



Московский педагогический
государственный университет

Е. В. Бродовская, А. Ю. Домбровская

БОЛЬШИЕ ДАННЫЕ В ИССЛЕДОВАНИИ ПОЛИТИЧЕСКИХ ПРОЦЕССОВ

Москва
2018

Елена Бродовская

**Большие данные в исследованиях
политических процессов**

«МПУ»

2018

УДК 31(075.8)
ББК 66.04я73

Бродовская Е. В.

Большие данные в исследованиях политических процессов /
Е. В. Бродовская — «МПГУ», 2018

ISBN 978-5-4263-0712-4

В учебном пособии раскрываются теоретико-методологические и методические аспекты больших данных как метода анализа политических процессов по цифровым маркерам. Раскрываются основы и типы киберметрического анализа как способа изучения больших данных. Показаны особенности таких методов интеллектуального поиска цифровых маркеров политических процессов в интернет-контенте, как Predictor Mining, Data Mining, проанализированы эвристические и методические основы этих методов социального компьютеринга. Представлены образцы разработки поисковых запросов и техник киберметрии с применением онлайн-сервисов для мониторинга социальных медиа. Приведены и проанализированы конкретные примеры киберметрических исследований в сфере изучения политических процессов по цифровым следам. Приводятся задания для самостоятельного освоения студентами техники применения больших данных в анализе политических процессов. Рекомендуется магистрантам направлений «Интернет-технологии в политике и управлении», «Политическое прогнозирование и проектирование» и аспирантам направления обучения «Политология», преподавателям высшей школы, читающим предметы, связанные с оценкой роли Интернета в современных реалиях, ученым, исследующим проблемы интернет-коммуникации с применением киберметрического анализа.

УДК 31(075.8)
ББК 66.04я73

ISBN 978-5-4263-0712-4

© Бродовская Е. В., 2018

© МПГУ, 2018

Содержание

Введение	6
Глава I	8
Конец ознакомительного фрагмента.	11

Е. В. Бродовская, А. Ю. Домбровская

Большие данные в исследовании политических процессов Учебное пособие

Введение

Появление в конце XX в. интернет-коммуникации и ее интенсивное развитие оказало колоссальное воздействие на все сферы жизнедеятельности общества. С одной стороны, киберпространство предложило новый способ социального поведения людей, в том числе гражданского и политического участия. С другой – интернет-технологии обеспечили возможность аккумулирования цифровых «следов» поведения пользователей. Эти данные, идущие от поисковых систем, социальных медиа, мессенджеров, сенсорных устройств, огромные по объему и собираемые специальными машинами, называют большими данными – Big Data. Большие данные обладают рядом особых свойств:

- большим объемом, который несовместим с хранением на персональном компьютере;
- большой скоростью производства;
- содержательным многообразием;
- хаотичностью, неструктурированностью;
- гибкостью (совместимостью, подстраиваемостью под различные базы данных);
- корреляционностью.

Вместе с тем величина больших данных – это не главное в определении этого понятия. Основное – это природа их формирования. Большие данные создаются без исследовательского участия, это неспровоцированная информация, это автоматически сгенерированные данные о фактах социального поведения, которые невозможно обработать имеющимися сегодня пакетами для статистической обработки данных.

Большие данные впервые в истории прикладного политического анализа поставили вопрос о том, что основная проблема – не сбор информации, так как большие данные – это уже аккумулированный огромный массив постоянно пополняемой информации, вопрос заключается в способах анализа неструктурированных данных.

Вполне очевидно, что Big Data, этот глобальный массив социальной информации, представляет собой объект исследования большого числа научных дисциплин. Интегральной областью, изучающей взаимодействие информационных, политических, социальных и духовных процессов, является *социальный компьютеринг*. Его основной метод – киберметрический анализ, который позволяет аккумулировать большие данные и анализировать репрезентированные в интернет-контенте цифровые маркеры политических процессов.

Для современного профессионала в сфере гуманитарного знания, и прежде всего политолога, реализующего интернет-технологии в политике и управлении (социально-медийную аналитику, технологии создания и продвижения онлайн-сетевых политических и коммерческих брендов и т. д.), весьма важной компетенцией служит умение применить техники киберметрии для оценки состояния и выявления тенденций в развитии политических процессов по цифровым маркерам. Немаловажно умение применить в научном исследовании методы киберметрии для аспирантов общественных и гуманитарных направлений подготовки, в том числе по научной специальности «Политология».

