

Е.В. Щекотин

ЦИФРОВЫЕ СЛЕДЫ КАК НОВЫЙ ИСТОЧНИК ДАННЫХ О КАЧЕСТВЕ ЖИЗНИ И БЛАГОПОЛУЧИИ: ОБЗОР СОВРЕМЕННЫХ ТЕНДЕНЦИЙ

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-011-00391.

Обсуждается возможность использования нового источника данных – цифровых следов пользователей для изучения качества жизни и благополучия населения. Представлен обзор актуального состояния исследований качества жизни и благополучия населения, которые используют в качестве источника данных цифровые следы (социальные медиа, поисковые запросы, сервис Google Books и т.д.) и методы цифровой социологии (большие данные, машинное обучение) для анализа и обработки этих данных.

Ключевые слова: цифровые следы; качество жизни; благополучие; большие данные; машинное обучение; цифровая социология.

Введение

Развитие информационно-коммуникационных технологий оказывает огромное влияние и на исследования в области качества жизни и благополучия. Это влияние носит двойственный характер: с одной стороны, информационно-коммуникационные технологии изменяют сам образ жизни людей, генерируют массу новых возможностей для человека, которые меняют жизнь на уровне повседневно-бытовой активности, помогают преодолеть определенные ограничения, связанные со здоровьем, и т.п., тем самым способствуя повышению качества жизни [1]; с другой стороны, эти технологии создают предпосылки для применения в социальных науках и практике государственного управления новых методологических инструментов для изучения качества жизни. Помимо опросного метода, который является основным в изучении субъективных аспектов качества жизни, сегодня активно используются данные о поведении пользователей в интернете, а также информация, которую они размещают в социальных сетях. Таким образом, для исследований качества жизни цифровые технологии важны не только с точки зрения изменения образа жизни людей, трансформации труда, социальных институтов и тому подобных процессов, получивших наименование цифровой трансформации, но и с точки зрения обновления методологии научного познания.

Как отмечает О.В. Крыштановская, «раньше мы изучали социальные группы путем физического соприкосновения: надо было подойти к человеку, задать ему вопросы, получить ответы, закодировать их, ввести коды в компьютер, создать матрицы, подвергнуть их статистическому анализу, декодировать и интерпретировать. С распространением социальных сетей появилась возможность проводить опросы онлайн, затрачивая на это меньше ресурсов, экономя время и собирая значительно большие объемы информации. Такого рода опросы подвергались критике со стороны профессионального сообщества в связи с тем, что невозможно было доказать репрезентативность таких выборок. Однако постепенно стали проступать контуры новой социологии – социологии, связанной с цифровой эпохой, с новыми методами сбора и обработки информации» [2. С. 5]. О.В. Крыштановская называет

цифровую социологию бесконтактной, ее особенность в том, что социологи никого не опрашивают, они получают информацию из потока сообщений, из открытых источников. «Можно не выполнять традиционные манипуляции, если люди в социальных сетях и блогах и так высказываются относительно всех политических событий. Можно не спрашивать их об этом, если они сами пишут свое мнение, не ожидая анкет социологов» [2. С. 6].

Можно с уверенностью утверждать, что социальные медиа стали сегодня одним из основных (если не основным) источников данных о чувствах, настроениях, переживаниях, мнениях и поведении людей. Обсуждая использование технологии больших данных для исследования благополучия, К. Беллет и П. Фрайтерс отмечают: «Утверждения о том, что социальные медиа могут, таким образом, предсказывать наше благополучие весьма точно, вообще не должно быть удивительным, поскольку это часто является смыслом социальных медиа» [3. Р. 102]. В конце обзора возможностей применения цифровых технологий для исследования благополучия, который вошел во Всемирный доклад о счастье 2019, авторы приходят к выводу, что «большие данные расширяют возможности исследователей, правительств, компаний и других субъектов измерять и прогнозировать благополучие и внутреннюю жизнь отдельных людей. Следует ожидать, что это повысит возможность анализировать воздействие политики и крупных изменений на благополучие в целом, что должно повысить интерес и знания о благополучии» [3. Р. 118].

Достоинства и недостатки цифровых следов как источника данных

Социальные медиа как источник информации обладают своими достоинствами и недостатками в сравнении с традиционными опросными методами. Так, С. Санчес, М. Краглиа и А. Брегт рассматривают Twitter как источник данных о различных социальных индикаторах и предлагают использовать эти данные для измерения качества жизни [4]. В своей работе С. Санчес, М. Краглиа и А. Брегт исследуют только один социальный индикатор – коммуникацию с политиками. Данное исследование представляет значи-

тельный интерес, так как авторы подробно проанализировали возможности и ограничения предлагаемого метода. Достоинствами метода измерения качества жизни по данным Twitter, по мнению авторов, является более высокая степень детализации распределения оценок пользователей во времени и пространстве, т.е. полученная информация является более подробной и точной. Эти данные доступны в любой момент – их данные можно получать в реальном времени. Также достоинствами этого метода являются быстрота реализации проекта исследования, относительно невысокая стоимость (за счет экономии на расходах на проведение опросов) и гибкость методологии – такой метод позволяет легко подстраиваться под поставленные задачи, в то время как опросы включают в себя фиксированные вопросы, что облегчает сравнение результатов и повторные исследования, но в случае изменения целей анализа не позволяет гибко реагировать на актуальные запросы. Главным недостатком Twitter как источника данных являются проблемы, связанные с репрезентативностью выборки, – пользователи Twitter не представляют всего населения, более активными пользователями являются молодые граждане. Также есть некоторые сложности с определением местоположения пользователей, что отражается на точности измерения. Авторы исследования приходят к выводу, что данные из Twitter вполне могут быть использованы для измерения социальных индикаторов, и если пока еще не удастся нивелировать отмеченные выше ограничения, то совершенно точно эти данные могут быть использованы в качестве дополнения к результатам официальных опросов. Особенно это актуально для малых поселений, где официальные опросы не дают репрезентативных оценок населения.

М. Шобер с соавторами высказывают более осторожную точку зрения относительно возможностей анализа социальных сетей в качестве инструмента, который может заменить опросы и официальную статистику [5]. С одной стороны, очевидно, что анализ социальных сетей обладает огромным потенциалом для быстрого и доступного измерения не только тех аспектов, которые замеряют опросы, но и тех социальных явлений, которые никогда ранее не оценивались. В то же время возникает много вопросов о том, в какой мере анализ социальных сетей будет соответствовать результатам опросов, т.е. в какой мере он дает надежные данные. Решение этих вопросов далеко от окончательных выводов. Основная проблема, по мнению авторов, заключается в том, что исследователи, занимающиеся анализом социальных сетей и опросами, имеют различный опыт, знания и языки описания. Так, ученые в области компьютерных технологий и инженеры используют широкий спектр методов для анализа данных в социальных сетях (такие как машинное обучение, анализ естественного языка, лексический анализ), социальные исследователи ориентируются на методы социологических опросов («золотой стандарт»). Сюда же подключаются исследователи из области коммуникации и медиа, рынка и потребительского поведения и т.д. Все это

крайне усложняет ситуацию и нуждается в дальнейших обсуждениях и исследованиях.

В.И. Дудина и Д.И. Юдина отмечают, что «хотя измерение общественного мнения через анализ информации из социальных сетей, блогов и других социальных медиа постепенно становится все более распространенной практикой полстерских служб и компаний, методологических работ, описывающих не просто процесс сбора и анализа такой информации, но и сравнение полученных данных с результатами опросных методов, на данный момент очень мало» [6. С. 65]. Можно выделить три основных подхода к решению задачи извлечения мнения из неструктурированных текстов, которые представляют собой посты в социальных сетях, блогах и т.д.: это машинное обучение (тематическое моделирование, методы обучения с учителем), лингвистические методы (словари, лингвистические парсеры) и гибридный подход, соединяющий два указанных подхода [7]. В своем исследовании В.И. Дудина и Д.И. Юдина сравнивали реакции на резонансное событие общественного мнения (демонстрация фильма «Чайка» на видеохостинге Youtube в 2015 г.), полученные двумя различными способами – с помощью традиционного опроса общественного мнения и тематического моделирования мнений, извлеченных из интернета.

