

УДК 70

Научная статья

## **BIG DATA КАК ФЕНОМЕН ИННОВАЦИОННОЙ ЖУРНАЛИСТИКИ**

**Шаймерген Г., Абдиназар С.**

*Международный университет информационных  
технологий,  
г. Алматы, Казахстан*

### **АННОТАЦИЯ**

Исследование посвящено изучению истории появления и развития журналистики данных, анализу актуальности работы с данными для журналистики на современном этапе, а также феномену явления big data в журналистике данных и журналистике в целом. Изучается история развития журналистики данных. Проводится анализ основных характеристик больших данных.

Представлен обзор опыта анализа big data в журналистике данных в зарубежных СМИ и казахстанских журналистов, успешно использующих большие данные в своих исследованиях. Приведены результаты влияния цифровых технологий на различные типы журналистских ролей и как они изменились, а также затронуты тренды развития отечественных изданий по журналистике данных.

**КЛЮЧЕВЫЕ СЛОВА:** журналист, журналистика данных, big data, большие данные, СМИ, анализ данных, инфографика, сбор данных.

**DOI: 10.24334/KBSU.2021.1.3.009**

Поступила в редакцию: 10.12.2021

В окончательном варианте: 15.12.2021

**Для цитирования:** Шаймерген, Г. Big Data как феномен инновационной журналистики / Г. Шаймерген, С. Абдиназар. – Текст : непосредственный // Вестник Кабардино-Балкарского гос-

ударственного университета: Журналистика. Образование. Словесность. – 2021. – Т. 1, № 3. – С. 99-121. – DOI 10.24334/KBSU.2021.1.3.009

## **ВВЕДЕНИЕ**

В наше время быстрого развития цифровых и информационных технологий трансформации подвергаются все сферы жизнедеятельности. Не стала исключением и журналистика. Существует несколько видов журналистики, каждая из которых может быть представлена в разных формах.

В настоящее время открывается всё больше возможностей для передачи, хранения и анализа больших массивов данных, которые ежегодно увеличиваются. Эта тенденция привела к возникновению в начале 2000-х гг. термина “большие объемы данных” (big data). Появление этого термина обозначает начало формирования кардинально новых инструментов, предназначенных для работы с большими объемами данных и те возможности, которые они привнесут в журналистику. Несмотря на то, что понятие больших объемов данных не имеет конкретно очерченных границ, следует разделять работу с данными в общем смысле слова и работу именно с крупными массивами данных в силу различия в подходе и использовании для работы с ними новых инструментов и методов. Появление новых журналистских проектов и материалов, основанных на работе с крупными массивами данных, все больший доступ к массивам открытой информации, а также распространение в Интернете в открытом виде крупных объемов информации (в том числе, считавшейся до недавнего времени секретной) представляет совершенно новые задачи и дает уникальные возможности для развития современной журналистики. Эти изменения в журналистском процессе объясняют актуальность выбранных задач исследования.

Большие данные в свою очередь стали неотъемлемой частью современного направления журналистики – журналистики данных – специализированной деятельности, заключающейся в поиске и распространении новостей, где традиционные журналистские методы работы сочетаются с методами анализа данных, программирования и визуализации.

Сейчас, когда у каждого пользователя Интернета появляется все больший доступ к данным, к возможностям для поиска, сбора и обработки огромного объема информации, вопросы работы с данными становятся как никогда актуальны. При этом, получить необходимые данные стало возможно, как из открытых источников, так и в результате журналистских запросов или поиска данных через не открытые источники. Следует понимать, что направление “журналистики данных” включает в себя весь спектр работ над журналистским материалом, в основе которого лежит анализ большого объема данных. То есть, это и поиск первичных данных, и очистка массива данных, и его последующая аналитика, трактовка и визуальное представление для читательской аудитории. В результате по своей сути журналистика данных служит связующим звеном для двух различных направлений работы – технической (которая предполагает проведение работ по сбору и аналитике большого объема данных) и журналистской (которая предполагает исследование темы, а также трактовку и подготовку концепции для наилучшего представления выводов исследования). Сейчас, когда журналисты все чаще сталкиваются с наборами “сырых” данных, для профессионального журналиста необходимым становится умение работать с ними, ориентироваться в данных, анализировать и получать новые сюжеты. Понимание данных как в ключе возможности углубиться в уже существующий сюжет, так и в ключе основы для создания сюжета “с

нуля” по итогам изучения имеющегося массива информации служит новым этапом в развитии современной журналистики.

