги проходили в разных городах страны и алгоритм не связал их в один сюжет из-за разницы в географии.

Кроме того, к провалу факт-чекинга и объективности, что становится значительным изъяном, может привести заранее определенная выборка источников, поскольку разработчик может изначально закрыть доступ алгоритму к определенным ресурсам.

В российских редакциях относятся к первичному факт-чекингу и мониторингу информации алгоритмами с предосторожностью. Более того, информация в большинстве СМИ подтверждается людьми в источниках, что алгоритм, конечно, сделать не может: «Полностью роботу формирование инфоповода и повестки на данный момент дать не готовы, это будет требовать постоянной постпроверки. Факт ошибки очень велик» (Александр Маас, ТАСС).

Отношение внутри редакций к ошибкам, которые допускают роботы, неоднозначное. Для одних изданий даже небольшая фактическая неточность — это серьезная ошибка, которая наносит вред изданию (Sports.ru, ТАСС), в других редакциях к ошибкам роботов («Яндекс») относятся достаточно спокойно. Вопрос об этической стороне ответственности изданий за совершенную алгоритмом ошибку оказался достаточно субъективным и спорным (важно отметить, что вопрос рассматривался в приоритете не с точки зрения законодательства, а с точки зрения этики и договоренностей внутри каждого СМИ, с целью выявить отношение индустрии к роботизированной журналистике в целом).



[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналисткой этики

Для респондентов на данном этапе был задан следующий вопрос: «Кто виноват, если будет допущена ошибка алгоритмом, я использую эти данные в своей заметке и в итоге выпущу ее с ложной информацией? Разработчик или журналист?» Меньшинство респондентов выступило за однозначную ответственность разработчика за ошибку. Данная позиция объясняется тем, что разработчик предоставляет журналистам услугу, и такой «товар» должен проходить сертификацию — необходимое количество тестов и проверок, поэтому программист должен давать гарантии. «Большую ответственность несет программист, но тут, как и в самолетостроении, других областях, должен быть принцип сертификации» (Иван Макаров, РБК).

Большинство респондентов выступило за то, что ответственность остается на плечах журналиста, учитывая уязвимость алгоритма. Во-первых, алгоритм, который собирает информацию и выдает данные, — внутренний источник редакции, а журналист, предполагается, должен каждый источник проверить. В том числе и Терминал «Яндекса» как новостную ленту можно отнести к такому источнику. «Это, если честно, проблема журналиста. Если он отвечает за публикацию, он должен проводить факт-чекинг тех материалов, которые он использует» (Мария Петрова, «Яндекс. Медиа»).

Позиция ответственности журналиста подтверждается и тем, что журналист якобы не должен доверять сообщениям алгоритма, и, даже если робот со временем будет способен «хорошо и безошибочно» генерировать текст на русском языке, обязательно в процессе должен быть ответственный человек, который будет самостоятельно отслеживать повестку дня.

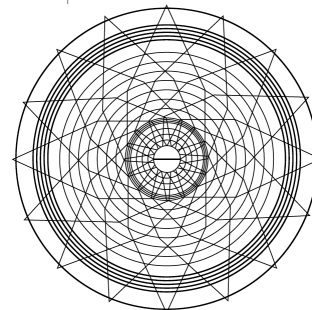
Существует еще один немаловажный аспект. Искусственный интеллект, связанный с нейронными сетями, находится в постоянном обучении. База, которую «сдавал» изготовитель в редакцию, была одним алгоритмом, через несколько запусков и процессов обучения она становится уже «другим алгоритмом» с новыми нейронными связями, поэтому за трансформации алгоритма и возможные ошибки программист может не нести ответственности. Более того, разработка алгоритма — это работа целой команды не только разработчиков, но и журналистов, на текстах которых машина проходит обучение. «Это ведь не конечный продукт, рано или поздно они [алгоритмы] начинают принимать решения самостоятельно» (Александр Маас, ТАСС).

