

Применение больших данных в социальных науках: история и современное состояние

Научный руководитель – Мартыненко Татьяна Сергеевна

Путиенко Александра Сергеевна

Студент (магистр)

Московский государственный университет имени М.В.Ломоносова, Социологический факультет, Москва, Россия

E-mail: putienko900@gmail.com

Тема использования больших данных в современной социологии все чаще поднимается на круглых столах, конференциях и других видах встреч научного сообщества. Появление больших данных считают “угрозой”, “проблемой”, “вызовом” для традиционных методов сбора данных в социологии.

Прежде чем говорить об использовании больших данных в современной социологии, рассмотрим основные характеристики данного понятия.

Как феномен большие данные появились в рамках 3D-моделирования, их основные характеристики были выделены Д. Леини [6] и обозначены как 3V:

- volume (размер): огромные по объему, состоят из терабайтов и петабайтов данных;
- velocity (скорость): высоко динамичные, создаются в реальном времени или близком к нему;
- variety (разнообразие): различающиеся по типам данных, структурированные и неструктурированные, часто связаны с временем и местом (объединены метаданными).

Активное развитие в исследовательских науках большие данные начали получать с 2014 года, тогда ученые выделили следующие характеристики этого феномена:

- изменчивость, которая предполагает, что данные изменчивы и их интерпретация существенно зависит от контекста;
- исчерпывающая полнота;
- «высокое разрешение»;
- реляционность, то есть наличие у различных баз данных общих полей, которые позволяют их объединять в единые массивы;
- расширяемость и масштабируемость - предполагают, что в зависимости от задач исследования массив данных можно расширять за счет присоединения новых массивов;

Исследуя использования больших данных в социологических исследованиях, мы можем обратиться к теории информационного детерминизма и информационных технологий. Данную проблематику изучали американский ученый Т. Нейсбит [3], отечественные исследователи Пушкин В.Г. и А.Д.Урсул в 1989 году [3]. По мнению ученых разрабатывающих данную теорию: использование больших данных стало возможным благодаря технологиям, которые связаны с цифровыми аппаратами и программными средствами.

Использование больших данных относят к разделу цифровой социологии, так как накопленные теоретические решения в вышеупомянутых теориях от званий до методов стали применяться в цифровой социологии.

Началом использования в социологии больших данных мы можем считать статью, которую выпустили в 2009 в журнале «Science» «Вычислительные социальные науки» [2], которая вызвала наибольший интерес у социологов, и является одной из самых цитируемых по проблематике больших данных в социальных науках.

Гэри Кинг [7] говорил, что наибольший результат в социальной науке возможен, когда присутствуют три условия: инновационные статистические методы, новая компьютерная наука и оригинальные теории отдельных областей знаний.

Для социологов внедрение технологий больших данных в их работы стало возможностью продвинуться в точности анализа проводимого исследования. Особенно полезными становятся online-данные, благодаря которым исследователь может в режиме реального времени получать данные о поведении пользователя и придавать их анализу без задействования таких способов исследования, как интервью или опрос, например.

Кроме того появление социальных сетей и регистрация online-поведения позволила проводить online-эксперименты. Один из самых известных экспериментов в 2012 году был основан на данных более 60 миллионов пользователей Facebook. Эксперимент тестировал возможность социального влияния при политической мобилизации на примере выборов в Конгресс.

В 2014 году в обзоре «Annual Review Sociology» подводятся итоги о вкладе данных о поведении пользователей из источников больших данных в развитие социальных наук. В итогах значится, что, несмотря на инновационность, исследования проводятся низкого качества. У социологов появляется доступ к большим количествам данных о поведении пользователя, однако, выборка этих данных не всегда дает осмысленные и качественные результаты из-за своего количества.

В одной из своих работ в 2016 году Волков В., Скугаревский Д., Титаев К. писали, что при использовании больших данных «исследователь не только лишен возможности влиять на инструмент, но и нередко не может наблюдать его в действии» [1].

Таким образом, мы можем сказать, что активное использование больших данных в социальных науках, а в частности в социологии началось с 2009 года, и, на данный момент, исследования с применением технологий больших данных не получили широкого распространения. Более того, точность данных исследований из-за возможных масштабов данных не доказана и подвергается сомнению.

Источники и литература

- 1) Волков В., Скугаревский Д., Титаев К., Проблемы и перспективы исследований на основе Big Data (на примере социологии права) // Социологические исследования, 2016. No. 1. С. 48–57.
- 2) Губа К. Большие данные в социологии: новые данные, новая социология? // Социологическое обозрение. 2018., № 1. С. 41–64.
- 3) Ницевич В. Ф., Цифровая социология: теоретико-методологические истоки и основания // Цифровая социология, 2018., № 1. С. 18-28.
- 4) Пирогов А.И. (2011). Интернет и информационная безопасность личности. // Вестник Московской государственной академии делового администрирования., 2011. №1 (7). С. 45–54.
- 5) Пушкин В.Г., Урсул А.Д., Информатика, кибернетика, интеллект, философские науки. Кишинев: Штиинца., 1989.
- 6) Сивков Д., Большие данные в этнографии: вызовы и возможности // Социология науки и технологий., 2017. №1. С. 56–68.
- 7) King G., Restructuring the Social Sciences: Reflections from Harvard's Institute for Quantitative Social Science // PS: Political Science & Politics., 2013. Vol 47. 1. С. 165–72.
- 8) Laney D. 3D data management: Controlling data volume, velocity and variety. In: Meta Group, (2001)
- 9) Lazer D., Pentland A., Adamic L., Aral S., Barabasi A-L., Brewer D., Christakis N., Contractor N., Fowler J., Gutmann M., Jebara T., King G., Macy M., Roy D., Van Alstyne M. (2009). Computational Social Science // Science. Vol 323.5915. С. 721–723.