Целью статьи является исследование журналистики данных как объекта и практики использования big data в журналистике данных, а именно выявление тенденций применения big data в профессиональной журналистской деятельности.

Пионером направления журналистики данных, считается исследование The Guardian 1821 г., в котором были отражены сводные таблицы с указанием цен на обучение в высших учебных заведениях Британии. Материал вызвал положительные отзывы, так как впервые все данные были открыто представлены для читателей.

Другой полноценный материал, представляющим журналистику данных, принадлежит легендарной Флоренс Найтингейл. В 1858 г. автор на основе исследования, написала учебное пособие, посвященное анализу состояния и ежегодных потерь британской армии. В 54-страничном материале автор приводит огромное количество данных, переведенных в табличный вид и диаграммы. Книга представляет собой нарисованную от руки инфографику, посвященную анализу состояния ежегодных потерь в штате британской армии в 1858 г., а также конкретным причинам смертей солдат в цифро-процентном соотношении [6]. Новаторство работы заключалась не в использовании графиков и диаграмм (их в журналистских материалах использовали и до работы Найтингейл), а в подходе к “сырым” данным, с помощью которых автор оформила по своей сути полностью прозрачный (в значении используемых источников и данных) материал, который при желании может самостоятельно проанализировать каждый читатель, используя опубликованные в статье данные.

В современном своём виде журналистика данных появилась в 1952 г. в тот момент, когда в работе журналиста впервые был использован компьютер. Телеканалом CBS были арендованы вычислительные мощности компьютера для расчета прогнозов на результаты президентских выборов и оценки результатов каждого из кандидатов. Подобная практика использования компьютера для журналистики стала носить понятие “компьютерной журналистики” (Computer-assisted reporting (CAR)). Первым СМИ, который занялся журналистикой данных в наше время по праву считается The Guardian (2009 г.) запустивший раздел Datablog, в котором все материалы создаются с помощью обработки и представления данных [7].

В связи с тем, что в настоящее время в Интернете есть возможность найти всё больше данных в открытом доступе в отличие от ситуации, которая существовала ранее, когда журналисты в большинстве своём пытались найти новые данные и информацию, ускорило развитие журналистики данных, повысился спрос и ценность на умение обрабатывать данные и находить в них журналистские истории.

Следующей ступенью в развитии журналистики данных оказался тренд “big data” (рус. большие данные). Большие данные являются вторым по значимости трендом развития современного общества (после виртуальности). Понятие big data, обозначает структурированные и неструктурированные (данные представлены в форме текста с датами, цифрами, фактами, расположенными в нём в произвольной форме) данные огромных объёмов, которые возможно эффективно обрабатывать программными инструментами.

На техническом языке определение “big data” можно сформулировать как специальный алгоритм для обработки постоянно

увеличивающегося объема данных. Сегодня информация собирается огромными объемами из разных источников: интернет, контакт-центры, мобильные устройства и т. д. Чаще всего такие данные не имеют четкой структуры и упорядоченности, поэтому человек не может использовать их для какой-либо деятельности. Для автоматизации анализа применяют технологии “big data”.

Первая платформа, предназначенная для взаимодействия с большими наборами данных Hadoop начала работать в 2005 г. Сегодня она представляет собой большой стек технологий для обработки информации. Чуть позже популярность начала набирать NoSQL – совокупность методов для создания систем управления большими данными.

Объем генерируемой информации стал увеличиваться с появлением крупных интернет-сервисов. Пользователи загружают фотографии, просматривают контент, ставят “лайки” и т.п. Вся эта информация собирается в больших объемах для дальнейшего анализа, после которого можно вносить улучшения в работу сервисов. Например, социальные сети используют большие данные для показа пользователям релевантной рекламы (то есть той, которая соответствует их потребностям и интересам) в таргетинге, то позволяет соцсетям продавать бизнесу возможность проведения точных рекламных кампаний.

Чтобы массив информации обозначить приставкой “big” он должен обладать следующими тремя свойствами:

- объем (volume) – данные измеряются по физической величине и занимаемому пространству на цифровом носителе;
- скорость, обновление (velocity) – информация регулярно обновляется и для обработки в реальном времени необходимы интеллектуальные технологии больших данных; разнообразие;

– разнообразие (variety) – информация в массивах может иметь неоднородные форматы, быть структурированной частично, полностью и скапливаться бессистемно. например, социальные сети используют большие данные в виде текстов, видео, аудио, финансовых транзакций, картинок и прочего.

