ЗАРУБЕЖНАЯ ЖУРНАЛИСТИКА

Шилина Александра Геннадьевна, аналитик, венчурный фонд активного участия в криптосетях «Парадайм Фанд» (Paradigm Fund), Москва, Россия; e-mail: alexandrashilina@mail.ru

ЖУРНАЛИСТИКА ДАННЫХ В ЗАРУБЕЖНОЙ КАЧЕСТВЕННОЙ ПРЕССЕ (КЕЙС СПЕЦИАЛИЗИРОВАННЫХ РЕСУРСОВ ГАЗЕТ *THE GUARDIAN* И *THE NEW YORK TIMES*)

В статье представлены основные результаты исследования публикаций, относимых специалистами к журналистике данных, в зарубежной качественной прессе на примере специализированных ресурсов авторитетных англоязычных изданий «The Guardian» и «The New York Times» за 2014–2016 гг. Проведенный анализ публикаций демонстрирует сходные характеристики по контенту и субъекту. Так, большие данные пока не являются обязательным компонентом публикаций журналистики данных. Источники информации редко обрабатываются самими журналистами. Подготовка материалов ориентирована на моносубъектность: авторство большинства публикаций единоличное. Материалы по жанру представляют собой комментарии к данным экономической и политической тематики, сопровождающиеся одним или двумя вариантами визуализации данных в виде инфографики.

Ключевые слова: *The Guardian Datablog, The New York Times' Upshot, большие данные, количественные данные, журналистика данных.*

Alexandra G. Shilina, analyst at Paradigm Fund (Active Participation in Crypto Networks Venture Fund), Moscow, Russia; e-mail: alexandrashilina@mail.ru

DATA JOURNALISM IN THE FOREIGN QUALITY PRESS (THE CASE OF SPECIALIZED RESOURCES OF THE GUARDIAN AND THE NEW YORK TIMES NEWSPAPERS)

This paper presents the main results of the study of data journalism in the foreign press on the example of materials from the global authoritative newspapers, The Guardian and The New York Times (2014–2016). The analysis of articles shows similar characteristics. Big data is not an essential component of foreign data journalism. Sources of information are rarely processed by the journalists themselves. Foreign data journalism is primarily monosubjective: the authorship of most articles is individual. Data journalism in foreign quality press has certain characteristics that can be described as normative: in terms of genre most of the

materials are comments on economic and political data with one or two data visualizations.

Key words: *The Guardian Datablog, The New York Times' Upshot, big data, quantitative data, data journalism.*

DOI: 10.30547/vestnik.journ.5.2019.135162

В мировой практике СМИ последнего десятилетия прослеживаются очевидные тенденции в использовании данных: влияние новых технологий на процесс создания и тематику журналистского материала; появление новых возможностей доставки информации; расширение интерактивного участия аудитории через творчество и самовыражение[1].

В 2010-х гг. благодаря активному внедрению и использованию больших данных (англ. *big data*) для подготовки материалов СМИ, оформляется новое направление – журналистика данных (англ. *data journalism*). Журналистика данных опирается на традиции расследовательской журналистики, *компьютерной,* или *вычислительной,* журналистики (англ. *computer-assisted reporting – CAR*) и *прецизионной журналистики* (англ. *precision journalism*), основанной на применении научных методов сбора и анализа данных (Meyer, 2002).

Одна из наиболее распространенных дефиниций журналистики данных принадлежит А. Б. Говарду, который определяет ее как процесс «сбора, очистки, организации, анализа, визуализации и публикации данных для создания журналистских материалов» (Howard, 2014). Однако, согласно исследованию шведских ученых Э. Аппельгрен и Г. Нюгрена, журналистику данных пока трудно определить, она состоит из комбинации компетенций и навыков нескольких профессий (Appelgren, Nygren, 2013).

Журналистика данных остается дискуссионным форматом в профессиональном и академическом сообществе. За прошедшие годы данное понятие не получило четкой дефиниции. Если первоначально термин употреблялся для обозначения структурированных, машиночитаемых данных, используемых в журналистике вместе с традиционным текстом (Holovaty, 2006), то сегодня, как показывает практика, рамки того, что можно называть *большими данными* в материалах журналистики данных, становятся всё более размытыми. Так, не всегда публикации, относимые сообществом к

журналистике данных, содержат в качестве первоисточника именно большие данные в их общепризнанном понимании – как серии подходов, инструментов и методов обработки структурированных и неструктурированных машиночитаемых данных объёмом от одного петабайта[2] для получения воспринимаемых человеком результатов[3]. Часто материалы основываются на крупных наборах цифровых данных, не требующих сложных программных инструментов обработки.

В данном исследовании мы разграничиваем большие данные и выделяемые нами *количественные данные* – гибридный тип данных, являющийся основой большинства публикаций зарубежной журналистики данных.

Мы определяем большие данные как массивы данных огромных объемов и значительного многообразия, которые из-за их количества и сложности не могут быть собраны, сохранены и проанализированы без помощи специальных горизонтально масштабируемых программных инструментов обработки, появившихся в конце 2000-х гг.. Чаще всего такие данные оформляются в базы размером от нескольких петабайт, генерируемые самыми разными источниками: социальными медиа, веб-сервисами, мобильными приложениями, интернетом вещей и т. д. Под количественными данными мы предлагаем понимать наборы данных небольших объемов и разнообразия, которые могут быть обработаны человеком с использованием любого из традиционных статистических инструментов. Это могут быть статистические данные, данные социологических исследований, рейтинги и др., которые являются привычными информационными поводами в журналистике.

Данное положение позволяет нам провести разграничение между *журналистикой данных* в её широком понимании, охватывающей целый спектр материалов, публикуемых в соответствующих рубриках, и условной *журналистикой больших данных* – как специализированным направлением в журналистике, материалы которого создаются с использованием информации, полученной из соответствующих по масштабам и сложности обработки цифровых баз данных. Соответственно, в рамках работы мы используем следующее определение журналистики данных: журналистика данных – это направление в журналистике, основой материалов которого является обработка любых так называемых *количественных данных*, которые могут служить информационным поводом

или быть частью материала, обогащая его. При этом особое значение приобретает визуальная презентация этих данных аудитории. В то же время в работе отдельно обращается внимание на журналистику больших данных – направление, в основе которого лежит работа именно с большими данными, в соответствии с их пониманием современными специалистами, как цифровой информации, которую невозможно обрабатывать традиционными способами.

Британский «Гардиан датаблог» *(The Guardian datablog)* является признанным экспертами (Borges-Rey, 2016; Knight, 2015; Stalph, 2017) показательным примером современной европейской журналистики данных. Он создается в 2009 г. в качестве отдельного раздела издания «Гардиан» журналистом и впоследствии его редактором С. Роджерсом. Материалы, основанные на крупных цифровых данных, публикуются регулярно, что иллюстрирует растущую востребованность подобной информации. Помимо текстового контента авторы предоставляют читателям возможность ознакомиться с проанализированными данными, размещая на сайте ссылки на базы данных, послужившие основой публикаций, визуализируя данные различными способами, в том числе интерактивно. Публикации «Гардиан датаблог» исследователи относят к «основной (англ. *general*) ежедневной простой форме журналистики данных» (по Э. Борхес-Рэю), которая создается преимущественно одним автором[4]. При этом, как отмечает М. Найт – исследователь британской журналистики данных, публикации «Гардиан датаблог» тяготеют к варианту «комплексной формы журналистики данных» (Knight, 2015: 69) – мультимедийному конвергентному контенту, включающему в себя текстовую и визуальную составляющую (согласно проведенному автором в 2015 г. контент-анализу публикаций журналистики данных Великобритании).

