

## ИССЛЕДОВАНИЯ СОДЕРЖАНИЯ ИНТЕРНЕТ-КОММУНИКАЦИИ: ИДЕИ И ВОЗМОЖНОСТИ

**Н. В. ЕФИМОВА, кандидат философских наук, доцент**  
Белорусский государственный университет  
г. Минск, Республика Беларусь

Internet studies – это область исследований, выросшая на почве социальных наук. В данном случае изучаются не технические характеристики интернета а взаимоотношения людей и коммуникация, которая осуществляется в интернете или посредством интернета, а также влияние интернета на процессы, внешние по отношению к нему, например, на образование.

Социальные проблемы, возникающие в процессе пользования интернетом (включая неравенство в доступе, создание цифровых барьеров, последствия для социальных отношений, сообществ, участия, идентичности и популярной культуры) не сразу привлекли внимание ученых. Б. Уэллман выделяет и описывает следующие периоды Internet studies:

**1978 – 1998 год.** В данный период фокус в основном был сосредоточен на технических возможностях интернета. Его перспективы видели в создании отдельных приложений для узких групп пользователей. Позже в рамках этого же периода Интернет стал восприниматься как такое изобретение, которое положило начало принципиально новой эпохе в жизни человечества. Но при этом ученые полностью отделяли онлайн-взаимодействие от офлайн-взаимодействия, игнорируя тот факт, что они неразрывно связаны друг с другом. Социальные аспекты Интернета практически игнорировались.

**1998 – 2009 годы.** Новые коммуникационные технологии стали стремительно проникать в жизнь людей, и, в связи с этим сообщество ученых перестало воспринимать интернет как изолированный феномен, признав его неразрывную связь с повседневной практикой. Акцент был сделан на изучении пользователей интернета: на их социально-демографических и социально-психологических характеристиках, а также на интересующей пользователей тематике интернет-публикаций.

**С 2009 года и по сей день продолжается третий период.** Его фокус внимания сосредоточен на междисциплинарном подходе к исследованиям интернета, адекватном его многоаспектной роли в обществе. Интернет проникает все глубже в жизнь современного общества, развиваются беспроводные технологии и устройства, университеты по всему миру создают уникальные дисциплины, направленные на изучение самых различных сторон Интернета. В целом сообщество ученых старается исследовать каждую составляющую феномена интернета, как часть целого, а сфера Internet Studies становится все более популярной как в научной среде, так и за ее пределами. В рамках данной сферы ученые из множества смежных дисциплин объединились для изучения и прогноза *социальных и культурных перспектив* использования Интернета в современном обществе [1, с. 123-129].

Можно выделить и охарактеризовать два вида исследований: интернет-центрированное и общество-центрированное. В первом случае анализируется содержание интернета как таковое, например, тональность описания события в публикациях онлайн-медиа или тематическая направленность сайта. Во втором случае ставится цель через анализ содержания Интернета описать или спрогнозировать события, процессы, происходящие или ожидаемые в офлайн-среде. Так, например, через анализ поисковых запросов пользователей можно узнать их предпочтения по отношению к товарам конкретной категории или брендам и предсказать потребительский спрос.

Преимуществом Internet Studies является то, что существует возможность найти и изучить написанное и опубликованное «респондентами» в интернете, не прибегая к прямой коммуникации с ними, не требуя эксплицитного осознания субъектами их роли или самого факта участия в исследовании («нереактивные методы онлайн исследований»). И. Ф. Девятко делит все нереактивные методы сбора интернет-данных: на простое наблюдение и анализ цифровых отпечатков и готовых данных [2, с. 17].

Простое наблюдение означает скрытое (без манифестации присутствия исследователя) отслеживание сетевого поведения индивидов и групп пользователей в сообществах, группах, на форумах и иных площадках для коммуникации в интернете. Анализ цифровых отпечатков и готовых данных включает анализ цифровых следов интернет-активности и изучение онлайн-документов, архивов и записей.

«Цифровые следы» интерпретируются как результат манипуляций пользователей с цифровым кодом сайтов и браузеров, это могут быть запросы в поиске, переходы по ссылкам, лайки и репосты в социальных сетях и тому подобное. Анализ онлайн-документов и архивов подразумевает исследование текстов. Это онлайн-дневники, блоги, комментарии, а также публикации в новых медиа, регулярные электронные отчёты компаний.

Таким образом, нереактивные методы позволяют раскрывать поведение пользователей в сети, результаты их сетевой активности, а также анализировать интернет как медийную реальность, исследуя циркулирующие в этой реальности тексты, их содержание. Содержательные данные интернета могут быть изучены с применением целого ряда методик, как ручных, так и автоматизированных (первые предусматривают активность исследователя на всех стадиях процесса сбора и обработки данных, вторые либо полностью заменяют исследователя программой, либо оставляют за ним сбор массива для анализа и(или) интерпретацию результатов).

