

Использование больших данных в междисциплинарных исследованиях на примере макрорегиона Большое Средиземноморье

УДК 311.2

DOI 10.26425/2658-347X-2022-5-3-24-30

Получено 29.04.2022

Доработано после рецензирования 27.06.2022

Принято 05.07.2022

Ярмак Ольга Валерьевна

Канд. социол. наук, доц., зав. каф. социальных коммуникаций, Севастопольский государственный университет, г. Севастополь, Россия

E-mail: olga_yarmak@inbox.ru

ORCID: 0000-0002-5025-9112

SPIN-код: 8215-7744

Саркисов Дмитрий Григорьевич

Студент магистратуры, Севастопольский государственный университет, г. Севастополь, Россия

ORCID: 0000-0002-5166-7628

E-mail: Sarkisoff.mitya@yandex.ua

Жаркова Татьяна Владимировна

Студент магистратуры, Севастопольский государственный университет, г. Севастополь, Россия

ORCID: 0000-0003-3993-1366

E-mail: tayanatsu@gmail.com

АННОТАЦИЯ

В настоящее время происходит интенсивное внедрение технологий сбора и анализа больших данных в социальную, экономическую, политическую сферы в странах и макрорегионах, том числе в Большом Средиземноморье. Этот факт актуализирует перспективы междисциплинарных исследований на стыке социологии, политологии и регионоведения. В статье рассмотрены преимущества и риски использования указанных технологий в перечисленных сферах на примере Испании и стран Ближнего Востока. К основным проблемам и угрозам авторы отнесли: нарушение конфиденциальности персональных данных и рост манипуляций с ними

в политических целях; возникновение неправомерных форм самоорганизации граждан; переход интернет-пользователей на анонимные аккаунты (вследствие практики «приватизации» персональных данных негосударственными и террористическими формированиями в региональных пространствах), препятствующий качественному сбору данных. Отмечена необходимость решения перечисленных проблем на международном уровне. Показан диапазон возможностей, открываемый работой с большими данными, для обеспечения прагматичного и эффективного социально-политического прогнозирования и оперативного регионоведения.

Ключевые слова

Большие данные, оперативное регионоведение, макрорегион, Большое Средиземноморье, информационные технологии, сбор и анализ данных, персональные данные, конфиденциальность, прогнозирование, таргетирование, политические технологии

Для цитирования

Ярмак О.В., Жаркова Т.В., Саркисов Д.Г. Использование больших данных в междисциплинарных исследованиях на примере макрорегиона Большое Средиземноморье // Цифровая социология. 2022. Т. 5. № 3. С. 24–30.

Благодарности

Статья выполнена при поддержке программы «Приоритет-2030» ФГАОУ ВО «Севастопольский государственный университет» (стратегический проект № 5).

© Ярмак О.В., Жаркова Т.В., Саркисов Д.Г., 2022.

Статья доступна по лицензии Creative Commons «Attribution» («Атрибуция») 4.0. всемирная (<http://creativecommons.org/licenses/by/4.0/>).



The use of big data in interdisciplinary research on example of the Greater Mediterranean macroregion

Received 29.04.2022 Revised 27.06.2022 Accepted 05.07.2022

Olga V. Yarmak

Cand. Sci. (Sociol.), Assoc. Prof., Head of the Department of Social Communications, Sevastopol State University, Sevastopol, Russia

ORCID: 0000-0002-5025-9112

SPIN-code: 8215-7744

E-mail: olga_yarmak@inbox.ru

Dmitriy G. Sarkisov

Graduate Degree, Sevastopol State University, Sevastopol, Russia

ORCID: 0000-0002-5166-7628

E-mail: Sarkisoff.mitya@yandex.ua

Tatiana V. Zharkova

Graduate Degree, Sevastopol State University, Sevastopol, Russia

ORCID: 0000-0003-3993-1366

E-mail: tayanatsu@gmail.com

ABSTRACT

Big data collection and analysis technologies are being intensively introduced into the social, economic and political spheres in countries and macroregions, including the Mediterranean region. This fact actualises the prospects of interdisciplinary research at the intersection of sociology, political science and regional studies. The article examines the benefits and risks of using these technologies in listed areas, taking Spain and the Middle East as examples. Among the main problems and threats, the authors consider the following: personal data confidentiality breach, and increased data manipulation for political purposes, the emergence

of illegitimate forms of citizens self-organisation, the transition of Internet users to anonymous accounts due to the practice of "privatization" of personal data by non-state and terrorist groups in regional spaces, preventing the qualitative collection of data. The need to solve these problems at the international level was noted. The range of possibilities offered by working with big data to provide pragmatic and effective socio-political forecasting and operational regional studies is shown.