Позднее список свойств приобрел два дополнительных свойства: ценность и достоверность. Ценность определяется ответом на вопрос – принесет ли полученная информация пользу деятельности, а достоверность – можно ли доверять используемым данным (насколько они правдивы), ведь неточная информация может навредить организации и ее деятельности.

Работа с большими данными делится на 3 этапа: интеграция, управление и анализ.

1 этап. Интеграция. Компания интегрирует в свою работу технологии и системы, позволяющие собирать большие объемы информации из разных источников. Внедряются механизмы обработки и форматирования данных для упрощения работы аналитиков с “big data”.

2 этап. Управление. Полученные данные нужно где-то хранить, этот вопрос решается до начала работы с ними. Решение принимается на основе множества критериев, главными из которых считаются предпочтения по формату и технологии обработки. Как правило, для хранения организации используют локальные хранилища, публичные или частные облачные сервисы.

3 этап. Анализ. Большие данные начинают приносить пользу после анализа. Это заключительный этап взаимодействия с ними. Для этого применяют машинное обучение, ассоциацию правил обучения, генетические алгоритмы и другие технологии. После анализа данных остается только самое ценное для деятельности.

Большие данные становятся доступными благодаря растущему распространению мобильных устройств, датчиков, “умных” машин, цифровых данных отслеживания, цифровых хранилищ и архивов, а также других фрагментов социальной и природной деятельности, представленных по кликам, твитам, лайкам, GPS-координатам, временным меткам и т. д. Огромный объем цифровых данных и его безграничный рост поразителен, и точно так же значительным является возрастающая легкость, с которой стандартное компьютерное программное обеспечение может управлять и манипулировать наборами данных, которые когда-то требовали применение суперкомпьютеров, тем самым увеличивая эпизоды исследования цифровых данных. В определенное время “момент больших данных” осуществил технологический переход к потоку данных. Это социотехническое явление с культурными, экономическими и политическими истоками и последствиями; это действительно мифология, а также наука или бизнес.

Сущность big data в общем виде возможно описать как цифровые машиночитаемые постоянно обновляющиеся в онлайн-режиме объёмы данных, анализ которых позволяет определить существенные закономерности и интеллектуальные явления, значимые для оптимизации сферы их применения [3]. Главной качественной характеристикой big data являются результаты их обработки и анализа специалистами в сфере науки о данных (англ. data science), представление и использования результатов в различных сферах жизнедеятельности. То есть большие данные заведомо техносоциальный феномен [2].

## **МЕТОДЫ ИССЛЕДОВАНИЯ**

В этом разделе мы сосредоточимся на четырех концепциях, которые подчеркивают последствия больших данных для журналистики: эпистемология, экспертиза, экономика и этика. Эти четыре концепции позволят провести наблюдение за изменениями в различных отраслях новостей, СМИ и информационных технологий, а также затронет важную динамику больших данных и их значение для журналистики, как для академической теории, так и для профессиональной практики.

Феномен больших данных вызывает множество актуальных вопросов к новостям СМИ, некоторые наиболее важные из них связаны с легитимацией новых заявлений о знаниях и истине; переговорами о профессиональном статусе, полномочиях, наборов навыков по мере разработки и внедрения новых специализаций; потенциале и проблемах, связанных с новой рентабельностью, ресурсами, инновациями, созданием ценностей и возможности получения дохода. Эти четыре концепции служат динамическим набором линз, через которые появилась возможность просматривать современные разработки и предвидеть возможности будущих исследований.

## **РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ**

Подобно научному, корпоративному и государственному сектору в целом, медиаиндустрия должна решить вопрос: что делать с потоком увеличивающихся данных. Это связано с тем, что многие традиционные новостные организации пытаются найти свой путь в условиях подрыва профессионального авторитета, бизнес-моделей и традиционной логики производства и распространения информации.