Наконец, в каждой редакции есть «инструкция», по которой журналист должен проверять информацию, даже если его заметка опирается на внутренние источники. Читатели ожидают, что каждая информация прошла определенный «фильтр» (Flanagin, Metzger 2000) и формирует свое доверие к СМИ. «Если мы такие правила редакции будем перекладывать на нейронную сеть, то «сдвигения» этики с плеч журналиста происходить не будет, мы просто используем другие инструменты, чтобы выполнять работу» (Ольга Сидорова, Mail.ru).

[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики



Таким образом, практически все эксперты, опираясь на опыт своей редакции, выразили мнение, что ответственность остается за журналистом. Несмотря на то, что алгоритм способен брать на себя солидную часть работы, система еще слишком уязвима. По причине многих сложностей и проблем, которые возникают при использовании алгоритма, в российских редакциях относятся с недоверием и скептицизмом к автоматизации. Однако важно отметить, что разработки в некоторых редакциях ведутся, и в течение нескольких лет журналисты надеются получить более совершенный алгоритм.

Специалисты отмечают, что одна из главных задач журналиста — проверка фактов и информации, которую они обязаны совершать даже за внутренним источником информации, в виде которого выступает роботизированная журналистика. Алгоритм — это инструмент, и в российских редакциях он выступает лишь «младшим помощником», ассистентом журналиста, в связи с чем трансформации в этике журналиста не происходит, в отличие от опыта коллег на Западе, где алгоритм становится полноценным участником медиапроизводства.

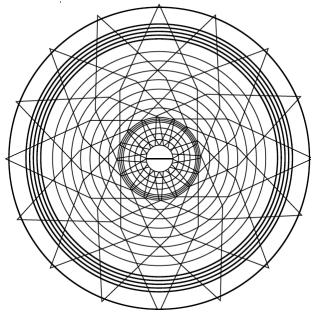
Важно отметить, что редакции приходят к тестированию автоматизированной журналистики не только потому, что это важно для упрощения работы редакции, но и из-за интереса СМИ к новому направлению. Как таковой особой потребности для использования алгоритмов, в том числе для создания сообщений, не было, практически все проекты по внедрению строились на чистом энтузиазме и любопытстве, что в итоге привело к многочисленным сложностям. Возможно, через несколько лет, когда разработчики редакций и журналисты, учитывая свой опыт, подойдут к направлению более осознанно, начнется новый этап развития алгоритмической журналистики в России.

Выводы и рекомендации

Таким образом, несмотря на потенциальные выгоды медиаиндустрии, когда алгоритм увеличивает скорость работы журналиста, позволяет расширить количество тем и информационных поводов (которые ранее не было возможности затронуть в связи с недостатком рабочей силы и времени), выявлять информационные поводы, в России к подобной автоматизации относятся с недоверием и скептицизмом.

Этому способствует ряд особенностей роботизированной журналистики: алгоритм кластеризации совершает ошибки и неправильно выделяет инфоповоды, не учитывая контекст; уязвимость алгоритма перед человеческим фактором при программировании алгоритмов и цензурой; сложности в морфологии. Алгоритм требует постоянной «тренировки», а также институционализация алгоритмов в качестве создателей контента поднимает этические вопросы: проблему авторства и ответственности за ложную, неправильно интерпретированную и искаженную информацию в тексте.

Мы сделали вывод, что в российских редакциях алгоритмам доверяют сбор и анализ данных, а не мониторинг, факт-чекинг и написание заметок (более



[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики

сложные и ответственные процессы). Таким образом, ответственность остается на плечах журналиста не только из-за уязвимости в российских реалиях алгоритма, который разработчики пока не могут довести до релевантного уровня выполнения работы, но и из-за отношения к алгоритму не как к самостоятельному производителю контента, а как к инструменту, с помощью которого человек может упрощать свою работу, осуществляя сбор данных быстрее. Использование таких новых инструментов не предполагает снятие ответственности и «сдвига» такой ответственности с журналиста на разработчика.

Мы предполагаем, что в процессе внедрения автоматизации через несколько лет этика роботизированной журналистики все же начнет движение к периферии на «этике СМИ и кибер-этики», как у западных коллег, когда ответственность за ошибки внутреннего ресурса смогут брать на себя и разработчики в процессе внутреннего производства.