Вывод, который делают авторы данного исследования, состоит в том, что анализ текстов из сети «Интернет» наделен как определенными достоинствами, так и недостатками. К преимуществам метода можно отнести то, что мнения и обсуждения в интернете позволяют получить больше информации о контексте формирования общественного мнения относительно какого-либо события или персоналии. Также результаты анализа мнений в интернете могут послужить для усовершенствования опросов (структуры опросника, вопросов). В то же время авторы заключают, что «анализ текстов из Интернета не может рассматриваться как полноценная альтернатива массовым опросам» [6. С. 75]. Основная проблема, связанная с экстраполяцией результатов изучения онлайн реальности на оффлайн-реальность, заключается в отсутствии теоретических оснований для генерализации данных на более широкие («нецифровые») слои населения. Как отмечают В.И. Дудина и Д.И. Юдина, «традиционная модель массового опроса предполагает привязку мнений к социально-демографическим группам, а результаты анализа текстов из сети Интернет дают возможность представить лишь спектр обсуждаемых тематик и их относительную популярность, но не позволяет сопоставить мнения с их носителями, поскольку при анализе данных из социальных медиа возникает проблема с получением надежной демографической информации» [6. С. 76]. Данные об онлайн активности пользователей социальных сетей, которые проще и быстрее получить, чем результаты опросов, пока не вписываются в существующую модель общественного мнения.

Преодоление этого ограничения В.И. Дудина и Д.И. Юдина видят в совершенствовании инструментов анализа онлайн-текстов, и, в частности, в разработке специфически социологических анализаторов,

которые будут способны структурно воспроизводить традиционные опросы. Такая точка зрения хорошо иллюстрирует основную причину, препятствующую более активному использованию цифровых технологий в социологических исследованиях, – это интуитивное стремление социологов свести новые методы изучения социальной реальности к привычным и хорошо обоснованным стандартам, которые гарантируются тщательным следованием предписанным традиционной статистикой процедурам. Новые технологии требуют от исследователя более свободного и творческого взгляда, полета «социологического воображения», что вполне успешно реализуется в других дисциплинах (например, в математике и физике) и сферах знания (например, в бизнесе).

Вместе с тем, несмотря на отмеченные методологические ограничения, исследование различных социальных процессов с помощью анализа социальных сетей становится все более популярным. Так, Э. Коуди, А. Риган, П. Доддс и С. Данфорт утверждают, что изучение общественного мнения при помощи анализа текстов в интернете (в данном случае это Twitter) обладает такой же предикативной точностью, как и традиционные оффлайн-опросы. Они полагают, что распространение социальных медиа создают предпосылки для проведения опросов общественного мнения, в которых не требуется приглашать респондентов к участию [8]. То есть исследователь получает возможность изучать мнение людей, которое они высказывают добровольно, по собственному желанию, без посредничества интервьюеров и использования заранее сконструированных опросных листов.

Также, помимо упомянутой выше статьи С. Санчес, М. Краглия и А. Брегг, можно отметить работу Д. Антенуччи с соавторами [9], в которой исследуется проблема безработицы на материале Twitter. Большое значение методу изучения онлайн-активности в социальных сетях придается в исследованиях, посвященных изучению актуальных политических процессов (прежде всего, прогнозирование результатов голосования, но не только), и маркетинге. В первом случае любимым объектом изучения для исследователей является мессенджер Twitter (не в последнюю очередь это связано с бесплатным доступом к открытым данным в Twitter, что позволяет исследователям свободно собирать интересующие их данные). Во втором случае изучаются возможности анализа социальных сетей для достижения маркетинговых целей, продвижения продуктов и брендов и т.д.

По нашему мнению, глубокий количественный анализ онлайн-активности, т.е. различных действий людей в социальных сетях, позволяет получить информацию об их эмоциях и переживаниях и оценить субъективные аспекты качества жизни и благополучия людей. Информация об онлайн-активности пользователей получила название «цифровой след». Как отмечают Дж. Блюб Дж. Конделл и Т. Ланни, в современном мире человек наделен двумя идентичностями: первая, «реальная», формируется на основе официальных документов, вторая, «цифровая» включает в себя всю совокупность форм взаимодействия человека с интернетом (историю поисковых запросов,

онлайн-сервисы, форумы, блоги, социальные сети) [10]. Цифровой след отображает присутствие человека в интернете и служит доказательством его цифровой и реальной идентичности. Как заметил М. Конлин, описывая сущность саморепрезентации в интернете, «ты есть то, что ты постишь» [11].

Цифровые следы подразделяются на два типа – активные и пассивные [12]. Пассивные следы пользователь оставляет, когда посещает ресурсы в интернете, просматривает и загружает какую-либо информацию, взаимодействует с мобильной телефонной сетью (определение местоположения пользователя) и т.д. Пассивные цифровые следы собирают без ведома пользователей владельцы интернет-ресурсов, приложений и платформ. В англоязычной литературе пассивный цифровой след называют «выхлопными дымными» (data exhaust или exhaust data). Выхлопные данные являются частью более широкой категории нетрадиционных данных (unconventional data), которые включают в себя геопространственные, сетевые и временные ряды данных и могут быть полезны для прогнозной аналитики. Каждый посещенный сайт, нажатая ссылка и даже наведение курсора мыши фиксируются, оставляя за собой след данных, создается огромное количество часто необработанных данных, которые могут храниться в разных формах – в виде файлов cookie, временных файлов, лог-файлов, сохраненных выборов и многого другого.

Активным цифровым следом является та информация, которая создается и размещается в интернете самим человеком. Онлайн-аккаунты предоставляют множество проверенных ссылок на характеристики реальных личностей. Почти каждый созданный онлайн-аккаунт требует наличия адреса электронной почты. Официальные онлайн-сервисы требуют персональных идентификаторов, таких как имя, дата рождения, адрес и уникальные персональные идентификационные номера. Интернет-магазины требуют почтового адреса и платежных реквизитов, поиск в интернете и использование карт Google часто включают в себя использование текущего местоположения человека и потенциально того, где он будет находиться в будущем, а учетные записи в социальных сетях представляют собой подтверждение контактов, отношений, а также профессиональные и личные интересы.

Экстериорная и интериорная стратегии изучения качества жизни и благополучия в контексте цифровизации

Современные цифровые технологии открывают широкие возможности для измерения и оценки качества жизни с использованием цифровых следов пользователей. Выше мы рассмотрели основные аргументы в дискуссии, которая сегодня ведется относительно применимости анализа цифровых следов для измерения качества жизни. Далее рассмотрим некоторые наиболее примечательные исследования, в которых реализуется методологический подход, связанный с использованием цифровых следов для измерения качества жизни и благополучия. Можно выделить две основные стратегии в том, что касается взаимодей-

ствия двух направлений исследований – информационно-коммуникативных технологий и исследований качества жизни и благополучия. Первая стратегия состоит в рассмотрении цифровизации как процесса проникновения информационно-коммуникативных технологий в жизнь общества и применительно к проблематике качества жизни соотносится с изучением возможностей, которые появляются в этой связи, для улучшения качества жизни людей. Эту стратегию можно назвать экстериторной, так как цифровые технологии здесь выступают в роли внешнего фактора, который позволяет реализовать некоторые позитивные (или негативные) изменения в жизни людей. Такая установка реализуется в исследованиях, которые рассматривают социальные медиа как новое социальное явление, которое оказывает воздействие на различные аспекты качества жизни и благополучия людей [13–26]. В этом случае социальные медиа выступают как существенный фактор, который может негативно или позитивно воздействовать на такие аспекты благополучия как удовлетворенность жизнью [27–29], социальный капитал [30–32], удовлетворенность браком [33], депрессия и страхи [34–39], одиночество [40, 41], душевное здоровье [42–44], социальная поддержка [45], оценка своей жизни по степени счастья в сравнении с другими людьми [46], продуктивность работы [47] и т.д.

Вторая стратегия отталкивается от понимания цифровизации как уже состоявшего процесса и цифровых технологий как составного компонента повседневного существования людей, поэтому ее можно охарактеризовать как интериорную. В рамках этой стратегии информация о поведении пользователей в онлайн-пространстве (цифровые следы) рассматривается как достоверный источник о качестве жизни и благополучии людей. Жизнь человека разворачивается одновременно в двух реальностях – в оффлайн (физическом мире) и в онлайн (информационном пространстве), и эти реальности взаимно детерминируют друг друга, плавно перетекая одна в другую. Поэтому существование человека одновременно в двух реальностях необходимо рассматривать как комплексный и целостный феномен, в котором невозможно без упрощения и искажения отделить оффлайн-жизнь от онлайн. Данная стратегия требует создания нового методологического инструментария, присущего цифровой социологии (технологии больших данных, машинное обучение, анализ социальных сетей и т.д.) для сбора и обработки информации о качестве жизни в информационном измерении.