Keywords

Big data, regional studies, macroregion, Big Mediterranean region, information technologies, data collection and analysis, personal data, confidentiality, forecasting, targeting, political technologies

For citation

Yarmak O.V., Zharkova T.V., Sarkisov D.G. (2022) The use of big data in interdisciplinary research on example of the Greater Mediterranean macroregion. *Digital sociology*. Vol. 5, no 3, pp. 24–30. DOI: 10.26425/2658-347X-2021-5-3-24-30

Acknowledgements. The article was supported by the Priority-2030 programme of the Sevastopol State University, strategic project No. 5.



ВВЕДЕНИЕ / INTRODUCTION

В современных реалиях на смену привычным военно-политическим союзам приходит сетевая дипломатия, коррелирующая с многосторонними социальными структурами и требующая новых форм управления. В связи с этим широко используемые цифровые технологии в политике и регионоведении приобретают актуальность в силу как имеющихся угроз, так и открывающихся возможностей.

Обращение авторов к данной теме связано с неоднозначностью применения технологий анализа больших данных в процессе достижения политических целей, непрозрачностью процедуры сбора больших данных и дальнейшего их использования, различием типов данных для единого анализа, а также с вызовами, появляющимися вследствие использования больших данных в управлении массовым сознанием и коммуникациями.

Потенциал применения технологий сбора и анализа больших данных имеет спектр возможностей, которые могут приводить к конфликтным ситуациям в социально-политической сфере. Ярким примером таких ситуаций является избирательный процесс, когда использование цифровых технологий приводит к возникновению и распространению фейковых новостей (от англ. fake news). Негативный аспект использования больших данных в политических процессах проявляется в создании ситуаций, когда цифровые технологии способствуют возникновению неправомерных форм самоорганизации граждан.

ЛИТЕРАТУРНЫЙ ОБЗОР / LITERATURE REVIEW

Феномен больших данных в междисциплинарных исследованиях стал одним из наиболее актуальных и дискутируемых среди ученых. С одной стороны, он относится к технологическим возможностям, и позволяет повысить эффективность таргетирования и информационного воздействия на целевую аудиторию [Вершинин, 2004]. С другой стороны, имеется некоторая амбивалентность подходов к оцениванию технических возможностей больших данных в современной социологической науке, прежде всего потому, что, как упомянуто выше, прозрачность методов, используемых в современной практике политической сферы, крайне невысока. Несмотря на это, остается необходимость в рассмотрении указанного феномена относительно тех возможностей и угроз, которые могут возникать с введением технологий сбора и анализа больших данных в общественно-политической сфере.

Работа с большими данными позволяет регулярно отслеживать настроение общественности, создавать выгодный политический имидж, формировать наиболее успешную стратегию политического противостояния конкурентам [Николаенко, 2019]. При правильном применении технология сбора и анализа больших данных имеет огромный потенциал влияния на политическую повестку любого из государств, однако главным критерием успешности такой деятельности является степень распространенности Глобальной Сети в обществе.

Большие данные открывают новые перспективы, но вместе с тем создают риски. Например, облегчается работа с огромными объемами информации, появляется возможность выстроить верную стратегию на основе более точного прогноза. С другой стороны, возникает риск нарушения закона о персональных данных и использования политиками персональных данных пользователей Сети в личных интересах с целью манипуляции мнением общественности и пропаганды [Черникова, 2015; Bolsover and Howard, 2017].