В исследовании мы также рассматриваем публикации «Нью-Йорк Таймс» – одного из наиболее технически и технологически развитых американских печатных изданий, которое одним из первых практикует создание публикаций, основанных на новых форматах цифровых данных: в 2012 г. оно запускает проект «Год в графике» *(2012: The Year in Graphics)*[5], представив главные события года (выборы в США, Олимпиаду, разрушения от урагана и др.) в формате интерактивных визуализаций. 2013 год «Нью-Йорк Таймс» объявляет годом интерактивного сторителлинга *(2013: The Year in Interactive Storytelling)*[6], в этих рамках издание запускает

проект «Визуализация данных» *(Data Visualization)*. Ежегодно журналисты «Нью-Йорк Таймс» создают интерактивные онлайн-проекты, основанные на больших данных, в рамках серий публикаций, озаглавленных «Год в интерактивном сторителлинге, графике и мультимедиа» *(The Year in Interactive Storytelling, Graphics and Multimedia)*. С апреля 2014 г. материалы журналистики данных размещаются в специализированном новостном разделе – «Апшот» *(The Upshot)*, созданном лауреатом Пулитцеровской премии Д. Леонхардтом.

Для понимания значения больших данных в зарубежной журналистике данных и для выявления характерных особенностей её контента мы предпринимаем кейс-стади ресурсов-лидеров данного направления журналистики – «Гардиан датаблог» и «Апшот». На их примере мы рассматриваем, какое место в материалах журналистики данных отдельных газет занимают большие данные, которые в современных условиях потенциально должны определять её специфику, а также выделяем основные характеристики этого направления, представленные не в отдельных разовых проектах (как, например, в публикациях на основе «Панамского досье»), а в регулярных публикациях новостного типа.

Степень разработанности темы исследования

Практики журналистики данных как в России, так и за рубежом пока в достаточной степени не систематизированы и не отражены в научной литературе. Существующие исследования находятся на стыке различных научных направлений и концепций.

Во-первых, это национальные и институциональные исследования первой волны (2009–2015 гг.), посвященные разбору и описанию отдельных публикаций журналистики данных (Borges-Rey, Stalph, 2018).

Во-вторых, это исследования, связанные с компьютерной, вычислительной журналистикой. Например, работа С. Коэн, Д. Т. Гамильтона и Ф. Тёрнера, в которой авторы утверждают, что для развития медиаиндустрии и преодоления ею финансового и технологического давления журналистам необходимо работать сообща с техническими специалистами, способными предлагать иные возможные алгоритмы создания контента, к примеру путем извлечения данных (Cohen, Hamilton, Turner, 2011). Также это исследование работы зарубежных редакций Т. Флю, К. Спурджиона,

А. Дэниэля, А. Свифта, в котором утверждается, что применение методов вычислительной журналистики создает основу для оригинальных материалов расследовательской журналистики и расширяет возможности форм взаимодействия с аудиторией, привлекая её в Интернете (Flew, Spurgeon, Daniel, Swift, 2012). Это статья Н. Диакопоулоса об инновациях в компьютерной журналистике, в которой – с опирой на теоретическое обобщение пятидесяти пяти концепций, методику системного анализа и ценностно-ориентированного проектирования – производится градация аспектов влияния инноваций в журналистике и выявляются недооцененные области их применения (Diakopoulos, 2012, 2015). Другие работы – работа по социологии вычислительной и алгоритмической журналистики (Anderson, 2013), исследование о датификации в журналистике и о значении концепции открытых данных (Baack, 2015).

В-третьих, это работы по эпистемологии журналистики данных: научные статьи, посвящённые изучению влияния феномена больших данных на современную журналистику (Parasie, 2015; Lewis, Westlund, 2015); исследование влияния инноваций на работу журналистов (Gynnild, 2014) и анализ влияния вычислительной журналистики и журналистики данных на работу в редакциях (Coddington, 2015).

В-четвертых, это работы, посвященные рассмотрению подходов к обучению журналистов программированию, статистике, базовым инструментам, необходимым при создании публикаций журналистики данных: исследование Д. Хьюветта о возможностях и ограничениях в преподавании журналистики данных (Hewett, 2015); изучение преподавания политической расследовательской журналистики на примерах публикаций журналистики данных (Graham, 2015); исследование стратегий образования в журналистике данных на основе сравнительного исследования методик европейских стран (Splendore, Di Salvo, Eberwein, Groenhart et al, 2015); исследование готовности студентов Гонконга к созданию материалов журналистики данных (Yang, Du, 2016); статьи, раскрывающие значение обучения журналистов статистическим методам (Nguyen, Lugo-Ocando, 2015) и навыкам анализа данных в эпоху цифровых технологий (Yarnall, Johnson, Rinne, Ranney, 2008).

В-пятых, это работы исследователей, посвящённые национальным различиям и особенностям журналистики данных разных

стран. В российском академическом сообществе исследования журналистки данных малочисленны: И. Радченко и А. Сакоян (Radchenko, Sakoyan, 2014), А. Валеева (Valeeva, 2017) и С. Вартанов (2017).

Существующие концептуальные работы о сущности феномена журналистики данных, его характеристиках, месте в системе массмедиа и возможных социальных эффектах пока не изучены комплексно ни зарубежными, ни российскими авторами, что свидетельствует не только о его новизне, но и о сложности, обусловливает необходимость поиска релевантных подходов.

Методика исследования

В статье рассматриваются публикации разделов, посвященных журналистике данных, технологически развитых и авторитетных зарубежных изданий – «Гардиан» и «Нью-Йорк Таймс» («Гардиан датаблог» и «Апшот»). Основной исследовательский метод – кейс-стади на базе результатов количественного контент-анализа. Опыт двух изданий может рассматриваться в большей степени именно как кейс, но при этом – за счет показательности объектов – отражающий общее состояние журналистики данных в настоящий момент.

Согласно проведенным подсчетам, в год на сайте «Гардиан» в разделе «Гардиан датаблог» публикуется 60–70 материалов. На первом этапе исследования нами рассматриваются практически все материалы раздела за три года – выборка материалов близка к сплошной: 180 публикаций (по 60 в год). Данное число материалов выбирается для дальнейшего релевантного сравнения проанализированных публикаций «Гардиан датаблог» с материалами «Апшот», которых на первом этапе исследования также рассмотрено 180 (по 60 материалов в год). Публикации американского ресурса были отобраны методом простого случайного отбора (вероятностная выборка) из генеральной совокупности, равной примерно 1800 (в год в «Апшот» размещаются около 600 материалов).

Хронологические рамки охватывают 2014–2016 гг. – три года ровно в середине 2010-х гг., период, достаточный для сбора статистических данных на уровне пилотажного исследования, на этапе, когда основные практики можно считать устоявшимися и представляющими интерес для изучения.