Журналисты давно знакомы с данными и базами данных как с объектом новостной работы и журналистских свидетельств, что очевидно из многолетнего использования компьютерных репортажей (CAR) и даже более старых форм визуализации информации. Но база данных, которую Манович (1999) [5] смело предлагает для цифровой эры, повествование в романах и кино, которое относилось к современной эпохе, приобрело особенно заметное значение в современной журналистике. Значительный поворот к оцифровке информации в последнее время был связан с большой ролью в журналистике методов компьютерных наук и наук о данных – от программирования и алгоритмов до машинного обучения и вероятностных моделей, а также стиль разработки программного обеспечения с открытым исходным кодом и его акцент на обеспечении прозрачности и интерактивности наборов. Традиционное программное обеспечение для управления базами данных, или просто “большое” в своем потенциально преобразующем значении, приобрело особое значение для новостей. Главный вопрос использования журналистами больших данных заключался не в том, могут ли данные, компьютеры и алгоритмы использоваться журналистами в общественных интересах, а скорее “как?”, “когда?”, “где?”, “почему?” и “кем?” использоваться. Помимо журналистики, основанной на данных, в редакционной части новостных СМИ, манипуляции с данными представляют не меньший интерес и для бизнес-стратегий и приложений для прогнозирования рынка, выявление снижения доходов и дифференциации аудитории. По мере того, как большие данные приобретают все большее значение, необходимы технологии для определения и адаптации подходящих технологических систем и решений от внешних поставщиков или разработки и реинжиниринга этих систем и решений. Эти технологии, в свою очередь, могут

улучшить важную коммуникативную функцию в переговорах по технологическим системам и решениям в редакционных и деловых подразделениях организации, урегулировать давно существующие разногласия между отделами новостей и бизнеса, маркетинга. Таким образом, большие данные как социальный, культурный и технологический феномен служат концептуальным “микроскопом” в понимании журналистики как профессиональной сферы, так и коммерческого предприятия – стремясь действовать осмысленно и извлекать пользу из растущих массивов цифровых данных в общественной жизни.

Последствия больших данных для журналистики требуют концептуальной эвристики, которая может служить ориентиром для будущих исследований. Журналисты, как и астрономы, биологи, другие эксперты пытаются разобраться в больших данных, существует необходимость разрабатывать инструменты для понимания, что они значат для практики и восприятия новостей в журналистских исследованиях. Журналисты в новостях СМИ могут понять феномен больших данных по-разному, следовательно, и подход у всех разный: кто-то сопротивляется, кто-то приспосабливается, искажая или формируя данные. Они могут действовать проактивно или реактивно, сосредотачиваясь на отсутствии, одной или нескольких стратегических целях, которые находятся на стыке редакционных, деловых и технологических интересов и практик, из чего следует необходимость концептуальной отправной точки.

В материалах по журналистике до сих пор концептуальные разработки и исследовательские программы в основном сфокусированы на “вычислительной журналистике” или “вычислительных исследованиях в журналистике”. Но в этом анализе “вычис-

лений” отсутствует более целенаправленное рассмотрение “больших данных” как концепции и феномена. Тем не менее имеются примеры журналистских исследований в целевом маркетинге больших данных и производства средств массовой информации, рассмотрение пользователей смартфонов как основных клиентов больших данных, использование значений больших данных для методов исследования коммуникации. Журналисты и новостные организации стремятся осмыслить и извлечь пользу из больших данных во время исследования алгоритмов, вычислений и количественной оценки. В целом развитие больших данных потенциально имеет большое значение для журналистских методов познания и формирования ценностей. Данные и связанная с ними деятельность будет иметь большее значение в предстоящие годы, с результатами, которые могут оказаться утопическими, антиутопическим или иными [8]. Чтобы оценить значимость и получить максимально достоверные результаты в будущем, возможно будет эффективным рассмотреть особый вклад эпистемологии, экономики и этики больших данных как концептуальных точек опоры, признавая и применяя целостное взаимодействие между ними.

Журналистика быстро изменилась с начала цифровой революции в обществе. Традиционный идеал журналистики заключался в том, чтобы репортеры выступали в качестве независимых источников, пытаясь доносить новости объективно на основе фактов. Хотя эта традиция остается неизменной в некоторых новостных каналах, расширенный доступ к технологиям привел к увеличению числа гражданских журналистов и журналистов-активистов, которые открыто придерживаются субъективной точки зрения, но все же пытаются продвигать эту точку зрения через

призму, которая включает справедливое кадрирование, редактирование и отчеты. Некоторые также используют эти технологии для продвижения пропаганды под видом новостей. К сожалению, Интернет позволил всем стать авторами: любой может создать блог и опубликовать свое мнение в мире, где оно может стать вирусным без какой-либо проверки фактов или редактирования. Журналисты в эпоху цифровых технологий вынуждены работать в мире, где цикл новостей движется быстрее. В результате часто бывает затруднительно найти баланс между своевременной и подробной отчетностью.