ЦЕЛЬ И МЕТОДЫ ИССЛЕДОВАНИЯ / PURPOSE AND METHODS OF THE RESEARCH

Цель представленного исследования – выявление потенциала и рисков применения технологий сбора и анализа больших данных в политической сфере на примере отдельного макро-региона. Работа выполнена в структурно-функциональной ориентации, что позволило рассматривать анализ больших данных как функцию политических, управленческих процессов и процесса влияния на массовое сознание. В рамках исследования применялся метод киберметрии для контент-анализа медиаресурсов стран Большого Средиземноморья на предмет использования больших данных, в том числе примеров их внедрения в Испании и странах Ближнего Востока (Израиль, арабские государства). Эмпирической базой исследования стали вторичные социологические данные последних лет.

ИСПОЛЬЗОВАНИЕ БОЛЬШИХ ДАННЫХ В СОЦИОЛОГИИ / THE USE OF BIG DATA IN SOCIOLOGY

Тенденция актуализации цифровой социологии и применения в ней цифрового инструментария обусловлена не только интенсивным расширением интернет-пространства из-за количественного увеличения веб-ресурсов, но и качественными изменениями в нем: сайты кардинально

перестраивают внешний вид и функционал, активно отвечают на баг-репорты пользователей, интернет вещей приносит в финансовую систему все большую виртуальную денежную массу, а поведение интернет-пользователей становится все более активным и многогранным.

Наблюдается рост недоверия к традиционной социологии среди населения, связанный с активным воздействием интервьюера на респондента, в то время как кредит социального доверия переносится в интернет-среду и реализуется посредством цифровых коммуникаций [Журавлёва, 2015].

Традиционные социологи все больше обращаются к комбинированным методам, проводя опросы на интернет-площадках, экзитполы (от англ. exit poll – опрос на выходе) в онлайн-формах на избирательных участках и cookie-файлах подходящих по выборке групп [Крыштановская, 2018]. При использовании этих методик возникает проблема: данные могут быть некомплементарны, а качество сопоставлений незначительно ввиду отсутствия четких и надежных алгоритмов и инструкций анализа.

Современная социология выделяет четыре типа исследований, использующих большие данные:

- анализ массивов текста средствами вычислительной лингвистики;
- сетевой анализ;
- машинное обучение;
- онлайн-эксперименты [McFarland et al, 2015].

Следует учитывать, что этот процесс двухсторонний, и некоторые методологические находки из области больших данных могут быть использованы при анализе относительно небольших массивов данных, а возможно – в рамках качественных исследований.

По мнению Connelly и др. [2016], одним из ключевых отличий современных больших данных и данных, получаемых, например, в результате переписи населения, является то, что они генерируются самостоятельно, а не собираются с определенными целями.

В Сети можно найти множество коммуникативных площадок, каждая из которых потенциально представляет собой интересный кейс трансформации социального взаимодействия. Существует множество подходов к изучению новых форм коммуникации. В рамках настоящего исследования авторы проанализировали методологические возможности, которые предоставляются процессом сетевых коммуникаций.

Стоит отметить, что большая часть современных социологов использует метод анализа больших данных в своей практике, так как он позволяет оперативно получать информацию о происходящих общественных процессах из реакций в социальных

медиа [Щекотин, 2021]. К достоинствам метода анализа больших данных относится способность последних описывать генеральную совокупность, автоматически выгружаемую из исследуемого интернет-сегмента. Кроме того, рассматриваемый метод экономит временные, финансовые и человеческие ресурсы [Одинцов, 2017].

Однако, говоря о достоинствах данного метода, необходимо помнить о наиболее частом заблуждении исследователей: большие данные, с одной стороны, объективны, но с другой, исследователь не влияет на этот инструмент, так как ключевую роль выполняют электронные механизмы, не контролируемые человеком, поэтому крайне важно внимательно относиться к получаемой информации.

ИСПОЛЬЗОВАНИЕ БОЛЬШИХ ДАННЫХ В ПОЛИТИКЕ / THE USE OF BIG DATA IN POLITICS

Анализ больших данных в политической сфере стал одним из наиболее эффективных методов, имеющих новейшие технологические возможности и позволяющих повысить точность таргетирования целевых аудиторий. Однако существуют определенные риски и угрозы использования больших данных в политической сфере, которые рассмотрены ниже.