Для выявления тематической и жанровой специфики публикаций, относимых к журналистике данных, они сравниваются с материалами основных изданий «Гардиан» и «Нью-Йорк Таймс». Для этого специально изучаются материалы, размещенные на сайтах газет за один день 28 марта 2018 г. При отборе публикаций выборка сплошная – всего 126 публикаций в двух газетах. В указанную дату и предшествовавшие дни не происходило событий, которые привлекали бы к себе повышенное внимание и потенциально могли бы исказить тематический и жанровый баланс, складывающийся в изучаемых изданиях. При этом массив публикаций, вышедших за один день, видится нам достаточным для того, чтобы дать представление об использовании в газетах определенных тематик и жанров.

Как мы уже упоминали выше, не всегда публикации, относимые сообществом к журналистике данных и размещаемые в соответствующих рубриках, содержат в качестве первоисточника именно большие данные в их признанном понимании, часто материалы основываются на большом наборе количественных данных (статистических, социологических и т. п.). В данной статье мы разделяем *большие данные* в их принятом понимании, как базы данных огромных объемов и методы их обработки, и выделяемые нами *количественные данные* – гибридный тип данных, являющийся основой большинства публикаций зарубежной журналистики.

Наличие в анализируемых публикациях больших данных определяется нами исходя из информации, представленной в статье: прямое указание автора на то, что при создании материала используются большие данные; упоминание автором названия ресурса или размещение в публикации ссылки на базу данных, которые просматриваются нами и выявляется, что они содержат большие данные.

Для релевантного описания специфики журналистики данных (анализа субъекта и контента) публикации сравниваются по набору параметров. За основу взята дополненная схема анализа публикаций журналистики данных Ф. Штальфа (Stalph, 2017).

Количественные параметры анализа публикаций:

- число иллюстраций как результат визуализации данных,
- официальное число авторов публикации (подписей под статьей),
- число профессий, компетенции которых необходимы при ее создании (например, навыки графического дизайнера и разработчика).

Качественные параметры:
- тематика публикации,
- жанр публикации,
- источники информации и их типы.

Категория *тематика* разбивается на следующие достаточно крупные подкатегории: политика, экономика/бизнес, здравоохранение, спорт, преступность, культура, образование/технологии/социальные отношения/природные явления, другое (потенциально – наука, мода, путешествия и пр.). Применяемая в данном исследовании классификация создавалась с опорой на классификатор Ф. Штальфа (Stalph, 2017), который использовался при изучении журналистики данных в ведущих зарубежных изданиях. Кроме того, мы посчитали нецелесообразным выделять очень узкие тематические группы на этапе первичного анализа.

Жанр рассматривается по объединенной классификации А. В. Колесниченко 2008 и 2018 гг. В работе 2008 г. автор выделяет следующие жанры журналистики: комментарий, аналитическая статья, расширенная новость, короткая новость, «песочные часы», экспертное интервью, информационное интервью, личностное интервью, ньюс-фиче, фиче, репортаж, портрет (Колесниченко, 2008). В статье 2018 г. – расширенная новость, аналитическая статья, короткая новость, репортаж, портрет, колонка, случай/житейская история, подборка, рекомендация, реконструкция, тест, рецензия, тест-драйв, мультимедийный лонгрид (формат), опрос (знаменитостей, публики), кейс, ответ, карточки, онлайн-трансляция, расследование, анонсирование книг, обзор матча (Колесниченко, 2018).

Так, в данной работе при изучении жанровых особенностей публикаций категория *жанр* разбивается на следующие подкатегории: «песочные часы», информационное интервью, ньюс-фиче, экспертное интервью, личностное интервью, комментарий (в том числе комментарий к данным), фиче, расширенная новость, аналитическая статья, короткая новость, репортаж, портрет, колонка, случай/житейская история, подборка, рекомендация, реконструкция, тест, рецензия, тест-драйв, мультимедийный лонгрид (формат), опрос (знаменитостей, публики), кейс, ответ, карточки, онлайн-трансляция, расследование, анонсирование книг, обзор матча. Мы понимаем, что публикации журналистики данных могут не включать некоторые из перечисленных жанров, однако подробная классификация представляется нам необходимой при изучении публикаций традиционной журналистики рассматривае-

мых изданий и дальнейшем сравнении выявленных жанровых особенностей с характеристиками публикаций, основанных на больших данных.

Источники данных и их типы изучаются на основе классификации источников информации в публикациях журналистских материалов М. А. Разумовой, которая выделяет следующие их типы: *PR*-источники компании, под которыми подразумеваются официальные заявления компании, пресс-служба, руководство компании, пресс-конференции, сайт компании, корпоративные издания; *IR* источники компании – *IR*-служба, финансовые отчёты, существенные факты, совет директоров, уставные документы, миноритарные акционеры, инвестиционные меморандумы, встречи с аналитиками и инвесторами, годовые отчёты, базы данных по компаниям (СПАРК, ЕГРЮЛ); государственные источники, под которыми подразумеваются законопроекты, заявления официальных лиц, министерств и ведомств, страницы официальных лиц в Интернете, решения судов; экспертный блок источников информации: аналитика, участники рынка, профессиональные ассоциации, консультанты; источники – другие СМИ (российские и иностранные); неофициальные источники, а именно: анонимные источники и конфиденциальные документы; новые медиа – сайты и социальные сети (Разумова, 2016).

Так, категорию анализа *источники данных* публикации мы разделяем на следующие подкатегории:

- Данные, которые могут являться большими данными или содержаться в публикации в качестве контекстного источника к ним. Это данные организаций – международных (ЮНЕСКО, ООН, ОБСЕ, МВФ, ВОЗ и др.), государственных/национальных (министерств, национальных статистических служб и т. п.), общественных организаций/аналитических компаний; данные социальных медиа/поисковиков/приложений («Фэйсбук» *(Facebook),* «Твиттер» *(Twitter)*, «Гугл» *(Google)* и др.); данные других СМИ/онлайн-ресурсов; собственные данные (данные журналистов издания); данные результатов академических исследований;
- Контекстные данные в публикациях: *PR*-источники компании (официальные заявления компании, пресс-служба, руководство компании, пресс-конференции, корпоративные издания); *IR*-источники компании (финансовые отчёты, уставные документы, инвестиционные меморандумы, годо-

вые отчёты, базы данных по компаниям); другие государственные источники (законопроекты, заявления официальных лиц, министерств и ведомств, решения судов);
- Среди контекстных источников информации мы выделяем также экспертный блок источников информации: комментарии или данные научного/академического сообщества; общественных деятелей и представителей бизнеса; комментарии представителей государственных органов.
- Отдельно обращается внимание на то, содержат ли рассматриваемые публикации неофициальные источники (анонимные источники и конфиденциальные документы); или же источник неизвестен/не указан в публикации.

При рассмотрении публикаций отмечается, к каким типам источников информации относится первоисточник: база данных, данные социальных медиа/поисковиков, документ.