Область журналистики включает в себя множество различных узких специальностей, каждая из которых требует различных наборов навыков, инструментов и подходов. Независимо от роли, эпоха цифровой журналистики принесла в отрасль глобальные изменения. Вот как эпоха цифровых технологий повлияла на различные типы журналистских ролей и как они изменились.

В так называемой *традиционной журналистике* репортеры и редакторы работают в местной официальной газете, основной газете города или района. Независимо от месторасположения, у газет есть устоявшаяся история профессионального освещения новостей, нейтральная политическая позиция и тщательное освещение своих регионов.

Большинство традиционных СМИ не успели отреагировать на эпоху цифровых технологий и пострадали из-за этого. Ожидается, что количество рабочих мест в газетах как репортеров, так и редакторов сократится на 40 % в течение следующих 10 лет, а небольшие газеты закроются или сократят объемы операций. Снижение доходов от рекламы, потеря раздела с объявлениями в Интернете и рост затрат вынудит многие новостные агентства со-

кратить штат или напрямую продавать продукцию крупным конгломератам, которые в свою очередь снижают количество публикаций.

*Телерадиовещание и кабельная журналистика.* С момента изобретения радио в начале XX в. ведущие журналисты вещания собирали миллионы зрителей, которые слушали авторитетные голоса о национальных новостях и текущих событиях.

Телерадиовещание и кабельное телевидение по-прежнему актуально, ведущие привлекают миллионы зрителей, будь то вечерние политические шоу или такие программы, как “60 минут”. Это не означает, что переход к цифровым технологиям не повлиял на тележурналистику. Телевизионные новости охватили информационно-развлекательные программы, поскольку они конкурируют за зрителей, демонстрируя истории, вызывающие эмоции и возмущение, по сравнению с более сухими, политически ориентированными историями, оказывающими гораздо большее влияние на реальный мир.

Рост числа независимых телеканалов также поставил под вопрос существование тележурналистики. Приложения для потоковой передачи, такие как Facebook Live и Twitch, позволяют гражданам журналистам и журналистам-предпринимателям распространять интересные видео высокой четкости и привлекать широкую аудиторию в течение нескольких минут. Национальные протесты, такие как марши Джорджа Флойда в 2020 и беспорядки в Капитолии США в январе 2021 г. широко освещались гражданскими журналистами, которые запечатлели невиданные ранее углы социальных волнений и социальных комментариев в центре событий.

*Журналистские расследования.* Лауреат премии “Оскар” 2015 г. за лучший фильм “В центре внимания” рассказал правдивую историю следственной группы Boston Globe, расследовавшей скандал со злоупотреблениями в католической церкви, потрясший город на рубеже. Это было убедительное свидетельство тому, как команда провела трудоемкую работу для убеждения десятков жертв насилия рассказать свои истории, и как команда кропотливо изучала тома физических каталогов и микрофишных записей из печатной и электронной эпох.

Журналистские расследования в эпоху цифровых технологий стали опираться на новые инструменты. Важным событием стало появление больших данных. Возможность просматривать финансовые данные, историю просмотров, журналы сотовых телефонов и всевозможные другие наборы данных дает журналистам-расследователям бесконечные объемы информации, которую необходимо анализировать, чтобы собрать воедино свои истории.

*Спортивная журналистика.* Спортивные обозреватели оставили свой след в культуре своими статьями, а лучшие авторы многих местных газет работают в спортивных секциях, погружая в интересный мир местных спортсменов. Все любят спорт. Будь то профессиональная лига, или любимая местная школьная команда, людям нравится следить за своими любимыми командами и игроками. Сейчас мы становимся свидетелями, как ухудшается их присутствие в сети новых игроков, таких как спортивный сайт по подписке The Athletic, который быстро стал популярным цифровым новостным сайтом для спортивных обозревателей. Теперь, когда спортсмены обращаются к клиентам напрямую через такие сайты, как Players 'Tribune и через учетные записи в социальных сетях, спортивные журналисты занялись аналитикой и журналистикой, чтобы оставаться актуальными.