1. Инструменты, которые предоставляет рассматриваемая технология, дают возможность манипулировать мнением пользователей, формируя «воронку недоверия» к политическим организациям и субъектам. Феномен «воронки недоверия» раскрывается в возрастании «интеллектуального долга» (Дж. Зиттейн), возникновении угрозы использования персональных данных в политических интересах, возрастании недоверия к применяющим эту технологию политическим объединениям и кандидатам.

В работе американского исследователя Дж. Зиттейна описана категория «интеллектуальный долг» – это ситуация активного внедрения новейших технологий в повседневную жизнь, при этом без осмысления принципов их действия, что впоследствии приводит к возможности спекулировать информацией, скрывая некоторые детали и особенности технических приборов (например, «умные колонки» Amazon собирают широкий спектр данных о пользователе). Именно поэтому вопросом, наиболее волнующим экспертные сообщества, является безопасность персональных данных, и ключевая угроза здесь – нелегальное использование данных при легальном доступе к ним.

2. Происходит размытие институтов демократии. Политическая таргетированная реклама, использование анализа политического поведения в социальных медиа, воздействие на личность избирателя и его дальнейший электоральный выбор приводит к несоответствию с ключевыми положениями демократии. Такая ситуация моделирует решения индивидов и оказывает на них информационное воздействие, способствует уязвимости решений при электоральном выборе. Примером такого успешного «информационного заражения» избирателей является предвыборная компания Д. Трампа, по окончании которой возникли дискуссии о судьбе устройства плебисцитарной демократии [Соловей, 2017].

3. Возникает угроза нарушения конфиденциальности персональных данных в Интернете. Современные исследования доказывают, что благодаря цифровым технологиям стала возможна точная идентификация пользователей (их пола, дат рождения и т. д.) в публичном доступе, благодаря чему, казалось бы, бесполезная информация о личности человека стала значимой и необходимой для осуществления политических манипуляций.

4. Алгоритмы имеют ряд когнитивных ограничений, поэтому некоторые аспекты могут игнорироваться, а ошибочный алгоритм может и вовсе изменить итоговый результат, приведя к неожиданным последствиям. Такие сбои являются основными аргументами использования больших данных и их алгоритмирования в политических и электоральных процессах. Например, в Чикаго были разработаны специальные алгоритмы в области криминалистики для отслеживания правонарушений, которые анализировали взаимосвязь расы и количества совершенных преступлений, что привело к увеличению количества полицейских патрулей за представителями конкретных рас и создало напряженность и конфликтность в этой среде.

ИСПОЛЬЗОВАНИЕ БОЛЬШИХ ДАННЫХ В РЕГИОНОВЕДЕНИИ: БОЛЬШОЕ СРЕДИЗЕМНОМОРЬЕ / THE USE OF BIG DATA IN REGIONAL STUDIES: THE GREATER MEDITERRANEAN

В регионоведческом анализе большие данные позволяют регулярно отслеживать общественные настроения в регионах и больших пространствах, поэтому они, с одной стороны, способствуют созданию политических имиджей регионов, с другой, формируют стратегии политического противостояния между странами. Метод анализа

больших данных в регионоведении применяется, в частности, в динамично развивающемся макрорегионе Большого Средиземноморья.

Как отмечено выше, цифровые данные пользователей динамично применяются в политике и регионоведении с целью манипуляционных и пропагандистских воздействий на пользователей Интернета (это относится и к странам, входящим в указанный макрорегион, а также некоторым другим).

Рассмотрим кейс протестов в странах Ближнего Востока в 2011–2012 гг., причинами которых стал разрыв между реальным благосостоянием арабской молодежи и декларируемым экономическим уровнем молодежи западных государств [Беляков, Прохвятилов, 2012]. Многие аналитики «Арабской весны» назвали революционные события «твиттерной революцией», потому что использовались собранные большие данные (сообщения из социальной сети Twitter) с геолокациями авторов, полными текстами числом прочитавших и подписавшихся на определенного автора пользователей. Протесты «Арабской весны» начинались с обычных твитов пользователей, среди которых были те, кто фактически распространял инструкции протеста (куда необходимо сегодня прийти, как дойти до места проведения протеста и как следует себя вести). Анализ данных твитов позволил сделать вывод о том, что именно такие коммуникации и способствовали появлению протестов. «Открытые» пользователи отслеживались государственными органами, при этом не менее интересно то, что их нередко подкупали, чтобы организовать задержание или увести протестующих с места проведения акции без потерь. Таким образом, одной из задач использования больших данных в описываемой ситуации стало как формирование протеста, так и управление им [Элдииб, 2017].