Результаты исследования

На первом этапе исследования для понимания специфики контента зарубежной журналистики данных анализируются источники данных в публикациях. Выявляется, что менее половины из них содержат в качестве источника информации большие данные (139 из 360), что позволяет нам сделать вывод о том, что журналистика данных сегодня не обязательно основывается на этом цифровом формате данных как информационном: для написания материала большие данные не являются необходимой составляющей. Часто в разделах, посвященных журналистике данных, размещаются различные материалы, содержащие в основном *количественные данные*. Так, помимо больших данных в «Гардиан датаблог» встречаются следующие типы источников информации: статистические данные национальной службы статистики Великобритании *(Office for National Statistics, ONS)*, опросные данные (например, Института Гэллапа), данные аналитической компании «Ипсос Мори» *(Ipsos Mori)* и исследовательского центра «Пью» *(Pew),* открытые данные государственных порталов Великобритании *(Gov.Uk, YouGov)*. Источниками информации в публикациях «Апшот» часто становятся: статистические данные, данные Бюро переписи населения США *(National Fair Housing Alliance)*, Бюро статистики труда *(U.S. Bureau of Labor Statistics)*, данные аналитических компаний и государственных

порталов открытых данных. Также данные не обязательно собираются, обрабатываются и анализируются самими журналистами: лишь в 62 публикациях из 360 большие данные анализируются при помощи специальных программ обработки больших данных.

Таблица 1

Большие данные в материалах онлайн-ресурсов «Гардиан датаблог» и «Апшот», 2014–2016 гг.

Год	Количество материалов			
	Большие данные – инфоповод	**Большие данные как инфоповод + обработаны автором**	**Большие данные как иллюстрация**	**Большие данные как иллюстрация + обработаны автором**
2014	50	26	1	0
2015	40	14	8	3
2016	49	22	4	2
Итого	**139**	**62**	**13**	**5**

Далее в рамках исследования анализируется специфика контента материалов журналистики, основанной на больших данных. Из заданных 360 публикаций изучаются материалы, основанные на больших данных как информационном поводе, то есть 139 публикаций, отобранных в ходе первого этапа исследования. Так как далее анализируются статьи, основанные на больших данных как основном информационном поводе, источник, обозначаемый нами как инфоповод, соответствует источнику, представляющему собой большие данные.

Выявляется, что основные источники информации в публикациях, основанных на больших данных (139 из 360), – это данные организаций: государственных и национальных – 68 (в 50 из них подобные данные являются инфоповодом и представляют собой большие данные), общественных и аналитических кампаний – 52 (в 37 – инфоповод = большие данные), международных организаций – 19 (в 16 – инфоповод = большие данные). Часто публикации создаются на основе данных других СМИ и онлайн-ресурсов (41 публикация из 139, в 9 информация из подобных источников

является инфоповодом = большими данными). Данные социальных медиа, поисковых систем и приложений содержатся в 20 публикациях (в 18 – инфоповод = большие данные). Собственные данные журналистов «Гардиан» и «Нью-Йорк Таймс» содержат 7 статей. Материалы «Апшот», основанные на больших данных, включают в качестве источников академические исследования 22 (в 4 – результаты исследования представляют собой большие данные). Также в публикациях присутствуют контекстные источники информации: другие государственные источники – законопроекты, заявления министерств и ведомств, решения судов, *PR*-источники и *IR*-источники компаний.

Экспертный блок источников информации в публикациях, основанных на больших данных, формируется комментариями общественных и бизнес-деятелей (в 31 из 139), представителей академического сообщества (в 20 из 139), представителей государственных органов и организаций (в 3 из 139 публикаций).

Отдельно обращается внимание на то, содержат ли рассматриваемые публикации неофициальные источники или, возможно, источник информации неизвестен/не указан. В выборку не попала ни одна статья, содержащая подобную информацию.

Самыми распространенными типами источников информации являются именно базы данных (114 из 139). Также в части публикаций журналистами используются данные, генерируемые социальными сетями и поисковыми системами (в 19 из 139 изученных публикаций). Основанием шести публикаций являются большие данные, представленные в форме документов.

Таким образом, результаты второго этапа исследования показывают, что публикации выбранных для изучения ресурсов, основанные на больших данных, чаще всего содержат в качестве информационного повода или в качестве дополнительного источника информации данные государственных, общественных и аналитических организаций, а также данные других СМИ, онлайн-ресурсов, социальных медиа и поисковых систем, которые по форме являются базами данных. Данные часто дополняются комментариями общественных и бизнес-деятелей, высказываниями представителей академического сообщества.

Таблица 2

Источники и типы данных в материалах онлайн-ресурсов «Гардиан датаблог» и «Апшот», основанных на больших данных как информационном поводе, 2014–2016 гг.

Публикации 2014-2016 гг., n=139		
Данные организаций		
	Источники больших данных, ставшие инфоповодом	**Источники информации, создающие контекст**
Государственные/национальные	50	68
Общественные/аналитические компании	37	52
Международные	16	19
Другие данные		
	Большие данные	**Контекстный источник**
Другие СМИ/онлайн-ресурсы	9	49
Социальные медиа/поисковые системы/ приложения	18	20
Академические исследования	4	22
Собственные данные издания	5	7
Контекстные данные		
Другие гос. источники	8	
PR-источники компании	2	
IR-источники компании	3	
Экспертный блок источников информации (комментарии)		
Общество/бизнес	31	
Государственные органы	3	
Научное/академическое сообщество	20	
Другое		
Публикации, содержащие неофициальные источники	0	
Публикации, в которых источник не известен/не указан	0	
Тип источника данных		
Базы данных	114	
Данные социальных медиа/поисковиков	19	
Документы	6	

Для понимания особенностей контента журналистики данных на третьем этапе исследования изучается тематика и жанр публикаций, основанных на больших данных.

Выявляется, что тематика публикаций «Гардиан датаблог» и «Апшот», основанных на больших данных, не обладает особой спецификой. Большинство материалов посвящено бизнесу и различным аспектам экономики (37 публикаций из 139), а также политическим аспектам (29 из 139). Теме здравоохранения посвящено 20 публикаций, спорту – 11, преступности – 10, культуре – 4. Различным аспектам образования, технологиям, природным явлениям и социальным проблемам – по 2 статьи. В остальных 20 публикациях доминирует другая, специализированная тематика, не охваченная классификатором (путешествия, метеорология, наука и пр.).

Таблица 3

Тематика публикаций, основанных на больших данных, в материалах онлайн-ресурсов «Гардиан датаблог» и «Апшот», 2014–2016 гг.

Тематика публикаций, основанных на больших данных	
Тема	**Число публикаций, n=139**
Экономика/бизнес	37
Политика	29
Здравоохранение	20
Спорт	11
Преступность	10
Культура	4
Образование	2
Природные явления	2
Социальные проблемы	2
Технологии	2
Другое	20

Доминирующий жанр в публикациях – комментарий. Большинство материалов представляют собой авторский текст с пояснением к проанализированным данным (64 из 139). Другие жанры: аналитическая статья (27 из 139) и расширенная новость (25 из 139). Встречаются также материалы в форме карточек, подборок, тестов. Три публикации не были отнесены нами ни к одному жан-

ру, предложенному в классификации А. В. Колесниченко, так как они представляют собой размещенные на сайте иллюстрации без какой-либо текстовой составляющей.

Таблица 4

Использование жанров в материалах, основанных на больших данных, на онлайн-ресурсах «Гардиан датаблог» и «Апшот», 2014–2016 гг.