БИБЛИОГРАФИЯ

Caswell D., Dörr K. (2018). Automated Journalism 2.0: Event-driven narratives: From simple descriptions to real stories. *Journalism Practice*, 12(4), 477–496. <https://doi.org/10.1080/17512786.2017.1320773>

Clerwall, C. (2014). Enter the robot journalism: Users' perceptions of automated content. *Journalism Practice*, 8(5), 519–531. <https://doi.org/10.1080/17512786.2014.883116>

Coddington, M. (2014). Clarifying journalism's quantitative turn. A typology for evaluating data journalism, computational journalism, and computer-assisted reporting. *Digital Journalism*, 3(3), 331–348. <https://doi.org/10.1080/21670811.2014.976400>

de Sola Pool, I. (1987). Electronics Takes Command. In R. Finnegan, G. Salaman, K. Thompson (Eds.). *Information Technology: Social Issues. A Reader*. London: The Open University.

Diakopoulos, N. (2014). Algorithmic accountability: Journalistic investigation of computational power structures. *Digital Journalism*, 3(3), 398–415. <https://doi.org/10.1080/21670811.2014.976411>

Federman, M. (2004). What is the Meaning of the Medium is the Message? [URL:http://individual.utoronto.ca/markfederman/article_mediumisthemessage.htm](http://individual.utoronto.ca/markfederman/article_mediumisthemessage.htm)

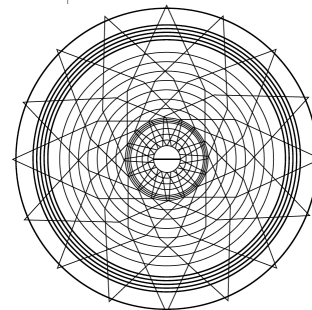
Flanagin, A., Metzger, M. (2000). Perceptions of Internet Information Credibility. *Journalism & Mass Communication Quarterly*, 77(3), 515–540.

Gordon, R. (2013). Want to build a data journalism team? You'll need these three people. [URL:http://knightlab.northwestern.edu/2013/06/28/want-to-build-a-data-journalism-team-youll-need-these-three-people/](http://knightlab.northwestern.edu/2013/06/28/want-to-build-a-data-journalism-team-youll-need-these-three-people/)

[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики



Gray, J., Chambers, L., Bounegru, L. (2012). *The Data Journalism Handbook. How Journalists Can Use Data to Improve the News*. Newton, MA: O'Reilly Media Inc.

Herman T. (2011). *Ethics and technology: controversies, questions, and strategies for ethical computing*. Hoboken, N.J.: Wiley.

Hollnbuchner, K., Dörr, K. (2017). Ethical challenges of algorithmic journalism. *Digital Journalism*, 5(4), 404–419. <https://doi.org/10.1080/21670811.2016.1167612>

Kent, M. (2015). An ethical checklist for robot journalism. URL:<https://medium.com/@tjrkent/an-ethical-checklist-for-robot-journalism-1f41dcbd7be2>

Marconi, F., Siegman, A., Journalist, M. (2017). *The future of augmented journalism: A guide for newsrooms in the age of smart machines*. New York: The Associated Press.

McLuhan, M. (2002). *Understanding Media: The Extensions of Man*. Cambridge, MA, USA: The MIT Press.

McLuhan H., McLuhan E. (1992). *Laws of Media: The New Science*. Toronto: University of Toronto Press.

Meyer, F. (2002). *Precision Journalism: A Reporter's Introduction To Social Science Methods*. Lanham MD: Rowman and Littlefield.