Большие данные необходимы для мирных целей. Например, в июле 2014 г. в столице Испании были улучшены электронные сервисы, позволяющие гражданам поддерживать связь с чиновниками. Данный проект был оценен в 14,7 млн долл., а в его основу легли инструменты анализа больших данных, с помощью которых стало возможным мгновенно реагировать на законодательные инициативы.

В еще одной стране средиземноморского региона, Израиле, компания Silicon Wadi стала серьезным конкурентом США в области цифровых технологий, в том числе по сбору и анализу больших данных в государственном и частном секторах. Компания в будущем планирует создать сети, содержащие данные датчиков визуального

интеллекта. Создаваемая база данных (англ. combat computing cloud) упростит поддержание коммуникации в онлайн-режиме, передачу информации, упорядочит действия между подразделениями Армии обороны Израиля.

ЗАКЛЮЧЕНИЕ / CONCLUSION

В настоящее время технологии сбора и анализа больших данных максимально интенсивно внедряются в социальную, экономическую, политическую сферы стран и регионов, что обусловлено возможностями поиска неясных взаимосвязей. Благодаря этому совершенствуются качество управления регионами и повышается эффективность социально-политического прогнозирования и оперативного регулирования процессов больших пространств.

Несмотря на преимущества использования больших данных, имеются серьезные риски и ограничения, связанные с нарушением конфиден-

циальности, ростом числа манипуляций данными, реакцией онлайн-пользователей на практику «приватизации» персональных данных в результате деятельности негосударственных и террористических формирований в региональных пространствах. Вследствие этого сформировалась тенденция к переходу пользователей Интернета на анонимные аккаунты, что, в свою очередь, препятствует качественному сбору и анализу больших данных для дальнейшего прогнозирования. Существует не менее значимый спектр угроз, связанный с этической стороной использования больших данных. Поэтому, по мнению авторов, необходимо на международном уровне инициировать процесс переговоров по вопросу законодательного регулирования использования больших данных.

СПИСОК ЛИТЕРАТУРЫ

- Фролов Д. Б., Прохвятилов В. В., Беляков Н. В. (2012). New media и Арабская весна // *Международные отношения*. № 1. С. 60–67. (Англ. яз).
- Вершинин М. С. (2004). Политическая коммуникация в информационном обществе: перспективные направления исследований: перспективные направления исследований // *Актуальные проблемы теории коммуникации*. Сборник научных трудов. СПб. Изд-во СПбГПУ. С. 253–270.
- Журавлёва Е. Ю. (2015). Социология в сетевой цифровой среде: от вычислительных и электронных социальных наук к цифровым социальным исследованиям // *СОЦИС*. № 8 (376). С. 25–34.
- Крыштановская О. В. (2018). Бесконтактная социология: новые формы исследований в цифровую эпоху // *Цифровая социология*, Т. 1, № 1. С. 4–8. <https://doi.org/10.26425/2658-347X-2018-1-4-8>
- Николаенко Г. А. (2019). Перспективы использования цифровых следов исследователей для анализа их коммуникативных стратегий (на примере социальной сети ResearchGate) // *Социология науки и технологий*. Т. 10, № 2. С. 15–17. <https://doi.org/10.24411/2079-0910-2019-12005>
- Одинцов А. В. (2017). Социология общественного мнения и вызов Big Data // *Мониторинг общественного мнения: экономические и социальные перемены* № 3. С. 30–43. <https://doi.org/10.14515/monitoring.2017.3.04>
- Соловей В. Д. (2017). Цифровая мифология и избирательная кампания Дональда Трампа // *Полис. Политические исследования*. № 5. С. 122–132. <https://doi.org/10.17976/jpps/2017.05.09>
- Черникова В. Е. (2015). Манипуляция массовым сознанием как феномен информационного общества // *Теория и практика общественного развития*. № 3. С. 141–144.
- Щекотин Е. В. (2021). Цифровые следы как новый источник данных о качестве жизни и благополучии: обзор современных тенденций // *Вестник Томского государственного университета*. № 467. С. 170–181. <https://doi.org/10.17223/15617793/467/21>
- Элдиш А. М. (2017). Революционные события в Египте 2011–2012 гг. в контексте «Арабской весны»: общее и особенное: дисс. ... канд. ист. наук: 07.00.03. Нижегород. гос. ун-т им. Н.И. Лобачевского. Нижний Новгород. 218 с.
- Bolsover G., Howard P. (2017). Computational propaganda and political big data: moving toward a more critical research agenda // *Big Data*. V. 5, no. 4. Pp. 273–276. <https://doi.org/10.1089/big.2017.29024.cpr>
- Connelly R., Playford C. J., Gayle V., Dibben C. (2016). The role of administrative data in the big data revolution in social science research // *Social Science Research*. V. 59. Pp. 1–12. <https://doi.org/10.1016/j.ssresearch.2016.04.015>
- McFarland D. A., Lewis K., Goldberg A. (2015). Sociology in the era of big data: the ascent of forensic social science // *The American Sociologist*. No. 1. Pp. 12–35. <https://doi.org/10.1007/s12108-015-9291-8>