Жанр	Число публикаций, n=139
Комментарий	64
Аналитическая статья	27
Расширенная новость	25
Карточки	5
Короткая новость	3
Кейс	2
Подборка	2
Тест	2
Экспертное интервью	2
Информационное интервью	1
Репортаж	1
Ньюс-фиче	1
Портрет	1
Другое	3

Как было оговорено выше, для выявления тематической и жанровой специфики журналистики данных статьи специализированных разделов сравниваются с материалами основных изданий «Гардиан» и «Нью-Йорк Таймс». Так, дополнительно изучаются 126 материалов, размещенных на сайтах газет за один день 28 марта 2018 г.

Выявляется, что публикации «Гардиан» и «Нью-Йорк Таймс» в основном посвящены политике (24 публикаций из 126), культуре (19) и спорту (13). Часть статей раскрывает темы преступности (8), экономики (7), технологий (9), здравоохранения (6). В нескольких публикациях освещаются события, связанные с природными явлениями (3) и социальными проблемами (3). Около четверти публикаций (33 из 126) затрагивают другие специализированные темы, не вошедшие в классификатор. Так, значительная часть ма-

териалов посвящена моде (6), науке (5), путешествиям (3), вопросам, связанным с недвижимостью (3), проведением досуга (2), едой (2) и др. В выборку попали также три некролога. Подобное разнообразие тем в публикациях классической журналистики наглядно подчеркивает её отличие от журналистики, материалы которой основываются на больших данных.

Мы видим, что публикации традиционной журналистики так же разнообразны по жанру. Чаще всего они представлены в форме расширенной новости, аналитической статьи, репортажа, короткой новости, интервью. Реже встречаются: обзор, комментарий, рецензия и портрет.

Таблица 5

Соотношение жанров материалов «Гардиан датаблог» и «Апшот», основанных на больших данных, и материалов Гардиан» и «Нью-Йорк Таймс» от 28 марта 2018 г.

Жанр	Доля жанров в публикациях «Гардиан датаблог» и «Апшот», основанных на больших данных, n=139 (%)	Доля жанров в публикациях «Гардиан» и «Нью-Йорк Таймс» от 28 марта 2018 г., n=126 (%)
Комментарий	46	6
Аналитическая статья	19	10
Расширенная новость	18	37
Карточки	4	0
Короткая новость	2	7
Кейс	1	0
Подборка	1	0
Тест	1	0
Экспертное интервью	1	4
Информационное интервью	1	0
Репортаж	1	6
Ньюс-фиче	1	0
Портрет	1	4
Личностное интервью	0	2
Колонка	0	5
Рецензия	0	5

Окончание таблицы 5

Обзор	0	5
Фиче	0	2
Житейская история	0	2
Некролог	0	2
Анкета	0	1
Обзор матча	0	1
Очерк	0	1
Мультимедийный лонгрид	0	2
Другое	2	0

Сравнение публикаций «Гардиан» и «Нью-Йорк Таймс» и публикаций журналистики данных, размещающейся в специализированных разделах, показывает, что публикации, основанные на больших данных, обладают меньшим тематическим и жанровым разнообразием, чем материалы изданий в целом.

Таблица 6

Соотношение тематики публикаций «Гардиан датаблог» и «Апшот», основанных на больших данных, и публикаций «Гардиан» и «Нью-Йорк Таймс» от 28 марта 2018 г.

Тематика	Доля в публикациях «Гардиан датаблог» и «Апшот», основанных на больших данных, n=139 (%)	Доля в публикациях «Гардиан» и «Нью-Йорк Таймс» от 28 марта 2018 г., n=126 (%)
Экономика/бизнес	27	6
Политика	21	19
Здравоохранение	14	5
Спорт	8	10
Преступность	7	6
Культура	3	15
Образование	1	1
Природные явления	1	2
Социальные проблемы	1	2
Технологии	1	7
Другое	14	26

На четвертом этапе исследования для понимания значения визуального компонента и особенностей подачи материала журналистики данных изучается объем иллюстративного материала (диаграмма, карта, график и т. п.) в рассматриваемых публикациях, основанных на больших данных. Из всей выборки (139 статей) только в двух публикациях отсутствует какая-либо визуализация. В большинстве материалов данные иллюстрируются одной (43 из 139), двумя (24 из 139), тремя (17) или четырьмя иллюстрациями (17). Более десяти содержатся в 7 изучаемых публикациях (их число может достигать 50, например в выборке «Апшот» встретилось несколько материалов о различиях в предпочтениях жителей штатов США при телепросмотре, представленные в виде тепловых карт).

Исследование визуальных компонентов подтверждает, что в журналистике данных презентация аудитории материала, в частности визуализация обработанных проанализированных данных, имеет большое значение. Формируется необходимость релевантных способов подачи такой информации в первую очередь за счет визуализации; большие данные обусловливают особый тип контента, представление информации в котором демонстрирует паритетное соотношение вербального текста и визуальной составляющей.

Таблица 7

Визуализация данных в публикациях, основанных на больших данных, на онлайн-ресурсах «Гардиан датаблог» и «Апшот», 2014–2016 гг.

Число иллюстраций к данным	Число публикаций, n=139
Одна	43
Две	24
Три	17
Четыре	17
Пять	12
Шесть	6
Семь	6
Восемь	2
Девять	1
Десять	2
Более десяти	7
Визуализация отсутствует	2

На следующем этапе описывается субъект журналистики данных. На основе изучения числа авторов и компетенций, потенциально необходимых для создания публикации, выявляется специфика работы специалистов по созданию материалов журналистики данных. Для определения числа компетенций анализируются следующие составляющие публикаций: текстовая, визуальная и интерактивная. Предполагается, что каждая составляющая соответствует наличию у автора определенной компетенции: для создания текста необходимы навыки пишущего журналиста; для создания инфографики, для оформления визуализации данных – навыки графического дизайнера (например, работа с *HTML5, CSS* и *Java Script*); для сложных программных решений и придания интерактивности элементам публикации – навыки программиста[7].

Таблица 8

Авторство в публикациях, основанных на больших данных, на онлайн-ресурсах «Гардиан датаблог» и «Апшот», 2014–2016 гг.

Число авторов	**Число публикаций, n=139**
Один	79
Двое	42
Трое	9
Четверо	3
Пятеро	3
Шестеро	2
Двенадцать	1
Число компетенций	**Число публикаций, n=139**
Одна	0
Две	90
Три	49

Выявляется, что для рассмотренного сегмента журналистики данных характерна моносубъектность. В большинстве публикаций авторство единоличное (79 из 139), 42 статьи подписаны двумя фамилиями, 9 – тремя. В большинстве публикаций (90 из 139) присутствует два типа работы (текст, основанный на анализе данных, их статическая визуализация или иллюстрация), требующие навыков двух профессий – журналиста, графического дизайнера или

иллюстратора. Для создания 49 из 139 рассматриваемых публикаций необходимы компетенции трех профессий (журналиста, графического дизайнера и программиста). Исходя из выявленного числа авторов материалов и используемых компетенций, можно было бы предположить, что сотрудникам присуща мультикомпетентность. Однако мы не можем с уверенностью утверждать, что число подписей под публикацией соответствует реальному числу её авторов: возможно, технических специалистов не всегда вписывают. В данной работе мы опираемся лишь на ту информацию, которая представлена в публикации.