Таким образом, интериорная стратегия исследований рассматривает цифровые следы как самостоятельный источник данных для оценки качества жизни и благополучия. Учитывая ту роль, которую социальные сети стали играть в последнее десятилетие в повседневной жизни людей, эти попытки представляются очень плодотворными. Так, Б. Хао с соавторами использовали технологию машинного обучения для того, чтобы предсказывать субъективное качество жизни пользователей социальных медиа. Авторы использовали данные в социальной сети Sina Weibo 1785 волонтеров для обучения алгоритма, предвари-

тельно попросив их заполнить опросники для оценки позитивных и негативных аффектов (PANAS) и оценки психологического благополучия (PWBS) [48]. Схожую работу, но на материалах Facebook, проделали Шварц и др. [49]. В другой работе Шварц с соавторами на основании данных Twitter дали оценку субъективного благополучия в разных графствах США [50]. Л. Чен с соавторами анализировали обновления статусов на страницах пользователей Facebook для предсказания субъективного благополучия пользователей [51]. К. Ву с соавторами использовали данные социальной сети Sina Weibo для построения City Happiness Index [52].

Н. Ванг с соавторами в течение года изучали профили пользователей Facebook и оценивали уровень их благополучия по шкале удовлетворенности жизнью Э. Диенера, затем и сравнили полученные данные с индексом Facebook's Gross National Happiness, который калькулируется на основе анализа количества позитивных и негативных слов, используемых в обновлениях статусов пользователей [53]. В данном исследовании авторы ставят под сомнение возможность применения лингвистического анализа интернет-сообщений для изучения психологического состояния пользователей. С. Янг и П. Сринивасан удовлетворенность жизнью населения исследуют на материале Twitter [54]. Сюда же можно отнести исследования, изучающие «географию счастья», например связь между эмоциями и мобильностью [55]. Интересные результаты были получены в нескольких исследованиях того, насколько положительные и негативные эмоции коррелируют с удовлетворенностью жизнью, измеренной на основе данных Facebook. Как оказалось, положительные эмоции слабо коррелируют с данными по удовлетворенности жизнью по Facebook, в то время как отрицательные эмоции демонстрируют такую связь. Отсюда можно предположить, что социальное давление может побуждать несчастливых людей притворяться счастливее, чем они есть на самом деле, что менее вероятно в случае проявления негативных эмоций.

Я. Алган с соавторами в качестве источника данных использовали запросы пользователей в поисковой системе Google, которые они анализируют с помощью сервиса Google Trends, за период 2008–2013 гг. [56]. В качестве поисковых запросов авторы использовали два перечня слов, которые связаны с субъективным благополучием. Первый список взят из онлайн-базы данных Better Life Index, в которой содержатся ответы на вопрос «что значит для вас лучшая жизнь?». Второй список основан на американском опросе использования времени, который фиксирует повседневную деятельность американских граждан, а также положительные или отрицательные эмоции, связанные с этими эпизодами. Авторы также добавили ряд слов, которые имеют отношение к различным жизненным переживаниям, связанным с субъективным благополучием, например, проблемы с работой (например, «безработица»), бедностью («купоны») или семейный стресс («приют для женщин»). В общей сложности исходная база данных содержит 827 слов, из которых 201 относится к материальным усло-

виям (доход, богатство, занятость и жилье), 529 – к качеству жизни (здоровье, досуг, образование, окружающая среда, гражданская жизнь, личная безопасность, субъективное благополучие и социальные связи) и 97 – к потенциально запретным категориям (порнография, ненависть и расизм, теории заговора).

Отдельные поисковые слова авторы объединили в 12 категорий, которые в свою очередь сгруппировали в три класса, отражающие важнейшие аспекта жизни: материальные условия (поиск работы, рынок труда, финансовая безопасность и домашние финансы), социальные (семейный стресс, семейное время, гражданская вовлеченность и личная безопасность) и здоровье и благополучие (состояние здоровья, здоровые

привычки, летняя активность, образование и идеалы). Авторы исключили категорию домашние финансы из модели, поскольку в этот период (2008–2013 гг.) преобладают слова, критически связанные с финансовым кризисом (например, ипотека), что делает важность этих слов для прогнозирования субъективного благополучия весьма специфичной во времени. В табл. 1 представлены ключевые слова, которые характеризуют каждую категорию. Так как компоненты категорий в оригинале изложены на английском языке не всегда есть возможность корректно перевести их в форму поискового запроса, поэтому некоторые запросы мы оставили на языке оригинала и дали пояснение в скобках, часть опустили по указанной выше причине.

Таблица 1

Компоненты категорий [56]

Аспекты жизни	Категория	Альфа Кронбаха	Ключевые слова
Материальные условия	Поиск работы	0,8720	Неполный рабочий день, увольнения, ярмарка вакансий, обучение, выходное пособие, уровень безработицы, ярмарка карьеры, работа, пособие по безработице
	Рынок работы	0,9315	Трудовые льготы, трудовой договор, карьерный центр, сопроводительное письмо, пенсия, сертификационная программа, выход на пенсию, опыт работы, начальный уровень, квалификация, дискриминация, льготы работникам
	Финансовая безопасность	0,9101	401k, банковские операции, семейный бюджет, жилищное управление, раздел 8, инфляция, студенческие кредиты, школьные кредиты, процентная ставка, увольнение, финансовый кризис, кредиты, eic (налоговый кредит на заработанных доход), социальное обеспечение, купоны, medicaid (медицинская помощь немущим), заявление о банкротстве, жилье, выселение, продовольственный банк, бездомные
	Домашние финансы	0,9618	Ипотечные ставки, ипотечный калькулятор, ипотечное рефинансирование, цена дома, ипотечный кредит, жилищный кризис, ипотечный платеж
Социальные аспекты	Семейный стресс	0,8818	Домашнее насилие, помощь в браке, проблемы брака, опека, брачные консультации, поддержка семьи, приют для женщин
	Семейное время	0,9140	Закуски для детей, выходные, adhd (синдром дефицита внимания с гиперактивностью), детский сад, детский центр, pta (родительский комитет), детская вечеринка, волейбол, играть в футбол, чтение, рецепт, домашняя работа, прачечная, малыш, детский душ, издевательства, детские книги, идеи для детей, обучение
	Гражданская вовлеченность	0,9148	Волонтерство, донорство крови, ассоциация домовладельцев, бойскауты, киванис, городской совет, мasons, бинго, чаепитие, клуб львов, клуб, собрание сообщества, гражданская активность, ротари-клуб, муниципальный совет
	Личная безопасность	0,8657	Огнестрельное оружие, жертвы, оружие, безопасность оружия, насильственная преступность, уровень преступности, нападение, камера безопасности, убийство, самооборона, агрессия, риски, домашняя сигнализация, ограбление
Здоровье и благополучие	Состояние здоровья	0,9051	Стресс, гипертония, диабет, ожирение, паническое расстройство, болезнь, употребление табака, рак легких, сердечные заболевания, обсессивно-компульсивные заболевания, рак, методы релаксации, антидепрессанты, состояние здоровья, переломы, артрит, астма, релаксация, уход за собой, проблемы со сном, клиника майо, симптомы депрессии, проверка симптомов, частота самоубийств, употребление наркотиков, хроническая усталость
	Здоровые привычки	0,8441	Физические упражнения, вес, здоровое питание, уход за зубами, фрукты и овощи, ожидаемая продолжительность жизни, похудеть, здравоохранение, бросить курить
	Летняя активность	0,9417	Гольф, рыбалка, мотоцикл, пруды, пешие прогулки, езда на велосипеде, водные виды спорта, катание на лодках, туры, пляжный коттедж, достопримечательности, ночлег и завтрак, бейсбол, софтбол, играть в бильярд
	Образование и идеалы	0,9862	Средняя школа, экономика, домашнее задание, государственный университет, мораль, индивидуализм, билль о правах, человеческий капитал, конституция, политические действия, социальная справедливость, отсев, чтение книг, грамматика, репетитор, буддизм, школа, этика, философия, учеба, изучение, математика, обучение, навыки, среднее образование, письмо, почитание, творчество, психолог, терапия, самоуважение, нравственность, бедность, онлайн-курсы, литература, ученая степень, язык, наука, грамотность, феминистка, права, отношения, свобода слова, гражданские права, религия, религиозные права, имущественные права, расизм, правительства, свобода высказывания, ритуалы, детская смертность, духовность

Используя методы машинного обучения, Я. Алган с соавторами построили модели предсказания для оценки благополучия населения как США в целом, так и на уровне отдельных штатов. Для установления взаимосвязи между данными о благополучии населения, полученными с помощью цифровых методов, с данными, полученными традиционными методами,

авторы использовали результаты ежедневных телефонных опросов, которые проводила компания Gallup Analytics на выборке не менее 500 американцев в возрасте старше 18 лет с 6 января 2008 г. по 4 января 2014 г. Опрос включает в себя 6 показателей самооценки положительных эмоций (счастье, учеба, оценка жизни в настоящий момент и через 5 лет, смех,

уважение), а также 4 показателя отрицательных эмоций (гнев, печаль, стресс, беспокойство), Я. Алган с соавторами отмечают интересное наблюдение – последствия «Великой рецессии» 2008 г. видны по большинству субъективных показателей благополучия: оценка жизни в настоящий момент и через 5 лет, счастье и смех значительно снизились в 2008–2009 гг., в то время как процент людей, испытывающих беспокойство, гнев, стресс и печаль, одновременно возрос. Второе наблюдение касается цикличности этих переменных, которые демонстрируют большие сезонные колебания.