*Новости печати.* Помимо газет, журналисты работают в журналах, таблоидах и других печатных изданиях. Публикации часто оказываются одними из самых тяжелых в отрасли из-за высоких затрат на печать цветных страниц и фотографий без полей, а также высоких затрат на качественных писателей, которым часто требуется продолжительное время для подготовки материала и на выполнение которых уходят недели или месяцы. Полноформатная журналистика, которая традиционно появлялась в журналах, переместилась на сайты. Бывшие ведущие журналы всё чаще переходят на онлайн-формат.

*Новости социальных сетей.* Журналисты социальных сетей появились только с момента зарождения цифровой журналистики. Первые журналисты в социальных сетях начали появляться на таких сайтах, как Facebook и Twitter, но теперь работают и на фото-сайтах, таких как Instagram, а также на специализированных платформах, таких как Signal и VSCO.

Журналисты социальных сетей идут в ногу со временем: по мере появления новых платформ и привлечения пользователей журналисты социальных сетей оценивают, выгодно ли их личному бренду завоевать там подписчиков. Они также нашли способы монетизировать свою работу с помощью рекламы, спонсируемых сообщений или платного доступа, которые резервируют контент для подписчиков через различные службы.

### **ПЕРСПЕКТИВЫ РАЗВИТИЯ ЖУРНАЛИСТИКИ ДАННЫХ В КАЗАХСТАНЕ**

Несмотря на трудности, казахстанская журналистика успешно осваивает новый мировой тренд – журналистику данных. Согласно исследованию, проведённому в 2014 году технологические преобразования в виде инфографики и визуализации информации, были найдены лишь в двух казахстанских изданиях

как “Казахстанская правда” и “Егемен Қазақстан”. Из онлайн изданий наиболее активно использовали инфографику сайт [www.bnews.kz](http://www.bnews.kz), [www.tengrinews.kz](http://www.tengrinews.kz) [4].

Изучение СМИ Казахстана за 2018-2019 гг. показал оптимистичный результат. Это заметно по методам, которые используются казахстанскими СМИ для создания журналистских историй. Развитию журналистики данных способствуют появление новых государственных порталов, одним из которых является портал “Открытые данные”, с помощью которого возникла возможность находить, загружать и использовать наборы данных, которые создаются государственными организациями. Также использование данных из официального сайта Комитета по статистике Министерства национальной экономики РК (КС МНЭ РК) позволяет трансформировать нечто абстрактное в доступный для понимания материал для читателей и телезрителей [1].

2020 г. нас порадовал появлением новых data-проектов, среди которых [Protenge.kz](http://Protenge.kz) созданный, одним из лучших дата-журналистов Казахстана Джамилей Маричевой. Это аккаунт в Instagram, где Джамия публикует результаты своих исследований. Все они основаны на данных, которые есть в многочисленных официальных документах, находящихся в открытом доступе.

Большого разнообразия проектов, которые были бы построены исключительно по принципам дата-журналистики, не так много, как в западных изданиях, на то имеются причины локального характера, связанные с доступом к открытым данным и др. Но несмотря на различные факторы новое направление развития современной журналистики – дата-журналистика продолжает набирать обороты.

## **ЗАКЛЮЧЕНИЕ**

Истории журналистики данных обычно основаны на больших наборах данных, которые часто состоят из общедоступных данных или данных, собранных с помощью широкой общественности. Data-журналистика принадлежит к виду “точной” журналистики, при которой журналисты используют возможности компьютера для анализа данных. В научном дискурсе еще недостаточно работ, отражающих специфику журналистики данных. В основном это зарубежные статьи, представляющие собой обзоры практик. Рекомендации и методические разработки по работе с большими данными содержатся лишь в материалах, предложенных авторами, которые работают в практической журналистике и являются сотрудниками известных зарубежных изданий - титанов журналистики как The New York Times, Financial Times, The Washington Post, The Pudding. Их исследования отличает от общего потока информации профессиональный и инновационный подход к обработке статистических, текстовых и визуальных данных. Специалисты data-журналистики успешно используют данные для анализа актуальных проблем в области медиа, культуры, экономике и др.