Таким образом, авторство большинства публикаций – единоличное. Изучение характеристик субъекта подтверждает, что журналистика данных, представленная в «Гардиан датаблог» и «Апшот» подпадает под категорию *обычной* журналистики данных (англ. *ordinary data journalism*), которая может создаваться одним человеком ежедневно, в отличие от *основательной* (англ. *thorough data journalism*) (De Maeyer, Libert, Domingo, Heinderyckx et al, 2015), которая создается группами, обладающими диапазоном различных навыков.

Выводы

Собранная в рамках данного исследования информация, на наш взгляд, дает возможность для развития академического дискурса, а также общего представления о журналистике данных. В настоящий момент в России она по-прежнему ассоциируется в первую очередь с крупными проектами, часто расследовательского характера, основанными на результатах обработки так называемых больших данных. Однако анализ кейса показательных зарубежных онлайн-ресурсов, принадлежащих ведущим англоязычным газетам, дает основания для выделения нескольких важных для понимания современного состояния журналистики данных тенденций.

Так, в условиях, когда ни в профессиональных, ни в академических кругах пока не сложилось устойчивого и четкого определения журналистики данных, полезным представляется ориентация на использование понятия *больших данных*. Изучение попавших в выборку текстов показало, что при потоковом, конвейерном производстве материалов, определяемых редакциями как журналистика данных, непосредственно большие данные являются основой не

более чем для трети публикаций. В этой ситуации актуальным становится использование при определении, классификации, обсуждении и изучении журналистики данных дополнительных уточняющих критериев. Это позволит, например, разделять условные *журналистику количественных данных* и *журналистику больших данных*.

При этом контент-анализ материалов, которые в рамках рассматриваемого кейса можно отнести к журналистике (больших) данных, показывает, что практики производства и характеристики текста не демонстрируют существенных отличий большинства изученных публикаций от вполне традиционных представлений о современной журналистике. Тематические и жанровые характеристики (при сравнении в нашем случае с контентом материнских изданий) в целом мало отличаются от массива материалов, производимых ежедневными изданиями. Возможно, обращает на себя внимание только более узкий спектр используемых в журналистике (больших) данных жанров и явное доминирование среди них группы аналитических жанров во главе с комментарием.

Что касается подготовки материалов, то можно заметить, что даже журналистика данных, потенциально предполагающая развитие технологических навыков работающих в этом сегменте авторов, пока не свидетельствует о формировании широкого круга специалистов нового типа. Лишь в небольшой доле текстов непосредственный инфоповод на базе результатов обработки больших данных генерируется редакцией (и, возможно, автором), а не сторонними организациями. Если основываться на информации об авторстве изученных материалов, наличие в большинстве публикаций одного автора может говорить о формировании у журналистов компетенций графического дизайнера или программиста. Однако невысокая сложность решенных в большинстве материалов технологических задач может говорить о справедливости предложенного зарубежными специалистами разделения журналистики данных на *обычную* и *основательную*, в которой большая часть текстов укладывается в первую категорию.

Примечания

[1] *UNESCO Future arts, media, and entertainment: seeds for 2020.* New York. 2011.
[2] Петабайт – единица измерения количества информации, равная 10^{15} байт.

[3] Manyika J. et al. (2011) *Big data: The next frontier for innovation, competition, and productivity.* McKinsey Global Institute, June. McKinsey (9 August 2011). P. 11.

[4] Э. Борхерс-Рэй выделяет три формы журналистики данных: а) основная, общая (англ. *main, general*) ежедневная форма журналистики данных; б) сложные расследовательские проекты в) геймированная (англ. *gamified*) журналистика данных, ориентированная на развлечение и максимальное вовлечение аудитории. Borges-Rey, E (2016) Unravelling data journalism: A study of data journalism practice in British newsrooms. *Journalism Practice* 10 (7): 833–843.

[5] 2012: The Year in Graphics. *The New York Times*. Available at: http://www.nytimes.com/interactive/2012/12/30/multimedia/2012-the-year-in-graphics.html (accessed: 16.09.2019)

[6] 2013: The Year in Interactive Storytelling. *The New York Times*. Available at: http://www.nytimes.com/newsgraphics/2013/12/30/year-in-interactive-storytelling/ (accessed: 16.09.2019)

[7] Работа со специализированными инструментами обработки больших данных и умение их анализировать рассматриваются нами как отдельная компетенция только на первом этапе исследования – для понимания уровня самостоятельности журналиста при поиске информационного повода и работе с ним и степени оригинальности материала.

Библиография

Вартанов С. А. Большие данные в онлайн-СМИ: подходы и стратегии использования // Медиаскоп. 2017. Вып. 4.

Колесниченко А. В. Практическая журналистика: учеб. пособие. М.: Изд-во Моск. ун-та, 2008.

Колесниченко А. В. Востребованность жанров журналистских текстов в онлайновых СМИ // Вестн. Моск. ун-та. Сер. 10: Журналистика. 2018. № 1. С. 26–42. DOI: 10.30547/vestnik.journ.1.2018.2642

Разумова М. А. Тематические приоритеты и база источников информации деловых СМИ (на примере газет «Коммерсантъ» и «Ведомости»): дис. ... канд. филол. наук. М., 2016.

Anderson C. W. (2013) Towards a sociology of computational and algorithmic journalism. *New Media & Society* 15 (7): 1005–1021. DOI: https://doi.org/10.1177/1461444812465137

Appelgren E., Nygren G. (2013) *Data Journalism in Sweden – Opportunities and Challenges. A Case Study of Brottspejl at Sveriges Television (STV)*. Stockholm: Södertõrn University.

Baack S. (2015) Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism. *Big Data & Society* 2 (2). DOI: https://doi.org/10.1177/2053951715594634

Borges-Rey E. (2016) Unravelling data journalism: A study of data journalism practice in British newsrooms. *Journalism Practice* 10 (7): 833–843. DOI: https://doi.org/10.1080/17512786.2016.1159921

Borges-Rey E., Stalph F. (2018*) A random walk through data + journalism: assessing the data journalism ecosystem. Paper presented on NODA-2018 (The Nordic Data Journalism Conference) 15–17 Mar.* Stockholm: Södertörn University.

Coddington M. (2015) Clarifying Journalism's Quantitative Turn: A typology for evaluating data journalism, computational journalism, and computer-assisted reporting. *Digital Journalism* 3 (3). DOI: https://doi.org/10.1080/21670811.2014.976400

Cohen S., Hamilton J. T., Turner F. (2011) Computational journalism. *Communications of the ACM* 54 (10): 66–71. DOI: 10.1145/2001269.2001288

De Maeyer J., Libert M., Domingo D., Heinderyckx F., Le Cam F. (2015) Waiting for Data Journalism. A qualitative assessment of the anecdotal take-up of data journalism in French-speaking Belgium. *Digital Journalism* 3 (3): 432–446. Available at: http: www.tandfonline.com/doi/full/10.1080/21670811.2014.976415 (accessed: 16.09.2019). DOI: https://doi.org/10.1080/21670811.2014.976415

Diakopoulos N. (2012) *Cultivating the Landscape of Innovation in Computational Journalism*. Tow-Knight Center for Entrepreneurial Journalism.