В результате этого интереснейшего исследования Я. Алган с соавторами делают несколько выводов. Ряд категорий поисковых запросов имеет очень высокую предсказательную точность – это поиск работы,

гражданская активность и здоровые привычки. Из всех показателей субъективного благополучия, которые используются в опросах, за исключением стресса, материальные условия являются наиболее важным семейством предикторов, за которыми следуют социальные факторы и категории здоровья/хорошего самочувствия. На уровне категорий наиболее важными переменными являются поиск работы, финансовое обеспечение, летний отдых и семейная жизнь.

К. Беллет и П. Фрайтерс в своем обзорном исследовании обращаются к изучению благополучия методами больших данных. Результаты сравнительного анализа данных о благополучии, полученных цифровыми и традиционными опросными методами, представлены в табл. 2. Для определения наличия связи авторы используют коэффициент детерминации R-квадрат.

Т а б л и ц а 2

Предсказание уровня благополучия методами Больших данных [3. С. 104]

Источник	Показатель субъективного благополучия	Показатель, измеряемый методами Больших данных	Источник Больших данных	Единица анализа	Коэффициент детерминации R ²
Предсказание на основе данных индивидуального уровня					
Collins S. et al. [57]	Удовлетворенность жизнью	Обновления статуса	Facebook	Facebook users	0,02
Kosinski M. et al. [58]	Удовлетворенность жизнью	Тип понравившихся страниц Facebook	Facebook	Facebook users	0,028
Liu P. et al. [59]	Удовлетворенность жизнью	Обновления статуса (положительные эмоции)	Facebook	Facebook users	0,03
Liu P. et al. [59]	Удовлетворенность жизнью	Обновления статуса (отрицательные эмоции)	Facebook	Facebook users	0,026
Schwartz H.A. et al. [49]	Удовлетворенность жизнью	Обновления статуса (темы, лексика)	Facebook	Facebook users	0,09
Предсказание на основе агрегированных данных					
Algan Y. et al. [60]	Удовлетворение жизнью	Поисковые запросы	Google Trends	Еженедельные наборы данных по США	0,760
Algan Y. et al. [60]	Счастье	Поисковые запросы	Google Trends	Еженедельные наборы данных по США	0,328
Collins S. et al. [57]	Удовлетворение жизнью	Средний размер персональной сети	Facebook	LS bins	0,7
Collins S. et al. [57]	Удовлетворение жизнью	Среднее количество обновлений статусов	Facebook	LS bins	0,096
Collins S. et al. [57]	Удовлетворение жизнью	Среднее количество тегов для фотографий	Facebook	LS bins	0,348
Hills I. et al. [61]	Удовлетворение жизнью	Слова	Google Books	Ежегодная панель из 5 стран	0,25
Schwartz H.A. et al. [49]	Удовлетворение жизнью	Темы и лексика из твитов	Twitter	Округа США	0,094

К. Беллет и П. Фрайтерс выделяют два типа данных для предсказания уровня благополучия – данные индивидуального уровня (обновления статуса в аккаунте пользователя, лайки, которые поставил пользователь, и т.д.) и агрегированные данные, т.е. не привязанные к отдельному пользователю, данные, усредненные по какой-то группе (поисковые запросы, средние значения сети, тегов, обновлений, слова из сервиса Google Books и т.п.). Как показывают приведенные в табл. 2 данные, агрегированные показатели, извлеченные из текстового контента социальных сетей, имеют гораздо более высокий предсказательный потенциал в сравнении с предсказаниями на основе данных индивидуального уровня.

К. Беллет и П. Фрайтерс подчеркивают, что предсказание и изучение благополучия с помощью технологий больших данных открывают широкие возможности перед исследователями, так как не требуется

проведения дорогостоящих масштабных социологических опросов. Способность определять текущие настроения населения и уровень удовлетворенности жизнью, являющейся бесспорным преимуществом изучения благополучия при помощи больших данных, может пригодиться самым разным действующим лицам. Для правительства такие технологии позволяют отслеживать последствия политических изменений или внешних событий на местных сообществах, влияние этих локальных событий в пространстве и времени. Это особенно важно, когда речь идет о социальных мультипликаторах событий, которые непосредственно влияют на отдельные подгруппы населения, таких как безработица или связанные с идентичностью потрясения. Способность определять текущее настроение и уровень удовлетворенности жизнью населения как на индивидуальном, так и на агрегированном уровне также может быть использована для

целенаправленного управления: правительства и организации могут целенаправленно воздействовать на группы населения с низким уровнем настроения или удовлетворенностью жизнью посредством специальной политики, направленной на эти группы (например, оказывать больше помощи в сфере психического здоровья или открыть больше детских учреждений для раннего возраста). Оппозиционные партии могут намеренно понижать высокий уровень удовлетворенности жизнью и обвинять правительство в его низком уровне. Рекламодатели могут адаптировать свои сообщения к настроению отдельных людей и их избирателей [3. С. 106].

Хотя данные о настроении, извлеченные из социальных сетей, могут не полностью соответствовать показателям удовлетворенности жизнью или счастья, полученным с помощью социологических опросов, эти данные позволяют осуществить гораздо более детальный анализ благополучия в масштабе одного дня. К. Беллет и П. Фрайтерс приводят в качестве примера использования масштабных данных для оценки эмоционального состояния населения исследование Г. Лейва настроения населения Мексики по данным Twitter [62]. В этом исследовании анализируется, как менялось соотношение положительных и отрицательных слов ежедневно в период 2016–2018 гг. в Twitter. Исследование Г. Лейва демонстрирует резкий взлет положительного настроения в определенные дни, такие как Рождество 2017 г. или день, когда Мексика обыграла Германию на чемпионате мира по футболу 2018 г., и всплеск отрицательного настроения в такие дни, как землетрясение в 2017 г., проигрыш на чемпионате мира Бразилии или избрание Дональда Трампа на выборах Президента в США в 2016 г. К. Беллет и П. Фрайтерс делают вывод о том, что на основе анализа данных Twitter можно получить вполне правдоподобные профили настроения и что отдельные события довольно недолговечны с точки зрения их влияния на настроение в Twitter [3. С. 108].

К. Беллет и П. Фрайтерс также обращают внимание на изменение подхода к измерению благополучия с наступлением цифровой эпохи. В доинтернетный период измерялись в основном объективные показатели: количество собственности, социальные отношения (брак), рождаемость и смертность, формы аттестации (образование, профессиональная подготовка, гражданство), потоки доходов (занятость, социальное обеспечение, налоги), другие виды деятельности (преступность, судебные дела, общественные организации, крупные покупки). Измерение во всех этих случаях обычно было явным и осуществлялось с помощью форм и систем, о которых население было достаточно хорошо информировано. Цифровая эпоха, в частности внедрение технологии больших данных, существенно увеличила возможности получать информацию об индивидуальном поведении – о его перемещении в пространстве, о процессах, происходящих внутри организма человека (с помощью различных медицинских устройств), о таких физических атрибутах, как отпечатки пальцев, структура сетчатки, рост, вес, частота сердечных сокращений, мозговая активность [3. С. 109].