**КОНКУРИРУЮЩИЕ ИНТЕРЕСЫ.** Конфликтов интересов в отношении авторства и публикации нет.

**АВТОРСКИЙ ВКЛАД И ОТВЕТСТВЕННОСТЬ.** Автор участвовал в написании статьи и полностью несет ответственность за предоставление окончательной версии статьи в печать.

## **СПИСОК ЛИТЕРАТУРЫ**

[1]. Барлыбаева, С. Х. Дата-журналистика в Казахстане как медиатренд / С. Х. Барлыбаева, М. А. Малтабаров. – Текст : электронный // Хабаршы. Журналистика сериясы. – № 1 (55) 2020. – С. 22-32.– URL : <https://doi.org/10.26577/HJ.2020.v55.i1.03>

[2]. Шилина, М. Г. Big&Open Data как фактор трансформации профессиональной социальной коммуникации? / М. Г. Шилина. – Текст : непосредственный // Коммуникация. Медиа. Дизайн. – Москва, 2016. – № 3. – С. 19-33.

[3]. Шилина, М. Г. Big data и цифровая датификация как техносциальный феномен. К вопросу формирования научно-теоретической рамки исследования / М. Г. Шилина. – Текст : непосредственный // Социально-политические науки. – № 2. – 2018. – С. 60-65.

[4]. Ибраева, Г. Ж. Журналистика данных – новый тренд казахстанских медиа / Г. Ж. Ибраева. – Текст : непосредственный // Вестн. Казахского нац. ун-та. (“Қазақ университеті”). – 2014 – № 1 (35). – С. 3-10.

[5]. Манович, Л. Теории софт-культуры./ Л. Манович [перевод с английского Асмик Бадоян, Надежда Лебедева] – Нижний Новгород: Красная ласточка, 2017. – 203 с. – Текст : непосредственный.

[6]. Mortality of the British Army, 1858. Scribd.

[7]. Gray Jonathan, Chambers Lucy and Bounegru Liliana (2012). The Data Journalism Handbook. How Journalists Can Use Data to Improve the News, O'Reilly Media, August. Available at: [https://data-journalismhandbook.org/uploads/first\\_book/DataJournalismHandbook-2012.pdf](https://data-journalismhandbook.org/uploads/first_book/DataJournalismHandbook-2012.pdf) (accessed 07 November 2021). (In English).

[8]. Seth C. Lewis, Westlund O. (2014). Big Data and Journalism: Epistemology, expertise, economics, and ethics. November Digital Journalism 3(3). Pp. 447-466. (In English).

## ИНФОРМАЦИЯ ОБ АВТОРЕ

**Гульназ Шаймерген**

магистрант, Международный университет информационных технологий, г. Алматы, Казахстан, улица Манаса 34/1, Алматы 050040/A15MOFO

E-mail: [gul105232@gmail.com](mailto:gul105232@gmail.com)

ORCID [0000-0002-6978-6975](https://orcid.org/0000-0002-6978-6975)

### **Салтанат Абдиназар**

магистрант, Международный университет информационных технологий, г. Алматы, Казахстан, улица Манаса 34/1, Алматы 050040/A15MOFO

ORCID [0000-0001-8616-066X](https://orcid.org/0000-0001-8616-066X)

**Научный руководитель** – Ниязгулова Айгуль Аскарбековна, заведующая кафедрой медиакоммуникации и истории Казахстана, кандидат филологических наук  
[a.nijasgulova@iitu.kz](mailto:a.nijasgulova@iitu.kz)

## **BIG DATA AS A PHENOMENON OF INNOVATIVE JOURNALISM**

**Gulnaz Shaimergeren, Saltanat Abdinazar**

*International University of Information Technologies,  
Almaty, Kazakhstan*

### **ABSTRACT**

The research is devoted to the study of the history of the emergence and development of data journalism, the analysis of the relevance of working with data for journalism at the present stage, as well as the phenomenon of the phenomenon of big data in data journalism and journalism in general. The history of the development of data journalism has studied. The analysis of the main characteristics of big data has carried out.

An overview of the experience of big data analysis in data journalism in foreign media and Kazakhstani journalists successfully using big data in their research have presented. The results of the influence of digital technologies on various types of journalistic roles and how they have changed are presented, as well as trends in the development of domestic publications on data journalism are touched upon.