Diakopoulos N. (2015) Algorithmic Accountability. Journalistic investigation of computational power structures. *Digital Journalism* 3 (3): 398–415. Available at: http://www.tandfonline.com/doi/full/10.1080/21670811.2014. 76411 (accessed: 16.09.2019). DOI: 10.1080/21670811.2014.976411

Flew T., Spurgeon C., Daniel A., Swift A. (2012) The promise of computational journalism. *Journalism Practice* 6 (2): 157–171. DOI: 10.1080/17512786.2011.616655

Graham C. (2015) By the Numbers. Data Journalism Projects as a Means of Teaching Political Investigative Reporting. *Asia Pacific Media Educator* 25 (2): 247–261. Available at: http://journals.sagepub.com/doi/full/10.1177/1326365X15604936 (accessed: 16.09.2019). DOI: https://doi.org/10.1177/1326365X15604936

Gynnild A. (2014) Journalism Innovation Leads to Innovation Journalism: The Impact of Computational Exploration on Changing Mindsets. *Journalism* 15 (6): 713–730. DOI: https://doi.org/10.1177/1464884913486393

Hewett J. (2015) Learning to teach data journalism: Innovation, influence and constraints. *Journalism* 17 (1): 119–137. Available at: http: journals.sagepub.com/doi/full/10.1177/1464884915612681 (accessed: 16.09.2019). DOI: https://doi.org/10.1177/1464884915612681

Holovaty A. (2006) *A fundamental way newspaper sites need to change*. Available at: http://www.holovaty.com/writing/fundamental-change/ (accessed: 16.09.2019).

Howard A. B. (2014) *The Art and Science of Data-Driven Journalism*. Columbia University Academic Commons. DOI: https://doi.org/10.7916/D8Q531V1.

Knight M. (2015) Data journalism in the UK: a preliminary analysis of form and content. *Journal of Media Practice* 16 (1): 55–72. Available at: http://www.tandfonline.com/doi/full/10.1080/14682753.2015.1015801 (accessed: 16.09.2019). DOI: https://doi.org/10.1080/14682753.2015.1015801

Lewis S. C., Westlund O. (2015) Big Data and Journalism: Epistemology, expertise, economics, and ethics. *Digital Journalism* 3 (3). DOI: 10.1080/21670811.2014.976418

Meyer P. (2002) *Precision Journalism: A Reporter's Introduction to Social Science Methods, 4th ed*. Oxford: Rowman & Littlefield.

Nguyen A., Lugo-Ocando J. (2015) The state of data and statistics in journalism and journalism education: Issues and debates. *Journalism* 17 (1): 3–17.

Available at: http://journals.sagepub.com/doi/full/10.1177/1464884915593234 (accessed: 16.09.2019). DOI: 10.1177/1464884915593234

Parasie S. (2015) Data-driven revelation? Epistemological tensions in investigative journalism in the age of ‚big data'. *Digital Journalism* 3 (3): 364–380. DOI: https://doi.org/10.1080/21670811.2014.976408

Radchenko I., Sakoyan A. (2014) The View on Open Data and Data Journalism: Cases, Educational Resources and Current Trends. In D.I. Ignatov, M.Y. Khachay., A. Panchenko, et al. (eds.) *Analysis of Images, Social Networks and Texts*. Cham: Springer. Pp. 47–54. DOI: https://doi.org/10.1177/1464884917700667

Royal C. (2010) *The journalist as programmer: A case study of the New York Times interactive news technology department. Paper presented at the Anais do International Symposium in Online Journalism*. Austin: The University of Texas at Austin, Austin.

Stalph F. (2017) Classifying Data Journalism. A Content Analysis of Daily Data-Driven Stories. *Journalism Practice* 12 (10): 1332–1350. DOI: 10.1080/17512786.2017.1386583.

Splendore S., Di Salvo P., Eberwein T., Groenhart H., Kus M., Porlezza C. (2015) Educational strategies in data journalism: A comparative study of six European countries. *Journalism* 2015 17 (1): 138–152. Available at: http://journals.sagepub.com/doi/full/10.1177/1464884915612683 (accessed: 16.09.2019). DOI: https://doi.org/10.1177/1464884915612683

Valeeva A. (2017) *Open Data in a Closed Political System: Open Data Investigative Journalism in Russia*. Reuters Institute for the Study of Journalism. University of Oxford. Available at: https://reutersinstitute.politics.ox.ac.uk/our-research/open-data-closed-political-system-open-data-investigative-journalism-russia (accessed: 16.09.2019).

Yang F., Du Y. R. (2016) Storytelling in the Age of Big Data: Hong Kong Students' Readiness and Attitude towards Data Journalism. *Asia Pacific Media Educator* 26 (2): 148–162. Available at: http:.journals.sagepub.com/doi/full/10.1177/1326365X16673168 (accessed: 16.09.2019). DOI: https://doi.org/10.1177/1326365X16673168

Yarnall L., Johnson J. T., Rinne L., Ranney M. A. (2008) How Post-secondary Journalism Educators Teach Advanced CAR Data Analysis Skills in the Digital Age. *Journalism & Mass Communication Educator* 63 (2): 146–164. Available at: http:.journals.sagepub.com/doi/abs/10.1177/107769580806300204 (accessed: 16.09.2019).

Notes

UNESCO Future arts, media, and entertainment: seeds for 2020. New York. 2011.

Manyika J. et al. (2011) Big Data: The Next Frontier for Innovation, Competition, and Productivity. McKinsey Global Institute, June. McKinsey (9 August 2011). P. 11.

2012: The Year in Graphics. *The New York Times*. Available at: http://www.nytimes.com/interactive/2012/12/30/multimedia/2012-the-year-in-graphics.html (accessed: 16.09.2019).

2013: The Year in Interactive Storytelling. The New York Times. Available at: http://www.nytimes.com/newsgraphics/2013/12/30/year-in-interactive-storytelling/ (accessed: 16.09.2019).

References

Anderson C. W. (2013) Towards a Sociology of Computational and Algorithmic Journalism. *New Media & Society* 15 (7): 1005–1021. DOI: https://doi.org/10.1177/1461444812465137

Appelgren E., Nygren G. (2013) *Data Journalism in Sweden – Opportunities and Challenges. A Case Study of Brottspejl at Sveriges Television (STV)*. Stockholm: Södertõrn University.

Baack S. (2015) Datafication and Empowerment: How the Open Data Movement Re-articulates Notions of Democracy, Participation, and Journalism. *Big Data & Society* 2 (2). DOI: https://doi.org/10.1177/2053951715594634

Borges-Rey E. (2016) Unravelling Data Journalism: a Study of Data Journalism Practice in British Newsrooms. *Journalism Practice* 10 (7): 833–843. DOI: https://doi.org/10.1080/17512786.2016.1159921

Borges-Rey E., Stalph F. (2018*) A Random Walk Through Data + Journalism: Assessing the Data Journalism Ecosystem. Paper presented on NODA-2018 (The Nordic Data Journalism Conference) 15-17 Mar*. Stockholm: Södertörn University.