В завершении представленного обзора обратимся к использованию цифровых данных как источника информации о качестве жизни и благополучии в российской научной литературе. К сожалению, нам не известны отечественные исследования, в которых используются цифровые следы для изучения качества жизни и благополучия, кроме тех, которые были реализованы под руководством авторы данной статьи. Можно выделить несколько исследований, в которых сделан акцент на влиянии процесса цифровизации на качество жизни населения, т.е. выдержанные в рамках экстернорной стратегии. Например, Н.А. Восколович исследует влияние цифровой трансформации сферы услуг на качество жизни населения и приходит к выводу, что этот процесс оказывает значительное влияние на качество жизни населения [63]. Г.П. Литвинцева и С.П. Петров выделяют 6 блоков качества жизни, которые обусловлены процессом цифровизации: качество населения, характеризуемое его цифровыми компетенциями; обеспеченность цифровыми благами (цифровое благосостояние населения); качество трудовой сферы; качество социальной сферы; качество электронных государственных и муниципальных услуг населению; безопасность информационной деятельности населения [64. С. 419]. Анализируя статистические данные по перечисленным блокам, авторы предлагают аналитический инструмент – российский региональный индекс цифровой составляющей качества жизни населения, который позволяет оценить уровень цифровизации качества жизни населения регионов РФ.

Интернорной стратегии исследования качества жизни и благополучия в контексте цифровизации придерживается автор данной статьи. Им реализованы два исследовательских проекта, в которых использовались в качестве источника данных цифровые следы пользователей. В первом исследовании на материале социальной сети Instagram исследовались образы благополучия, которые транслируют пользователи в своих постах [65]. Для этого были разработаны словари метафор благополучия и соответствующих им лингвистических маркеров. Далее авторы собирали сообщения (подписи и комментарии к фотографиям, изображениям и видео, которые разместили пользователи) в социальной сети Instagram, содержащие эти лингвистические маркеры. После тщательной фильтрации в выборке осталось 137 315 сообщений, которые были сгруппированы в крупные семантические кластеры, включающие тематически и семантически близкие метафоры. Авторы установили, что наиболее часто встречаются сообщения, характеризующие максимально интенсивное выражение положительных эмоций (счастье и/или наслаждение, удовольствие, радость и т.д.), далее следуют семантические кластеры, характеризующие любовь и просто позитивные эмоции, а также кластер, в котором присутствуют сообщения о позитивном опыте релаксации (развлечения, отдых, путешествия).

В другом исследовании были использованы для оценки уровня благополучия в различных регионах

РФ цифровые следы пользователей в социальной сети «ВКонтакте» [66, 67]. В данном исследовании авторы анализировали региональные сообщества в социальной сети «ВКонтакте», и в качестве источника данных выступали сообщения, которые размещают пользователи в этих сообществах и которые характеризуют те или иные аспекты благополучия жизни людей. Авторы выделили 19 социальных индикаторов (показателей) благополучия, в исследовании рассматривались только сообщения, которые тематически относятся к какому-либо из этих социальных индикаторов. Также с помощью машинного обучения для каждого сообщения определялась тональность сообщения. Далее, анализируя онлайн-активность пользователей по отношению к таким сообщениям (количество лайков, репостов и комментариев), авторы разработали индекс (не)благополучия регионов РФ. В статье [66] представлен индекс для 43 регионов РФ за 2018 г. Такое название инструмента измерения благополучия («индекс (не)благополучия регионов РФ») обусловле-

но спецификой коммуникации в онлайн-пространстве – поскольку большее внимание пользователей привлекают сообщения, которые содержат какую-либо негативную информацию, то в контексте анализа качества жизни и благополучия данный инструмент высвечивает в большей мере проблемные моменты в жизни конкретного региона, те аспекты, которые снижают качество жизни в этом регионе.

Таким образом, мы можем сделать вывод, что сегодня цифровые методы вызывают значительный интерес у исследователей и получили определенное признание в качестве достаточно надежного инструмента изучения благополучия и качества жизни. Вместе с тем существует целый ряд серьезных аргументов, заставляющих относиться к исследованиям с использованием методов цифровой социологии достаточно осторожно. Как представляется, дальнейшее совершенствование технологической составляющей данных исследований позволит сделать их столь же точными, как и традиционные опросные методы в социологии.

ЛИТЕРАТУРА

1. Фомин М.В. Технологии качества жизни и постиндустриальная эпоха // Вопросы философии. 2016. № 3. С. 139–146.
2. Крыштановская О.В. Бесконтактная социология: новые формы исследований в цифровую эпоху // Цифровая социология. 2018. № 1. С. 4–9.
3. Bellet C., Frijters P. Big Data and Well-being // World Happiness Report 2019 / eds. by J. Helliwell, R. Layard, J. Sachs. New York : Sustainable Development Solutions Network, 2019. P. 97–122.
4. Sanchez C.R., Craglia M., Bretz A.K. New data sources for social indicators: the case study of contacting politicians by Twitter // International Journal of Digital Earth. 2017. Vol. 10, is. 8. P. 829–845.
5. Schober M.F., Pasek J., Guggenheim L., Lampe C., Conrad F.G. Research Synthesis : Social Media Analyses for Social Measurement // Public Opinion Quarterly. 2016. Vol. 80, is. 1. P. 180–211.
6. Дудина В.И., Юдина Д.И. Извлекаем мнения из сети Интернет : могут ли методы анализа текстов заменить опросы общественного мнения? // Мониторинг общественного мнения : Экономические и социальные перемены. 2017. № 5. С. 63–78.
7. Ravi K., Ravi V. A survey on opinion mining and sentiment analysis: tasks, approaches and applications // Knowledge-Based Systems. 2015. Vol. 89. P. 14–46.
8. Cody E.M., Reagan A.J., Dodds P.S., Danforth C.M. Public Opinion Polling with Twitter. 2016. URL: <http://www.uvm.edu/pdodds/research/papers/cody2016b/>
9. Antenucci D., Cafarella M., Levenstein M., Re C., Shapiro M.D. Using Social Media to Measure Labor Market Flows // NBER Working Papers 20010. 2014. URL: <https://ideas.repec.org/p/nbr/nberwo/20010.html>
10. Blue J., Condell J., Lunney T. Digital Footprints: Your Unique Identity // Proceedings of the 32nd International BCS Human Computer Interaction Conference (HCI). Belfast, UK : BCS Learning and Development Ltd, 2018. P. 1–5.
11. Conlin M. You are what you post // BusinessWeek. 2006. March 27. URL: <https://www.bloomberg.com/news/articles/2006-03-26/you-are-what-you-post>
12. Флёров О.В. Цифровой след человека в Интернете: основные гуманитарные подходы // Образовательные ресурсы и технологии. 2018. № 4 (25). С. 79–82.
13. Awaworyi Churchill S., Mishra V. Trust, Social Networks and Subjective Wellbeing in China // Social Indicators Research. 2017. Vol. 132. P. 313–339.
14. Beyens I., Pouwels J.L., van Driel I.I., Keijsers L., Valkenburg P.M. The effect of social media on well-being differs from adolescent to adolescent // Scientific Reports. 2020. Vol. 10. Art. No. 10763. URL: <https://www.nature.com/articles/s41598-020-67727-7>
15. Bourgeois A., Bower J., Carrol A. Social Networking and the Social and Emotional Wellbeing of Adolescents in Australia // Australian Journal of Guidance and Counselling. 2014. Vol. 24, is. 2. P. 167–182.
16. Brusilovskiy E., Townley G., Sneath G., Salzer M.S. Social media use, community participation and psychological well-being among individuals with serious mental illnesses // Computers in Human Behavior. 2016. Vol. 65. P. 232–240.
17. Krishnan A., Zhou X. Modeling the effect of health antecedents and social media engagement on healthy eating and quality of life // Journal of Applied Communication Research. 2019. Vol. 47, is. 4. P. 365–380.
18. Lee P.S.N., Leung L., Lo V., Xiong C., Wu T. Internet communication versus face-to-face interaction in quality of life // Social Indicators Research. 2010. Vol. 100. P. 375–389.
19. McDool E., Powell P., Roberts J., Taylor K. The internet and children's psychological wellbeing // Journal of Health Economics. 2020. Vol. 69. Art. No. 102274. URL: <https://www.sciencedirect.com/science/article/pii/S0167629618311408?via%3Dihub>
20. Sabatini F., Sarracino F. Online Networks and Subjective Well-Being // Kyklos. 2017. Vol. 70, is. 3. P. 456–480.
21. Settanni M., Azucar D., Marengo D. Predicting Individual Characteristics from Digital Traces on Social Media : A Meta-Analysis // Cyberpsychology, Behavior, and Social Networking. 2018. Vol. 21, is. 4. P. 217–228.
22. Slegers K., Van Boxtel M.P.J., Jolles J. Effects of computer training and Internet usage on the well-being and quality of life of older adults: a randomized, controlled study // Journals of Gerontology : Psychological Sciences and Social Sciences. 2008. Vol. 63, is. 3. P. 176–184.
23. Summerskill B. Online social networks and wellbeing // Lancet. 2009. Vol. 374. P. 514.
24. Verduyn P., Ybarra O., Resibois M., Jonides J., Kross E. Do social network sites enhance or undermine subjective well-being? // A critical review. Social Issues and Policy Review. 2017. Vol. 11, is. 1. P. 274–302.
25. Wheatley D., Buglass S.L. Social network engagement and subjective well-being: a life-course perspective // British Journal of Sociology. 2019. Vol. 70, is. 1. P. 1971–1995.
26. Yang Y., Yuan Y., Archer N., Ryan E. Adoption of Social Media and the Quality of Life of Older Adults // 49th Hawaii International Conference on System Sciences (HICSS). 2016. P. 3133–3142. URL: <https://ieeexplore.ieee.org/document/7427577>