Автоматизированными методами анализа поведения пользователей и их «следов» являются сервисы типа Google Analytics, Яндекс.Метрика, и др. С их помощью можно анализировать посещаемость сайтов, «юзабилити» (удобство пользования сайтом), поведение пользователей на сайте, параметры электронной торговли, и некоторые другие. Автоматизированные методы анализа текстов: тематическое моделирование, классификация, кластеризация и сентимент-анализ. Например, метод тематического моделирования основан на «обучении» программы с помощью ключевых слов «видеть» тематику сообщений. Исследователь собирает материал (статьи, комментарии, посты) и определяет его темы. Далее исследователь обрабатывает несколько текстов вручную и с их помощью закладывает в программу алгоритм распознавания тематики сообщений. Программа связывает выбор темы исследователем с ключевыми словами в обработанных им текстах и запускает механизм обработки оставшегося массива текстов по аналогии. На данный момент существует две программы с указанными функциями: MALLET и TopicMiner. Существуют и упрощённые коммерческие аналоги.

Таким образом, метод тематического моделирования позволяет установить тематическую направленность определённого ресурса сети и тематический уклон публичной активности субъектов сети.

Сентимент-анализ используют для изучения мнений аудитории о каком-либо объекте, например продукте, бренде, событии. Его сущность заключается в маркировке сообщений конкретной тематики с точки зрения оценки, по таким категориям: положительное / отрицательное / нейтральное, негативное / позитивное, согласие / несогласие и т.п. Наиболее распространённым примером использования данного метода является обработка отзывов аудитории. Выборка для анализа отзывов формируется из обзоров и постов блогеров и обычных

пользователей на страницах и в социальных сетях, из онлайн-дневников, комментариев, отзывов на специализированных сайтах.

Таким образом, сентимент-анализ как метод исследования содержания интернета позволяет проанализировать отношение пользователей к определённым новостным темам, событиям, субъектам, компаниям или продуктам. Ручные методы анализа интернет-текстов не являются уникальными. Это давно известные контент-анализ, интент-анализ, дискурс-анализ и некоторые другие специальные методы анализа текстов. Особенностью применения этих методов при анализе интернет-среды являются единицы выборки: в сети интернет ими будут публикации сетевых СМИ, блоги, посты в социальных сетях, электронные рассылки, комментарии, отзывы, содержание страниц сайтов. Таким образом, ручные методы анализа текстового содержания в интернете отличаются от методов анализа печатных текстов параметрами создания выборки.

Техника контент-анализа заключается в представлении содержания текстов в численном выражении в соответствии с конкретными категориями и последующем анализе полученных числовых данных. Процедура контент-анализа начинается с установления системы категорий, или категориальной схемы. Далее необходимо идентифицировать в тексте единицы анализа и сопоставить их с категориями. Традиционными для контент-анализа СМИ единицами анализа являются тема, оценка, суждение, персонаж. На сегодняшний день этот метод имеет машинный аналог (например, программы ВААЛ и Vaal Toolbox для русскоязычных текстов), однако часто контент-анализ делается и вручную. Данный метод позволяет выявлять, степень внимания к темам, персонажам, общественные настроения аудитории и многие другие социальные явления.

Интент-анализ направлен на выявление интенции, т.е. намерения, стремления субъекта в созданном им тексте. Примером использования интент-анализа содержания интернета является исследование комментариев на новостном или ином портале [3, с. 15].

Очевидным отличием анализа текстов в интернете от анализа материалов несетевых ресурсов являются единицы отбора. Содержание Интернета включает в себя специализированный контент сайтов (коммерческих и некоммерческих), публикации сетевых СМИ, комментарии, отзывы и отклики, блоги, онлайн-дневники, электронные рассылки, электронные отчёты.

Кроме единиц отбора модифицируются также и категории анализа. К традиционным категориям добавляются такие как количество просмотров, пересылок, обсуждаемость публикаций или постов, тональность комментариев и отзывов и некоторые другие.

Таким образом, анализ содержания интернета создаёт новое поле исследовательской деятельности за счёт изменения объектов исследования, единиц анализа и инструментария.

#### **Список использованных источников**

1. Wellman, B. The three ages of internet studies: ten, five and zero years ago / B. Wellman // *New Media & Society*. – 2014 – Т. 6, № 1.
2. Девятко И. Ф. Инструментарий онлайн-исследований: попытка каталогизации // *Онлайн исследования в России 3.0/ Под редакцией Шашкина А. В., Девятко И. Ф., Давыдова С. Г.* – М.: Издательский дом «Кодекс», 2012. – С. 17–31
3. Радина, Н. К. Интент-анализ онлайн-дискуссий (на примере комментирования материалов интернет-портала ИноСМИ.ru) / Н. К. Радина // *Медиаскоп [Электронный ресурс]*. — 2016. Режим доступа: <http://www.mediascope.ru/2238>. Дата доступа: 26.05.2021.