

**Аванесов А.Ю.**,  
студент НИУ БелГУ (Белгород, Россия)  
Научный руководитель – Р.А. Травин

## **ТЕХНОЛОГИИ BIG DATA В СОЦИАЛЬНЫХ НАУКАХ И РОЛЬ НЕМАШИННЫХ МЕТОДОВ В АНАЛИЗЕ «БОЛЬШИХ ДАННЫХ»**

В условиях информатизации общества наука не в состоянии без помощи автоматизированных систем охватить и проанализировать все информационные потоки. Однако когда речь идет о социальных науках, необходимо понимать, что эти системы не могут быть использованы для выявления некоторых качественных характеристик «больших данных» ввиду их сложности, особенно когда речь идет о массивах информации, которые имеют субъективный характер: материалы СМИ, публицистические статьи, тексты в социальных сетях и т.д. В рамках проведения политических исследований показатели, рассчитанные с помощью автоматизированных систем на основе определенных индикаторов и моделей анализа, могут быть неверно интерпретированы в рамках использования «машинных» методов, и релевантный результат можно получить только на основе анализа, проведенного специалистом.

Тем не менее, развитие технологий Big Data позволило представителям социальных наук проводить принципиально новые исследования, охватывая огромные массивы информации, и ученые, в том числе политологи, активно занимаются разработкой как теоретических аспектов анализа «больших данных», так и практического их применения. В данном случае можно вспомнить работы С.В. Володенкова [4], Н.А. Цветковой, Н.М. Кузнецова [8] и других. А применение технологий Big Data в реальной политике можно увидеть на примере избирательной кампании Д. Трампа в 2016 году, в которой успеха удалось добиться во многом за счет продвижения таргетированной рекламы в социальных сетях на основе анализа «больших данных».

На практике недостаточно получить результаты машинного анализа данных, гораздо важнее умение работать с ними. И, например, в уже упомянутой президентской кампании в США 2016 года этой работой занималась компания Cambridge Analytica, с привлечением политологов, психологов и других специалистов для обработки данных [6].

Если говорить об исследованиях, связанных с анализом текстов, следует отметить, что работу автоматизированных систем анализа Big Data необходимо постоянно корректировать, так как они не в каждом случае способны эффективно анализировать подтексты и посылы. Следуя той же логике, можно сказать, что в некоторых случаях необходимо учитывать биографию или идеологические предпочтения автора, а также социально-политический контекст, цензуры, определяющие степень возможного отклонения позиции автора от каких-либо рамок, и только после этого делать реальные оценки, так как одни и те же данные в разных ситуациях могут быть оценены по-разному.

Таким образом, использование технологий Big Data нельзя идеализировать, и несмотря на их огромный потенциал в социальных науках, в их использовании есть ряд проблем, связанных, помимо технологических сложностей, с интерпретацией данных, доступом к ним, проблемой «фейков», если мы говорим о «цифровых следах», и др. Эти барьеры говорят о том, что в большинстве случаев нельзя делать выводы, основываясь исключительно на анализе Big Data, в отрыве от других методов и без критического осмысления.

Несмотря на это, некоторые сторонники эмпиризма, которых воодушевило развитие технологий Big Data, опираясь, например, на позицию аналитика Криса Андерсона, считают, что «социологи и политологи должны уступить свое место аналитикам данных, которые не обременены теоретическим багажом социальных наук» [5]. Однако такая позиция представляется неконструктивной ввиду сложности социально-ориентированных наук, с которой аналитики данных не в состоянии справиться самостоятельно. Один из примеров уже приведен выше. Кроме того, если говорить о прикладных задачах в политике, надо учитывать, что политическая сфера динамична, и для адаптации технологий Big Data необходимо постоянное переосмысление реалий специалистами-политологами. Технологии, используемые в одной ситуации, могут дать совсем другой результат в другой, даже в одном и том же обществе, так как условия, ожидания и восприятие людей меняются, и эти изменения не всегда могут просчитать аналитики данных.

Для подтверждения этой позиции приведем достаточно показательный, на наш взгляд, пример из политической жизни современной Армении. В стране весной 2018 года произошла так называемая «бархатная революция». На момент ее проведения социальные сети были полны гневных отзывов о прежних властях, а социально-экономическая обстановка в Армении вызывала недовольство населения. Для того, чтобы это понять, было достаточно применить использованный автором статьи метод включенного наблюдения, и можно с уверенностью предположить, что любой анализ «больших данных» это бы подтвердил. Косвенно вывод о высокой степени протестных настроений в армянском обществе 2018 года подтверждается количеством людей, вышедших на акции протеста с требованием отставки премьер-министра (более 100 тысяч человек), и результатами досрочных парламентских выборов 2018 года, при которых бывшая партия власти даже не преодолела порог голосов и не получила мест в парламенте [1, с. 62].

Далее обратимся к событиям в Армении, развивающимся с ноября 2020 года. Несмотря на то, что социально-экономическая и политическая обстановка в стране гораздо хуже, чем та, которая была двумя годами ранее (отсутствие реального роста благосостояния населения, сложная ситуация с распространением коронавирусной инфекции, при которой страна вышла на первое место в Европе по количеству случаев COVID-19 на 100 тыс. человек [2], беспрецедентный рост государственного долга [7, с. 122], поражение в войне и др.), нынешней оппозиции удалось собрать на акциях протеста изначально лишь

несколько тысяч человек [3]. Возникает закономерный вопрос: почему так произошло и как применить технологии Big Data для анализа этих причин? Так как для событий последних лет актуально их отражение в социальных сетях, целесообразно применить автоматизированные системы для анализа «цифровых следов» армянского сегмента этих сетей. Но если при интерпретации данных и при определении критериев анализа использовать логику, которая была актуальна для Армении 2018 года, результаты могут поставить аналитиков данных в тупик. А для того, чтобы понять, как в данном случае применить Big Data, необходимо использование «теоретического багажа» политологов и социологов.