REFERENCES

- Belyakov N. V. and Prokhvatilov V. V. (2012), “New Media and the Arab Spring”, *Mezhdunarodnye otnosheniya*, no. 1, pp. 60–67.
- Bolsover G. and Howard P. (2017), “Computer propaganda and political big data: the transition to a more critical research agenda”, *Big Data*, vol. 5, no. 4, pp. 273–276, <https://doi.org/10.1089/big.2017.29024.cpr>
- Chernikova V. E. (2015), “Manipulation of mass consciousness as a phenomenon of the information society”, *Theory and practice of social development = Teoriya i praktika obshchestvennogo razvitiya*, no. 3, pp. 141–144.
- Connelly R., Playford C. J., Gayle V. and Dibben C. (2016), “The role of administrative data in the big data revolution in social science research”, *Social Science Research*, vol. 59, pp. 1–12, <https://doi.org/10.1016/j.ssresearch.2016.04.015>
- Eldiib A. M. (2017), *Revolutionary events in Egypt in the context of the “Arab Spring”: general and special: dissertation ... cand. hist. sci.: 07.00.03*, National Research Nizhny Novgorod State University, Nizhny Novgorod, Russia (in Russian).
- Kryshtanovskaya O. V. (2018), “Contactless sociology: new forms of research in a digital age”, *Digital Sociology*, vol. 1, no. 1, pp. 4–8, <https://doi.org/10.26425/2658-347X-2018-1-4-8>
- McFarland D. A., Lewis K. and Goldberg A. (2015), “Sociology in the era of big data: the rise of forensic social science”, *American Sociologist*, no. 1, pp. 12–35, <https://doi.org/10.1007/s12108-015-9291-8>
- Nikolaenko G. A. (2019), “The Perspectives of Using Digital Traces of Researchers for Analyzing their Communication Strategies (by the Example of the Social Network “ResearchGate”)”, *Sociology of Science and Technology*, vol. 10, no. 2, pp. 15–17, <https://doi.org/10.24411/2079-0910-2019-12005>
- Odintsov A. V. (2017), “Sociology of public opinion and the Big Data challenge”, *The Monitoring of Public Opinion: Economic and Social Changes Journal*, no. 3, pp. 30–43, <https://doi.org/10.14515/monitoring.2017.3.04>
- Shchekotin E. V. (2021), Digital footprints as a new source of data on quality of life and well-being: An overview of current trends, *Tomsk State University Journal*, no. 467, pp. 170–181, <https://doi.org/10.17223/15617793/467/21>
- Solovey V. D. (2017), “Digital mythology and Donald Trump electoral campaign”, *Polis. Political Studies*, no. 5, pp. 122–132, <https://doi.org/10.17976/jpps/2017.05.09>
- Vershinin M. S. (2004), “Political communication in the information society: promising research directions”, *Actual problems of the theory of communication. Collection of scientific papers*, St. Petersburg, SPbPU Publ. House, pp. 98–107.
- Zhuravleva E. Yu. (2015), “Sociology in digital environment: towards digital social research”, *SOCIS*, no. 8(376), pp. 25–34.