Miller, R. (2015). AP's 'robot journalism' are writing their own stories now. *The Verge*. URL:<http://www.theverge.com/2015/1/29/7939067/ap-journalism-automation-robots-financial-reporting>

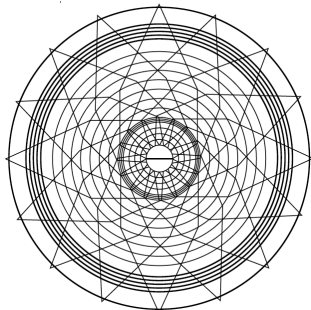
Petre, C. (2013). *A Quantitative Turn in Journalism?* Tow Center for Digital Journalism. URL:<http://towcenter.org/a-quantitative-turn-in-journalism>

Брызгалова Е., Болецкая К. (2017) Сюжет о задержании Навального исчез из главных новостей «Яндекса». *Ведомости*. URL:<https://www.vedomosti.ru/politics/articles/2017/03/26/682745-syuzhet-ob-areste-navalnogo>

Иванов А. (2015) Роботизированная журналистика и первые алгоритмы на службе редакций международных СМИ. *ЗНАК: проблемное поле медиаобразования*, 2(16): 32–40.

Иванов А. (2017) Транспарентность роботизированной журналистики: как новые технологии угрожают принципам профессии. *Вестник Челябинского государственного университета. Филологические науки*, 8(404): 28–33.

Кирия И. (2010) Теоретический... или что стоит за трудными словами. Что такое мультимедиа? Мультимедиа и новые принципы новостей. Организация труда в мультимедийной редакции конвергентного СМИ. А. Г. Качкаева (ред.) *Журналистика и конвергенция: почему и как традиционные СМИ превращаются в мультимедийные*, М.: Аспект Пресс: 12–62.



[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики

ROBOTIZATION OF NEWS JOURNALISM IN RUSSIA: NEW PRINCIPLES OF EDITORIAL WORK AND TRANSFORMATION OF JOURNALISTIC ETHICS

Isaev E.M.

Lecturer on Faculty of Communications, Media, and Design,
Higher School of Economics (HSE) (Moscow, Russia)

emisae@hse.ru

Kokoreva M.V.

Correspondent RBC (Moscow, Russia)

man-kokoreva@yandex.ru

Abstract:

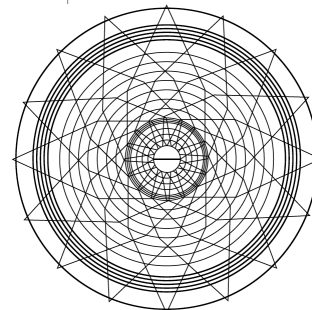
The purpose of this research is to point out the changes in journalistic ethics in Russia (in particular, the problem of fact-checking and responsibility for materials) caused by its robotization. We describe this through analyzing the media's editorial policy, their practice in creating a media content by algorithms, as well as a number of expert interviews with respondents who have encountered or face the processes of media robotization within their sphere of work. Potential benefits of the media industry from the introduction of algorithms in the work of the editorial staff and future revision of journalistic ethics do not inspire trust in Russian media. The responsibility remains on the journalists' shoulders not only because of the vulnerability of the algorithm in Russian realities, which developers can not yet bring to the relevant level of performance, and distrust of journalists to the direction, but also the attitude to the algorithm as a tool with which people can only simplify their work.

Keywords: robotic journalism, data journalism, journalistic ethics, convergent journalism, precision journalism, algorithmic journalism.

[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики



REFERENCES

Bryzgalova E., Boletskaya K. (2017) Syuzhet o zaderzhanii Naval'nogo ischez iz glavnykh novostey «Yandex». Vedomosti. URL:<https://www.vedomosti.ru/politics/articles/2017/03/26/682745-syuzhet-ob-areste-navalnogo>

Caswell D., Dörr K. (2018). Automated Journalism 2.0: Event-driven narratives: From simple descriptions to real stories. *Journalism Practice*, 12(4), 477–496. <https://doi.org/10.1080/17512786.2017.1320773>

Clerwall, C. (2014). Enter the robot journalism: Users' perceptions of automated content. *Journalism Practice*, 8(5), 519–531. <https://doi.org/10.1080/17512786.2014.883116>

Coddington, M. (2014). Clarifying journalism's quantitative turn. A typology for evaluating data journalism, computational journalism, and computer-assisted reporting. *Digital Journalism*, 3(3), 331–348. <https://doi.org/10.1080/21670811.2014.976400>

de Sola Pool, I. (1987). Electronics Takes Command. In R. Finnegan, G. Salaman, K. Thompson (Eds.). *Information Technology: Social Issues. A Reader*. London: The Open University.