27. Hawi N.S., Samaha M. The Relations Among Social Media Addiction, Self-Esteem, and Life Satisfaction in University Students // *Social Science Computer Review*. 2017. Vol. 35, is. 5. P. 576–586.
28. Longstreet P., Brooks S. Life satisfaction: A key to managing internet & social media addiction // *Technology in Society*. 2017. Vol. 50. P. 73–77.
29. Zhan L., Sun Y., Wang N., Zhang X. Understanding the influence of social media on people's life satisfaction through two competing explanatory mechanisms // *Aslib Journal of Information Management*. 2016. Vol. 68, is. 3. P. 347–361.
30. Burke M., Marlow C., Lento T. Social Network Activity and Social Well-Being // *Proceedings of the SIGCHI conference on human factors in computing systems*. 2010. P. 1909–1912.
31. Chen H.-T., Li X. The contribution of mobile social media to social capital and psychological well-being: Examining the role of communicative use, friending and self-disclosure // *Computers in Human Behavior*. 2017. Vol. 75. P. 958–965.
32. Ellison N.B., Vitak J., Gray R., Lampe C. Cultivating social resources on social network sites: Facebook relationship maintenance behaviors and their role in social capital processes // *Journal of Computer-mediated Communication*. 2014. Vol. 19. P. 855–870.
33. Valenzuela S., Halpern D., Katz J.E. Social network sites, marriage well-being and divorce: Survey and state-level evidence from the United States // *Computers in Human Behavior*. 2014. Vol. 36. P. 94–101.
34. Aalbers G., McNally R.J., Heeren A., de Wit S., Fried E.I. Social media and depression symptoms: a network perspective // *Journal of Experimental Psychology: General*. 2019. Vol. 148, is. 8. P. 1454–1462.
35. Alt D. Students' Wellbeing, Fear of Missing out, and Social Media Engagement for Leisure in Higher Education Learning Environments // *Current Psychology*. 2018. Vol. 37. P. 128–138.
36. Appel H., Gerlach A.L., Crusius J. The interplay between Facebook use, social comparison, envy, and depression // *Current Opinion in Psychology*. 2016. Vol. 9. P. 44–49.
37. Lin L.Y., Sidani J.E., Shensa A., Radovic A., Miller E., Colditz J.B., Hoffman B.L., Giles L.M., Primack B.A. Association between social media use and depression among U.S. young adults // *Depression and Anxiety*. 2016. Vol. 33. P. 323–331.
38. McCloskey W., Iwanicki S., Lauterbach D., Giammittorio D.M., Maxwell K. Are Facebook "Friends" Helpful? Development of a Facebook-Based Measure of Social Support and Examination of Relationships Among Depression, Quality of Life, and Social Support // *Cyberpsychology, behavior and social networking*. 2015. Vol. 18, is. 9. P. 499–505.
39. Reer F., Tang W.Y., Quandt T. Psychosocial well-being and social media engagement: The mediating roles of social comparison orientation and fear of missing out // *New Media & Society*. 2019. Vol. 21, is. 7. P. 1486–1505.
40. Hunt M.G., Marx R., Lipson C., Young J. No More FOMO: Limiting Social Media Decreases Loneliness and Depression // *Journal of Social and Clinical Psychology*. 2018. Vol. 37, is. 10. P. 751–768.
41. Song H., Zmyslinski-Seelig A., Kim J., Drent A., Victor A., Omori K., Allen M. Does Facebook make you lonely?: A meta analysis // *Computers in Human Behavior*. 2014. Vol. 36. P. 446–452.
42. Hardy B.W., Castonguay J. The moderating role of age in the relationship between social media use and mental well-being: An analysis of the 2016 General Social Survey // *Computers in Human Behavior*. 2018. Vol. 85. P. 282–290.
43. O'Reilly M., Dogra N., Whiteman N., Hughes J., Eruyar S., Reilly P. Is social media bad for mental health and wellbeing? Exploring the perspectives of adolescents // *Clinical Child Psychology and Psychiatry*. 2018. Vol. 23, is. 4. P. 601–613.
44. Rasmussen E.E., Puyanant-Carter N., LaFreniere J.R., Norman M.S., Kimball T.G. The serially mediated relationship between emerging adults' social media use and mental well-being // *Computers in Human Behavior*. 2020. Vol. 102. P. 206–213.
45. Lee K.-T., Noh M.-J., Koo D.-M. Lonely people are no longer lonely on social networking sites: The mediating role of self-disclosure and social support // *Cyberpsychology, Behavior, and Social Networking*. 2013. Vol. 16. P. 413–418.
46. Chou H.-T.G., Edge N. 'They Are Happier and Having Better Lives than I Am': The Impact of Using Facebook on Perceptions of Others' Lives // *Cyberpsychology, behavior, and social networking*. 2012. Vol. 15, is. 2. P. 117–121.
47. Priyadarshini C., Dubey R., Kumar Y.L.N., Jha R. Impact of Social Media Addiction on Employees' Wellbeing and Work Productivity // *Qualitative Report*. 2020. Vol. 25. P. 181–196.
48. Hao B., Li L., Gao R., Li A., Zhu T. Sensing Subjective Well-Being from Social Media // *Proceedings of 10th International Conference Active Media Technology*. Springer. 2014. P. 324–336.
49. Schwartz H.A., Sap M., Kern M.L., Eichstaedt J.C., Kapelner A., Agrawal M., Blanco E., Dziurzynski L., Park G., Stillwell D., Kosinski M., Seligman M.E., Ungar L.H. Predicting individual well-being through the language of social media // *Pacific Symposium on Biocomputing*. 2016. Vol. 21. P. 516–527.
50. Schwartz H.A., Eichstaedt J.C., Kern M.L., Dziurzynski L., Lucas R.E., Agrawal M., Park G.J., Lakshmikanth S.K., Jha S., Seligman M.E., Ungar L. Characterizing geographic variation in well-being using tweets // *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*. Cambridge, MA, 2013. P. 583–591.
51. Chen L., Gong T., Kosinski M., Stillwell D., Davidson R.L. Building a profile of subjective well-being for social media users // *PloS ONE*. 2017. Vol. 12, is. 11. Art. No. e0187278.
52. Wu K., Ma J., Chen Z., Ren P. Analysis of Subjective City Happiness Index Based on Large Scale Microblog Data // *Proceedings of Web Technologies and Applications : 17th Asia-Pacific Web Conference*. Springer, 2015. P. 365–377.
53. Wang N., Kosinski M., Stillwell D.J., Rust J. Can Well-Being be Measured Using Facebook Status Updates? Validation of Facebook's Gross National Happiness Index // *Social Indicators Research*. 2014. Vol. 115. P. 483–491.
54. Yang C., Srinivasan P. Life Satisfaction and the Pursuit of Happiness on Twitter // *PLoS ONE*. 2016. Vol. 11, is. 3. Art. No. e0150881.
55. Mitchell L., Frank M.R., Harris K.D., Dodds P.S., Danforth C.M. The Geography of Happiness: Connecting Twitter Sentiment and Expression, Demographics, and Objective Characteristics of Place // *PLoS ONE*. 2013. Vol. 8, is. 5. Art. No. e64417.
56. Algan Y., Murtin F., Beasley E., Higa K., Senik C. Well-being through the lens of the Internet // *PLoS ONE*. 2019. Vol. 14, is. 1. Art. No. e0209562.
57. Collins S., Sun Y., Kosinski M., Stillwell D., Markuzon N. Are you satisfied with life? Predicting satisfaction with life from Facebook // *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*. Springer, 2015. P. 24–33.
58. Kosinski M., Stillwell D., Graepel T. Private traits and attributes are predictable from digital records of human behavior // *Proceedings of the National Academy of Sciences*. 2013. Vol. 110, is. 15. P. 5802–5805.
59. Liu P., Tov W., Kosinski M., Stillwell D.J., Qiu L. Do Facebook status updates reflect subjective well-being // *Cyberpsychology, Behavior, and Social Networking*. 2015. Vol. 18, is. 7. P. 373–379.
60. Algan Y., Beasley E., Guyot F., Higa K., Murtin F., Senik C. Big Data Measures of Well-Being: Evidence from a Google Well-Being Index in the United States // *OECD Statistics Working Papers*. 2016. URL: https://www.oecd-ilibrary.org/economics/big-data-measures-of-well-being_5j1z9hpg0rd1-en
61. Hills T., Proto E., Sgroi D. Historical analysis of national subjective well-being using millions of digitized books // *Nature : Human Behaviour*. 2019. Vol. 3, is. 12. P. 1271–1275.
62. Leyva G. The mood of twitterers in Mexico. 2018. URL: <https://www.cepal.org/sites/default/files/presentations/presentation-the-mood-of-twitterers-mexico-inegi.pdf>
63. Восколович Н.А. Измерение влияния цифровой трансформации сферы услуг на качество жизни населения // *Государственное управление. Электронный вестник*. 2019. № 75. С. 6–23.