Учебное пособие имеет цель дать общее представление о понятии, эвристических возможностях метода больших данных в прикладном анализе политических процессов, а также сформировать навыки применения киберметрии в изучении цифровых маркеров политиче-

ских изменений. Учебные кейсы, показанные в издании, направлены на формирование компетенций, связанных со способностью осуществлять графическую вербальную интерпретацию больших данных, автоматически генерируемых специальными инструментами анализа интернет-контента. Список контрольных вопросов для самопроверки, рекомендуемая литература по теме и приложения помогут студентам освоить методы применения больших данных в исследовании политических процессов.

Рекомендуется магистрантам направления «Интернет-технологии в политике и управлении» и аспирантам научной специальности «Политология», преподавателям высшей школы, читающим предметы, связанные с оценкой роли Интернета в современных реалиях, ученым, исследующим проблемы интернет-коммуникации с применением киберметрического анализа.

Глава I

Теоретико-методологические основы исследований с применением больших данных в системе методов социального компьютинга

Термин «большие данные» (Big Data) используется в двух смыслах. Это гигантские массивы информации – цифровые следы, аккумулируемые специальными инструментами из различных источников: социальных медиа, мессенджеров, сенсорных устройств и т. д. Большие данные в этом смысле характеризуют глобальный объем, неструктурированность, неспровоцированность создания. Дефиниция «большие данные» также может обозначать метод анализа гигантского по объему неструктурированного массива информации. В этой ипостаси большие данные отличает высокий эвристический потенциал оценки состояния и динамики массового сознания, гражданского и политического участия, жизненных стратегий пользователей, а также установления значимых корреляций для прогнозирования социальных и политических процессов по цифровым маркерам. Анализ больших данных осуществляется с помощью методов социального компьютинга. Поэтому важным представляется вначале рассмотреть понятие и сущность данного направления современных исследований и определить место автоматизированного анализа больших данных в системе методов социального компьютинга.

При определении понятия социального компьютинга также необходимо учитывать существование двух его проявлений – как области научных исследований и как сферы компьютерных технологий.

Приоритетная задача настоящего пособия состоит в анализе первого значения понятия. В этом случае социальный компьютинг понимается как область исследований, основанная на междисциплинарности и полипарадигмальности подхода к изучению взаимосвязи информационных, коммуникационных и социальных процессов, которая формируется в условиях функционирования и развития глобальной электронной сети.

Термин “Social Computing” дословно переводится как социальные «вычисления». Другими словами, это сплав социальной и компьютерной науки, сформировавший новую парадигму междисциплинарных исследований. Эта парадигма основана на позитивистском восприятии цифровых следов социальных связей индивидов как маркеров социальных и политических процессов.

Наиболее понятное и адекватное понимание социального компьютинга предложено А. А. Давыдовым [28], [29]. В его публикациях социальный компьютинг характеризуется как общий термин для обозначения группы технологий на стыке исследований и приложений в областях социального поведения и вычислительных систем или как совокупность социотехнических средств для анализа и использования человеческих связей как части процесса анализа больших данных.

Вместе с тем в специальной литературе встречается сужение определения социального компьютинга до использования вычислительных методов с целью изучения социального поведения. Данная трактовка рассматриваемого понятия не учитывает такой важный элемент социального компьютинга, как наличие взаимосвязи людей посредством Интернета, электронных социальных сетей.

Заслуживает внимания опыт определения термина «социальный компьютинг» на основе облака тегов, составленного по материалам докладов конференции “Social Computing, Behavioral-Cultural Modeling and Prediction”. В фокусе семантического ядра оказались понятия сети, социального, моделей и моделирования, динамики, анализа общественного мнения, дву-

сторонней связи, реального и виртуального миров, симуляции, взаимодействия, информальных групп, социальных графов, анализа больших данных и проч.

Анализ облака тегов позволил выявить семантическое пространство «социального компьютинга» и редуцировать его к такому семантическому ряду: «моделирование социального взаимодействия общностей и групп с использованием методов анализа больших данных, репрезентированных в интернет-контенте».

Исходя из этого, многочисленные определения социального компьютинга сводятся к указанию на значимость синтеза точного и социально-гуманитарного знания в процессе исследования влияния интернет-коммуникации на сознание и поведение людей.