**KEY WORDS:** journalist, data journalism, big data, mass media, data analysis, infographics, data collection.

**DOI 10.24334/KBSU.2021.1.3.009**

**ORIGINAL ARTICLE SUBMITTED:**26.05.2021

**REVISION SUBMITTED:** 26.06.2021

**FOR CITATION:** Shaimergeren G., Abdinazar S. (2021) Big Data kak fenomen innovacionnoj zhurnalistiki [Big Data as a phenomenon of innovative journalism]. *Vestnik Kabardino-Balkarskogo gosudarstvennogo universiteta: Zhurnalistika. Obrazovaniye. Slovesnost'*. Vol. 1, no. 3, pp. 99-121. DOI 10.24334/KBSU.2021.1.3.009

**COMPETING INTERESTS.** The authors declare that there are no conflicts of interest regarding authorship and publication.

**CONTRIBUTION AND RESPONSIBILITY.** All authors contributed to this article. Authors are solely responsible for providing the final version of the article in print. The final version of the manuscript was approved by all authors.

## **BIBLIOGRAPHY**

[1]. Barlybaeva S.Kh., Maltabarov M.A. (2020). Data-zhurnalistika v Kazhastane kak mediatrend [Data journalism in Kazakhstan as a media trend]. *Khabarshy. Journalism series. No1 (55) 22-32.* <https://doi.org/10.26577/HJ.2020.v55.i1.03>. (In Russian).

[2]. Shilina M.G. (2016). Big&Open Data kak faktor transformacii professional'noj social'noj kommunikacii? [Big & Open Data as a Factor in the Transformation of Professional Social Communication?]. Communication. Media. Design. - Moskva.: NRU HSE, Faculty of Communications, Media and Design. M., No 3. Pp. 19-33. (In Russian).

[3]. Shilina M.G. (2018). Big data i cifrovaya datifikaciya kak tekhnosocial'nyj fenomen. K voprosu formirovaniya nauchno-teoreticheskoy ramki issledovaniya [Big data and digital dating as a techsocial phenomenon. On the question of the formation of the scientific and theoretical framework of the study]. Socio-political sciences. No. 2. Pp. 60-65. (In Russian).

[4]. Ibraeva G.Zh.(2014) ZHurnalistika dannyh – novyj trend kazahstanskih media [Data journalism is a new trend in Kazakhstani media]. Vestn. Kazakhstani nat. un-that. No. 1 (35). Almaty. ("Kazakh university"). (In Russian).

[5]. Manovich L. (2017). Theories of soft culture [translated from English by Hasmik Badoyan, Nadezhda Lebedeva]. - Nizhny Novgorod: Red Swallow. P. 203. (In Russian).

[6]. Mortality of the British Army (1858). Scribd. (In English).

[7]. Gray Jonathan, Chambers Lucy and Bounegru Liliana (2012). The Data Journalism Handbook. How Journalists Can Use Data to Improve the News, O'Reilly Media, August. Available at: [https://datajournalismhandbook.org/uploads/first\\_book/DataJournalismHandbook-2012.pdf](https://datajournalismhandbook.org/uploads/first_book/DataJournalismHandbook-2012.pdf) (accessed 07 November 2021). (In English).

[8]. Seth C. Lewis, Westlund O. (2014). Big Data and Journalism: Epistemology, expertise, economics, and ethics. November Digital Journalism 3(3):447-466. Available at: [https://datajournalismhandbook.org/uploads/first\\_book/DataJournalismHandbook-2012.pdf](https://datajournalismhandbook.org/uploads/first_book/DataJournalismHandbook-2012.pdf) (accessed: 11 December 2021). (In English).

### INFORMATION ABOUT THE AUTHOR

#### **Gulnaz Shaimerger**

Magistrant, International University of Information Technologies, Almaty, Kazakhstan.

Manas street 34/1, Almaty 050040 / A15MOFO

E-mail: [gul105232@gmail.com](mailto:gul105232@gmail.com)

ORCID [0000-0002-6978-6975](https://orcid.org/0000-0002-6978-6975)

#### **Saltanat Abdinazar**

Magistrant, International University of Information Technologies, Almaty, Kazakhstan.

Manas street 34/1, Almaty 050040 / A15MOFO

e-mail: [abdinazarsalta@gmail.com](mailto:abdinazarsalta@gmail.com)

ORCID [0000-0001-8616-066X](https://orcid.org/0000-0001-8616-066X)