64. Литвинцева Г.П., Петров С.П. Теоретические основы взаимодействия цифровой трансформации и качества жизни населения // Журнал экономической теории. 2019. Т. 16, № 3. С. 414–427.
65. Shchekotin E., Baryshev A., Mundrievskaya Y. Images of Well-Being in Everyday Communication // Proceedings of the 32nd International Business Information Management Association Conference (IBIMA 2018) : Vision 2020: Sustainable Economic Development and Application of Innovation Management from Regional Expansion to Global Growth. 15–16 November 2018, Seville Spain. P. 2663–2673.
66. Щекотин Е.В., Мягков М.Г., Гойко В.Л., Кашпур В.В., Коварж Г.Ю. Субъективная оценка (не)благополучия населения регионов РФ на основе данных социальных сетей // Мониторинг общественного мнения: Экономические и социальные перемены. 2020. № 1. С. 78–116.
67. Щекотин Е.В., Коварж Г.Ю., Гойко В.Л., Петров Е.Ю., Бакулин В.В. Оценка качества жизни населения регионов РФ на основе цифровых данных: методологические аспекты // Векторы благополучия: экономика и социум. 2020. № 3 (38). С. 138–156.

Статья представлена научной редакцией «Социология и политология» 25 мая 2021 г.

Digital Footprints as a New Source of Data on Quality of Life and Well-Being: An Overview of Current Trends

Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal, 2021, 467, 170–181.

DOI: 10.17223/15617793/467/21

Evgeny V. Shchekotin, Novosibirsk State University of Economics and Management (Novosibirsk, Russian Federation). E-mail: evgvik1978@mail.ru

Keywords: digital footprints; quality of life; well-being; big data; machine learning; digital sociology.

The study is supported by the Russian Foundation for Basic Research, Project No. 20-011-00391.

The article presents an overview of foreign research devoted to the study of the quality of life and well-being of the population by methods of digital sociology, i.e. with the help of new information and communication technologies (such as big data, machine learning, social network analysis, etc.). The most promising and actively developing area of research is the study of digital footprints of users, primarily in social networks (Facebook, Twitter, Weibo, etc.). The advantages (a high degree of detail in the distribution of user ratings over time and space, ability to get more information in general, availability at any time, speed of research, reducing the cost of conducting surveys, flexibility of methodology, etc.) and disadvantages (the problem of representativeness and reliability of data) of using such a methodology to study the quality of life and well-being of the population are discussed. Two main strategies for studying the quality of life and well-being of the population are considered, based on a different understanding of the role and significance of information and communication technologies in people's daily lives. The first strategy – external – considers digital technologies as an external and self-sufficient factor that affects people's daily lives. This strategy is to identify the impact of digital technologies (mainly social media) on individual nuances of quality of life and well-being (life satisfaction, social capital, social support, mental health, depression and fears, loneliness, etc.). The second strategy – internal – does not single out digital technologies as a separate external factor but considers digital technologies as an integral part of people's daily existence. In this case, life in the online information space and life in the physical offline environment are inseparable components of the holistic existence of a modern person, so the analysis of human behavior and reactions in the information space carries as reliable information about the quality of life and well-being of a person as traditional data sources (surveys, etc.). It is concluded that despite significant methodological difficulties and a certain caution of many sociologists regarding the methods of digital sociology, this direction has great prospects and further improvement of the technological component of these studies will make digital methods of studying the quality of life and well-being as reliable as traditional survey methods in sociology.

REFERENCES

1. Fomin, M.V. (2016) *Life Quality Technologies and Postindustrial Era. Voprosy filosofii*. 3. pp. 139–146. (In Russian).
2. Kryshatanovskaya, O.V. (2018) Kryshatanovskaya OV Contactless Sociology: New Forms of Research in a Digital Age. *Tsifrovaya sotsiologiya – Digital Sociology*. 1. pp. 4–9. (In Russian).
3. Bellet, C. & Frijters, P. (2019) Big Data and Well-being. In: Helliwell, J., Layard, R. & Sachs, J. (eds) *World Happiness Report 2019*. New York: Sustainable Development Solutions Network. pp. 97–122.
4. Sanchez, C.R., Craglia, M. & Bregt, A.K. (2017) New data sources for social indicators: the case study of contacting politicians by Twitter. *International Journal of Digital Earth*. 10 (8. pp. 829–845.
5. Schober, M.F. et al. (2016) Research Synthesis: Social Media Analyses for Social Measurement. *Public Opinion Quarterly*. 2016. 80, Is.1. pp. 180–211.
6. Dudina, V.I. & Yudina, D.I. (2017) Mining Opinions on the Internet: Can the Text Analysis Methods Replace Public Opinion Polls? *Monitoring obshchestvennogo mneniya: Ekonomicheskie i sotsial'nye peremeny – The Monitoring of Public Opinion Economic&Social Changes*. 5. pp. 63–78. (In Russian). DOI: 10.14515/monitoring.2017.5.05
7. Ravi, K. & Ravi, V. (2015) A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowledge-Based Systems*. 89. pp. 14–46.
8. Cody, E.M. et al. (2016) *Public Opinion Polling with Twitter*. [Online] Available from: <http://www.uvm.edu/pdodds/research/papers/cody2016/>
9. Antenucci, D. et al. (2014) Using Social Media to Measure Labor Market Flows. *NBER Working Papers 20010*. [Online] Available from: <https://ideas.repec.org/p/nbr/nberwo/20010.html>
10. Blue, J., Condell, J. & Lunney, T. (2018) Digital Footprints: Your Unique Identity. *Proceedings of the 32nd International BCS Human Computer Interaction Conference (HCI)*. Belfast, UK: BCS Learning and Development Ltd. pp. 1–5.
11. Conlin, M. (2006) You are what you post. *BusinessWeek*. March 27. [Online] Available from: <https://www.bloomberg.com/news/articles/2006-03-26/you-are-what-you-post>
12. Flerov, O.V. (2018) Human Digital Internet Trace: Humanitarian Approaches. *Obrazovatel'nye resursy i tekhnologii – Education Resources and Technologies*. 4 (25). pp. 79–82. (In Russian).
13. Awaworyi Churchill, S. & Mishra, V. (2017) Trust, Social Networks and Subjective Wellbeing in China. *Social Indicators Research*. 132. pp. 313–339.
14. Beyens, I. et al. (2020) The effect of social media on well-being differs from adolescent to adolescent. *Scientific Reports*. 10. Art. No. 10763. [Online] Available from: <https://www.nature.com/articles/s41598-020-67727-7>
15. Bourgeois, A., Bower, J. & Carrol, A. (2014) Social Networking and the Social and Emotional Wellbeing of Adolescents in Australia. *Australian Journal of Guidance and Counselling*. 24 (2). pp. 167–182.
16. Brusilovskiy, E., Townley, G., Snethen, G. & Salzer, M.S. (2016) Social media use, community participation and psychological well-being among individuals with serious mental illnesses. *Computers in Human Behavior*. 65. pp. 232–240.