Diakopoulos, N. (2014). Algorithmic accountability: Journalistic investigation of computational power structures. *Digital Journalism*, 3(3), 398–415. <https://doi.org/10.1080/21670811.2014.976411>

Federman, M. (2004). What is the Meaning of the Medium is the Message? URL:http://individual.utoronto.ca/markfederman/article_mediumisthemessage.htm

Flanagin, A., Metzger, M. (2000). Perceptions of Internet Information Credibility. *Journalism & Mass Communication Quarterly*, 77(3), 515–540.

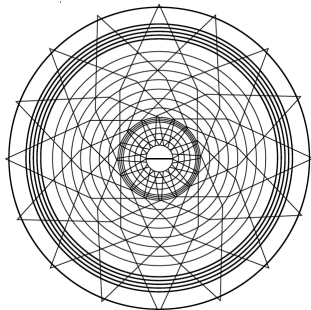
Gordon, R. (2013). Want to build a data journalism team? You'll need these three people. URL:<http://knightlab.northwestern.edu/2013/06/28/want-to-build-a-data-journalism-team-youll-need-these-three-people/>

Gray, J., Chambers, L., Bounegru, L. (2012). *The Data Journalism Handbook. How Journalists Can Use Data to Improve the News*. Newton, MA: O'Reilly Media Inc.

Herman T. (2011). *Ethics and technology: controversies, questions, and strategies for ethical computing*. Hoboken, N.J.: Wiley.

Hollnbuchner, K., Dörr, K. (2017). Ethical challenges of algorithmic journalism. *Digital Journalism*, 5(4), 404–419. <https://doi.org/10.1080/21670811.2016.1167612>

Ivanov A. (2015) Robotizirovannaya zhurnalistika i pervye algoritmy na sluzhbe redaktsiy mezhdunarodnykh SMI. *ZNAK: problemnoe pole mediaobrazovaniya*, 2(16): 32–40.



[Научные статьи]

Исаев Е.М., Кокорева М.В.

Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики

Ivanov A. (2017) Transparentnost' robotizirovannoy zhurnalistiki: kak novye tekhnologii ugrozhayut printsipam professii. Vestnik Chelyabinskogo gosudarstvennogo universiteta. Filologicheskie nauki, 8(404): 28–33.

Kent, M. (2015). An ethical checklist for robot journalism. [URL:https://medium.com/@tjrkent/an-ethical-checklist-for-robot-journalism-1f41dcbd7be2](https://medium.com/@tjrkent/an-ethical-checklist-for-robot-journalism-1f41dcbd7be2)

Kiriya I. (2010) Teoreticheskiy... ili chto stoit za trudnymi slovami. Chto takoe mul'timedia? Mul'timedia i novye printsipy novostey. Organizatsiya truda v mul'timediynoy redaktsii konvergentnogo SMI. A. G. Kachkaeva (Ed.) Zhurnalistika i konvergentsiya: pochemu i kak traditsionnye SMI prevrashchayutsya v mul'timediynye, M.: Aspekt Press: 12–62.

Marconi, F., Siegman, A., Journalist, M. (2017). The future of augmented journalism: A guide for newsrooms in the age of smart machines. New York: The Associated Press.

McLuhan, M. (2002). Understanding Media: The Extensions of Man. Cambridge, MA, USA: The MIT Press.

McLuhan H., McLuhan E. (1992). Laws of Media: The New Science. Toronto: University of Toronto Press.

Meyer, F. (2002). Precision Journalism: A Reporter's Introduction To Social Science Methods. Lanham MD: Rowman and Littlefield.

Miller, R. (2015). AP's 'robot journalism' are writing their own stories now. The Verge. [URL:http://www.theverge.com/2015/1/29/7939067/ap-journalism-automation-robots-financial-reporting](http://www.theverge.com/2015/1/29/7939067/ap-journalism-automation-robots-financial-reporting)

Petre, C. (2013). A Quantitative Turn in Journalism? Tow Center for Digital Journalism. [URL:http://towcenter.org/a-quantitative-turn-in-journalism](http://towcenter.org/a-quantitative-turn-in-journalism)