Например, если для 2018 года для анализа протестного потенциала более актуально измерение социально-экономического состояния граждан, то для 2021 года первостепенным будет измерение психологической составляющей, в том числе степени деморализации населения. И если в 2018 году состав оппозиции не являлся определяющим фактором «революции», то в 2021 году целесообразнее изучать именно состав оппозиции, направить системы Big Data на анализ этих составляющих. Однако аналитики данных вряд ли способны в данном случае эффективно задать направление прикладных исследований и верно интерпретировать некоторые результаты без помощи политологов и социологов.

На данном примере можно видеть одно и то же общество с разницей всего в два года, но принципиально разные индикаторы отражения протестного потенциала в, казалось бы, похожих ситуациях. Мы обозначили лишь некоторые факторы, но и их достаточно для того, чтобы понять, что в реальной политической практике нельзя опираться лишь на «сухие цифры», и аналитики данных не способны заменить политологов и социологов.

Также к числу наиболее актуальных проблем при анализе «больших данных» в социальных сетях в будущем можно отнести намеренное искажение информации, хакерство. И если на данный момент тренд автоматизированного анализа «больших данных» является новым и в целом такие методы эффективны, то, возможно, через 5-10 лет, в случае если для решения прикладных задач политологи будут применять системы анализа Big Data повсеместно, появятся силы, которые будут пытаться влиять на эти процессы с использованием «вбросов» ложных «цифровых следов». И представляется, что с нынешним уровнем развития «фабрик троллей» и индустрии искусственного интеллекта системы Big Data вряд ли смогут их распознать. Хотя эта проблема присуща скорее не автоматизированным технологиям как таковым, а среде, для которой они предназначены.

Конечно, «большие данные» не ограничиваются социальными сетями, и обозначенные проблемы не в равной степени характерны для всех массивов информации, однако они имеют место быть, поэтому полностью полагаться на автоматизированную обработку «больших данных» нельзя.

Исходя из этого, а также учитывая тренд информатизации общества, можно утверждать, что увеличение информационных потоков требует изменения логики их анализа, и в этой связи можно говорить о перспективности в том числе неавтоматизированного анализа больших объемов данных, при котором технологии Big Data эффективны по большей части при сборе первичной информации. А работа качественного характера, если объемы позволяют это сделать, проводится исследователями самостоятельно. Хотя она и требует больших временных затрат, результат от нее в некоторых случаях более релевантен. Такой подход к анализу Big Data может заметно облегчить задачи исследователей, а сама технология, хоть и является в определенной степени революционной, не способна полностью заменить традиционные методы социальных наук. При этом с развитием Big Data роль политологов и социологов, вопреки мнению сторонников крайнего эмпиризма, возрастает.

### **Литература:**

1. Айвазян, А. С. Парламентские выборы 2018 года в Армении: перспективы блока «Мой шаг» / А. С. Айвазян // Научно-аналитический вестник Института Европы РАН. – 2018. – № 6. – С.61–64.
2. Армения лидирует среди стран Европы по количеству случаев COVID-19 на 100 тыс. человек (ГРАФИК) [Электронный ресурс] // АМИ «Новости-Армения», 05.09.2020. – Режим доступа: <https://newsarmenia.am/news/armenia/armeniya-lidruet-sredi-stran-evropy-po-kolichestvu-sluchaev-covid-19-na-100-tys-chelovek-grafik/>. – Дата доступа: 24.02.2021.
3. В Ереване на площади Свободы прошел митинг противников Пашиняна [Электронный ресурс] // ТАСС, 16.11.2020. – Режим доступа: <https://tass.ru/mezhdunarodnaya-panorama/10018237>. – Дата доступа: 24.02.2021.
4. Володенков, С. В. Big Data как инструмент воздействия на современный политический процесс: особенности, потенциал и акторы / С. В. Володенков // Журнал политических исследований. – 2019. – Т. 3. – № 1. – С.7–13.
5. Губа, К. Большие данные в социологии: новые данные, новая социология? / К. Губа // Социологическое обозрение. – 2018. – Т. 17. – № 1. – С. 213–236.
6. Расследование Das Magazin: как Big Data и пара ученых обеспечили победу Трампу и Brexit [Электронный ресурс] // The Insider, 06.12.2016. – Режим доступа: <https://theins.ru/politika/38490>. – Дата доступа: 24.02.2021.
7. Социально-экономическое положение Республики Армения в январе-декабре 2020 г. (налогово-бюджетный сектор) [Электронный ресурс] // Статистический комитет Республики Армения. – Режим доступа: [https://www.armstat.am/file/article/sv\\_12\\_20r\\_221.pdf](https://www.armstat.am/file/article/sv_12_20r_221.pdf). – Дата доступа: 24.02.2021.
8. Цветкова, Н. А. Феномен дипломатии больших данных в мировой политике / Н. А. Цветкова, Н. М. Кузнецов // Вестник РГГУ. Серия: политология. История. Международные отношения. – 2020. – № 4. – С. 27–44.