Сегодня в мировом научном сообществе формируется и развивается ряд междисциплинарных направлений, близких к области исследований социального компьютинга.

Социология Интернета/социология киберпространства (Sociology of the Internet/Sociology of cyberspace) – отрасль социологии, объектом которой является исследование функционирующего в сети пространства социальных связей, а предметом анализа – социальные процессы и отношения в киберпространстве.

Цифровая социология (Digital sociology) – еще одна отрасль социологической науки, ориентированная на изучение цифровых медиа и их роли в повседневной жизни пользователей, социальных эффектов новых медиа, их воздействие на реальные социальные связи и процессы. Объект изучения цифровой социологии значительно шире, чем объект социологии Интернета, социологии онлайн-сообществ, социологии социальных медиа и т. п., и включает не только интернет-пространство как особый социокультурный и технологический феномен, но и те социальные среды, которые создают другие телекоммуникационные технологии. В предметное поле цифровой социологии также входят Интернет вещей и технологии работы с большими данными, которые, в свою очередь, формируют цифровую среду (digital space). Цифровая социология в этой связи может рассматриваться в самом широком понимании как сфера актуализации возможностей, которые дают цифровые инструменты для переосмысления структуры социологического знания. В этом смысле социальный компьютинг может рассматриваться как возможность «переформатирования» данных прикладных отраслей гуманитарных наук.

Культуральные «вычисления» (Cultural Computing) – направление исследований, сфокусированных на изучении культурных эффектов развития телекоммуникационных систем, выявлении цифровых маркеров культурных процессов, происходящих в виртуальном пространстве, в целом, воздействии интернет-коммуникации на развитие сферы культуры и искусства, *Когнитивные «вычисления» (Cognitive Computing)* – направление исследований, в которых применяются технологии оперативной обработки больших данных, структурирования хаотичных потоков глобальных по объему массивов информации.

Эмоциональные «вычисления» (Affective Computing) – междисциплинарная область, направленная на разработку интеллектуальных систем для обнаружения, распознавания, интерпретации и моделирования человеческих эмоций. Наибольшее применение эмоциональные «вычисления» находят в сфере развития виртуальной и дополненной реальности, цифровой медицины, игровой индустрии.

Контекстуальные «вычисления» (Context Computing) и иные виды автоматического и автоматизированного анализа больших данных, которые осуществляют программы искусственного социального интеллекта (Artificial Social Intelligence Agents). Контекстуальные «вычисления» ориентированы на создание систем, учитывающих социальный инвайронмент, ситуацию, в которой пребывает пользователь. Контекстуальные «вычисления» также неопределимы в разработке логики аккумулирования и анализа больших данных, позволяют учесть контексты формирования измеряемых информационных потоков.

Ключевые исследовательские задачи исследований в сфере социального компьютинга и смежных направлений:

- разработка методологических, методических и организационных основ изучения интернет-пространства и т. д.;
- исследование внутренних механизмов строения и развития социальных структур виртуального социального пространства;
- анализ процессов формирования и развития виртуальных сообществ;
- изучение особенностей социальных действий и поведения людей в условиях интернет-пространства;
- изучение влияния виртуальной среды (online) на сообщества, структуры и процессы, имеющие место в «реальной» жизни (offline), изучение феномена перехода виртуальных структур, сообществ, событий и т. д. в «реальное» пространство (так называемая девиртуализация);
- разработка методов автоматизированного анализа интернет-контента;
- создание методов автоматизированного структурирования информационных потоков в социальных медиа;
- формирование методов создания компьютерных корпусов основных языков России и постсоветского пространства;
- разработка методов создания базы цифровых маркеров социальной напряженности на основных языках народов России и постсоветского пространства;
- формирование методов создания базы цифровых маркеров межнациональных и межрелигиозных конфликтов на основных языках народов России и постсоветского пространства;
- разработка методов создания базы цифровых маркеров не-конвенциональных массовых; политических протестов на основных языках народов России и постсоветского пространства;
- формирование методов создания базы цифровых маркеров распространения экстремизма и терроризма на основных языках народов России и постсоветского пространства;
- разработка методов создания базы цифровых маркеров фальсификации отечественной и мировой истории на основных языках народов России и постсоветского пространства;
- формирование методов создания базы цифровых маркеров фейковой новостной информации на основных языках народов России и постсоветского пространства;

Конец ознакомительного фрагмента.

Текст предоставлен ООО «ЛитРес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на ЛитРес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.