Coddington M. (2015) Clarifying Journalism's Quantitative Turn: a Typology for Evaluating Data Journalism, Computational Journalism, and Computer-Assisted Reporting. *Digital Journalism* 3 (3). DOI: https://doi.org/10.1080/21670811.2014.976400

Cohen S., Hamilton J. T., Turner F. (2011) Computational Journalism. *Communications of the ACM* 54 (10): 66–71. DOI: 10.1145/2001269.2001288

De Maeyer J., Libert M., Domingo D., Heinderyckx F., Le Cam F. (2015) Waiting for Data Journalism. A Qualitative Assessment of the Anecdotal Take-up of Data Journalism in French-Speaking Belgium. *Digital Journalism* 3 (3): 432–446. Available at: http: www.tandfonline.com/doi/full/10.1080/21670811.2014.976415 (accessed: 16.09.2019). DOI: https://doi.org/10.1080/21670811.2014.976415

Diakopoulos N. (2012) *Cultivating the Landscape of Innovation in Computational Journalism*. Tow-Knight Center for Entrepreneurial Journalism.

Diakopoulos N. (2015) Algorithmic Accountability. Journalistic Investigation of Computational Power Structures. *Digital Journalism* 3 (3): 398–415. Available at: http://www.tandfonline.com/doi/full/10.1080/21670811.2014.976411 (accessed: 16.09.2019). DOI: 10.1080/21670811.2014.976411

Flew T., Spurgeon C., Daniel A., Swift A. (2012) The Promise of Computational Journalism. *Journalism Practice* 6 (2): 157–171. DOI: 10.1080/17512786.2011.616655

Graham C. (2015) By the Numbers. Data Journalism Projects as a Means of Teaching Political Investigative Reporting. *Asia Pacific Media Educator* 25 (2): 247–261. Available at: http://journals.sagepub.com/doi/full/10.1177/1326365X15604936 (accessed: 16.09.2019). DOI: https://doi.org/10.1177/1326365X15604936

Gynnild A. (2014) Journalism Innovation Leads to Innovation Journalism: the Impact of Computational Exploration on Changing Mindsets. *Journalism* 15 (6): 713–730. DOI: https://doi.org/10.1177/1464884913486393

Hewett J. (2015) Learning to Teach Data Journalism: Innovation, Influence and Constraints. *Journalism* 17 (1): 119–137. Available at: http: journals.sagepub.com/doi/full/10.1177/1464884915612681 (accessed: 16.09.2019). DOI: https://doi.org/10.1177/1464884915612681

Holovaty A. (2006) *A Fundamental Way Newspaper Sites Need to Change*. Available at: http://www.holovaty.com/writing/fundamental-change/ (accessed: 16.09.2019).

Howard A. B. (2014) *The Art and Science of Data-Driven Journalism*. Columbia University Academic Commons. DOI: https://doi.org/10.7916/D8Q531V1.

Knight M. (2015) Data Journalism in the UK: a Preliminary Analysis of Form and Content. *Journal of Media Practice* 16 (1): 55–72. Available at: http://www.tandfonline.com/doi/full/10.1080/14682753.2015.1015801 (accessed: 16.09.2019). DOI: https://doi.org/10.1080/14682753.2015.1015801

Kolesnichenko A. V. (2008) *Prakticheskaya zhurnalistika: ucheb. posobie* [Practical Journalism. A Study Guide]. Moscow: Moscow St. Univ. Publ. (In Russian)

Kolesnichenko A. V. (2018) Vostrebovannost' zhanrov zhurnalistskikh tekstov v onlaynovykh SMI [The Demand for Genres of Journalistic Texts in Online Media]. *Vestn. Mosk. un-ta. Ser. 10: Zhurnalistika* 1: 26–42. (In Russian)

Lewis S. C., Westlund O. (2015) Big Data and Journalism: Epistemology, Expertise, Economics, and Ethics. *Digital Journalism* 3 (3). DOI: 10.1080/21670811.2014.976418

Meyer P. (2002) *Precision Journalism: A Reporter's Introduction to Social Science Methods, 4th ed*. Oxford: Rowman & Littlefield.

Nguyen A., Lugo-Ocando J. (2015) The State of Data and Statistics in Journalism and Journalism Education: Issues and Debates. *Journalism* 17 (1): 3–17. Available at: http://journals.sagepub.com/doi/full/10.1177/1464884915593234 (accessed: 16.09.2019). DOI: 10.1177/1464884915593234

Parasie S. (2015) Data-Driven Revelation? Epistemological Tensions in Investigative Journalism in the Age of 'Big Data'. *Digital Journalism* 3 (3): 364–380. DOI: https://doi.org/10.1080/21670811.2014.976408

Radchenko I., Sakoyan A. (2014) The View on Open Data and Data Journalism: Cases, Educational Resources and Current Trends. In D.I. Ignatov, M.Y. Khachay., A. Panchenko, et al. (eds.) *Analysis of Images, Social Networks and Texts*. Cham: Springer. Pp. 47–54. DOI: https://doi.org/10.1177/1464884917700667

Razumova M. A. (2016) *Tematicheskie prioritety i baza istochnikov informatsii delovykh SMI (na primere gazet «Kommersant"» i «Vedomosti»)*: dis. ... kand. filol. nauk. [Thematic Priorities and the Database of Information Sources of

Business Media (on the example of the Kommersant and Vedomosti newspapers): PhD philol. sci. diss.]. (In Russian)

Royal C. (2010) *The Journalist as Programmer: a Case Study of the New York Times Interactive News Technology Department. Paper presented at the Anais do International Symposium in Online Journalism*. Austin: The University of Texas at Austin, Austin.

Stalph F. (2017) Classifying Data Journalism. A Content Analysis of Daily Data-Driven Stories. *Journalism Practice* 12 (10): 1332–1350. DOI: 10.1080/17512786.2017.1386583.

Splendore S., Di Salvo P., Eberwein T., Groenhart H., Kus M., Porlezza C. (2015) Educational Strategies in Data Journalism: a Comparative Study of Six European Countries. *Journalism* 2015 17 (1): 138–152. Available at: http://journals.sagepub.com/doi/full/10.1177/1464884915612683 (accessed: 16.09.2019). DOI: https://doi.org/10.1177/1464884915612683

Valeeva A. (2017) *Open Data in a Closed Political System: Open Data Investigative Journalism in Russia*. Reuters Institute for the Study of Journalism. University of Oxford. Available at: https://reutersinstitute.politics.ox.ac.uk/our-research/open-data-closed-political-system-open-data-investigative-journalism-russia (accessed: 16.09.2019). Measurements in the Era of Big Data: Concepts and Examples]. *Vestn. Mosk. un-ta. Ser. 10: Zhurnalistika* 3: 37–57. (In Russian)

Yang F., Du Y. R. (2016) Storytelling in the Age of Big Data: Hong Kong Students' Readiness and Attitude Towards Data Journalism. *Asia Pacific Media Educator* 26 (2): 148–162. Available at: http:.journals.sagepub.com/doi/full/10.1177/1326365X16673168 (accessed: 16.09.2019). DOI: https://doi.org/10.1177/1326365X16673168

Yarnall L., Johnson J. T., Rinne L., Ranney M. A. (2008) How Post-Secondary Journalism Educators Teach Advanced CAR Data Analysis Skills in the Digital Age. *Journalism & Mass Communication Educator* 63 (2): 146–164. Available at: http:.journals.sagepub.com/doi/abs/10.1177/107769580806300204 (accessed: 16.09.2019).

Поступила в редакцию
06.06.2019