17. Krishnan, A. & Zhou, X. (2019) Modeling the effect of health antecedents and social media engagement on healthy eating and quality of life. *Journal of Applied Communication Research*. 47 (4). pp. 365–380.
18. Lee, P.S.N., Leung, L., Lo, V., Xiong, C. & Wu, T. (2010) Internet communication versus face-to-face interaction in quality of life. *Social Indicators Research*. 100. pp. 375–389.
19. McDool, E., Powell, P., Roberts, J. & Taylor, K. (2020) The internet and children's psychological wellbeing. *Journal of Health Economics*. 69. Art. No. 102274. [Online] Available from: <https://www.sciencedirect.com/science/article/pii/S0167629618311408?via%3Dihub>
20. Sabatini, F. & Sarracino, F. (2017) Online Networks and Subjective Well-Being. *Kyklos*. 70 (3). pp. 456–480.
21. Settanni, M., Azucar, D. & Marengo, D. (2018) Predicting Individual Characteristics from Digital Traces on Social Media: A Meta-Analysis. *Cyberpsychology, Behavior, and Social Networking*. 21 (4). pp. 217–228.
22. Slegers, K., Van Boxtel, M.P.J. & Jolles, J. (2008) Effects of computer training and Internet usage on the well-being and quality of life of older adults: a randomized, controlled study. *Journals of Gerontology: Psychological Sciences and Social Sciences*. 63 (3). pp. 176–184.
23. Summerskill, B. (2009) Online social networks and wellbeing. *Lancet*. 374. p. 514.
24. Verduyn, P. et al. (2017) Do social network sites enhance or undermine subjective well-being? A critical review. *Social Issues and Policy Review*. 11 (1). pp. 274–302.
25. Wheatley, D. & Buglass, S.L. (2019) Social network engagement and subjective well-being: a life-course perspective. *British Journal of Sociology*. 70 (1). pp. 1971–1995.
26. Yang, Y., Yuan, Y., Archer, N. & Ryan, E. (2016) Adoption of Social Media and the Quality of Life of Older Adults. *49th Hawaii International Conference on System Sciences (HICSS)*. pp. 3133–3142. [Online] Available from: <https://ieeexplore.ieee.org/document/7427577>
27. Hawi, N.S. & Samaha, M. (2017) The Relations Among Social Media Addiction, Self-Esteem, and Life Satisfaction in University Students. *Social Science Computer Review*. 35 (5). pp. 576–586.
28. Longstreet, P. & Brooks, S. (2017) Life satisfaction: A key to managing internet & social media addiction. *Technology in Society*. 50. pp. 73–77.
29. Zhan, L., Sun, Y., Wang, N. & Zhang, X. (2016) Understanding the influence of social media on people's life satisfaction through two competing explanatory mechanisms. *Aslib Journal of Information Management*. 68 (3). pp. 347–361.
30. Burke, M., Marlow, C. & Lento, T. (2010) Social Network Activity and Social Well-Being. *Proceedings of the SIGCHI conference on human factors in computing systems*. pp. 1909–1912.
31. Chen, H.-T. & Li, X. (2017) The contribution of mobile social media to social capital and psychological well-being: Examining the role of communicative use, friending and self-disclosure. *Computers in Human Behavior*. 75. pp. 958–965.
32. Ellison, N.B., Vitak, J., Gray, R. & Lampe, C. (2014) Cultivating social resources on social network sites: Facebook relationship maintenance behaviors and their role in social capital processes. *Journal of Computer-mediated Communication*. 19. pp. 855–870.
33. Valenzuela, S., Halpern, D. & Katz, J.E. (2014) Social network sites, marriage well-being and divorce: Survey and state-level evidence from the United States. *Computers in Human Behavior*. 36. pp. 94–101.
34. Aalbers, G. et al. (2019) Social media and depression symptoms: a network perspective. *Journal of Experimental Psychology: General*. 148 (8). pp. 1454–1462.
35. Alt, D. (2018) Students' Wellbeing, Fear of Missing out, and Social Media Engagement for Leisure in Higher Education Learning Environments. *Current Psychology*. 37. pp. 128–138.
36. Appel, H., Gerlach, A.L. & Crusius, J. (2016) The interplay between Facebook use, social comparison, envy, and depression. *Current Opinion in Psychology*. 9. pp. 44–49.
37. Lin, L.Y. (2016) Association between social media use and depression among U.S. young adults. *Depression and Anxiety*. 33. pp. 323–331.
38. McCloskey, W., Iwanicki, S., Lauterbach, D., Giammitto, D.M. & Maxwell, K. (2015) Are Facebook "Friends" Helpful? Development of a Facebook-Based Measure of Social Support and Examination of Relationships Among Depression, Quality of Life, and Social Support. *Cyberpsychology, Behavior and Social Networking*. 18 (9). pp. 499–505.
39. Reer, F., Tang, W.Y. & Quandt, T. (2019) Psychosocial well-being and social media engagement: The mediating roles of social comparison orientation and fear of missing out. *New Media & Society*. 21 (7). pp. 1486–1505.
40. Hunt, M.G., Marx, R., Lipson, C. & Young, J. (2018) No More FOMO: Limiting Social Media Decreases Loneliness and Depression. *Journal of Social and Clinical Psychology*. 37 (10). pp. 751–768.
41. Song, H. et al. (2014) Does Facebook make you lonely?: A meta analysis. *Computers in Human Behavior*. 36. pp. 446–452.
42. Hardy, B.W. & Castonguay, J. (2018) The moderating role of age in the relationship between social media use and mental well-being: An analysis of the 2016 General Social Survey. *Computers in Human Behavior*. 85. pp. 282–290.
43. O'Reilly, M. et al. (2018) Is social media bad for mental health and wellbeing? Exploring the perspectives of adolescents. *Clinical Child Psychology and Psychiatry*. 23 (4). pp. 601–613.
44. Rasmussen, E.E., Punyanunt-Carter, N., LaFreniere, J.R., Norman, M.S. & Kimball, T.G. (2020) The serially mediated relationship between emerging adults' social media use and mental well-being. *Computers in Human Behavior*. 102. pp. 206–213.
45. Lee, K.-T., Noh, M.-J. & Koo, D.-M. (2013) Lonely people are no longer lonely on social networking sites: The mediating role of self-disclosure and social support. *Cyberpsychology, Behavior, and Social Networking*. 16. pp. 413–418.
46. Chou, H.-T.G. & Edge, N. (2012) 'They Are Happier and Having Better Lives than I Am': The Impact of Using Facebook on Perceptions of Others' Lives. *Cyberpsychology, Behavior, and Social Networking*. 15 (2). pp. 117–121.
47. Priyadarshini, C., Dubey, R., Kumar, Y.L.N. & Jha, R. (2020) Impact of Social Media Addiction on Employees' Wellbeing and Work Productivity. *Qualitative Report*. 25. pp. 181–196.
48. Hao, B., Li, L., Gao, R., Li, A. & Zhu, T. (2014) Sensing Subjective Well-Being from Social Media. *Proceedings of 10th International Conference Active Media Technology*. Springer. pp. 324–336.
49. Schwartz, H.A. et al. (2016) Predicting individual well-being through the language of social media. *Pacific Symposium on Biocomputing*. 21. pp. 516–527.
50. Schwartz, H.A. et al. (2013) Characterizing geographic variation in well-being using tweets. *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*. Cambridge, MA. pp. 583–591.
51. Chen, L. et al. (2017) Building a profile of subjective well-being for social media users. *PLoS ONE*. 12 (11). Art. No. e0187278.
52. Wu, K., Ma, J., Chen, Z. & Ren, P. (2015) Analysis of Subjective City Happiness Index Based on Large Scale Microblog Data. *Proceedings of Web Technologies and Applications: 17th Asia-Pacific Web Conference*. Springer. pp. 365–377.
53. Wang, N., Kosinski, M., Stillwell, D.J. & Rust, J. (2014) Can Well-Being be Measured Using Facebook Status Updates? Validation of Facebook's Gross National Happiness Index. *Social Indicators Research*. 115. pp. 483–491.
54. Yang, C. & Srinivasan, P. (2016) Life Satisfaction and the Pursuit of Happiness on Twitter. *PLoS ONE*. 11 (3). Art. No. e0150881.
55. Mitchell, L. et al. (2013) The Geography of Happiness: Connecting Twitter Sentiment and Expression, Demographics, and Objective Characteristics of Place. *PLoS ONE*. 8 (5). Art. No. e64417.
56. Algan, Y., Murtin, F., Beasley, E., Higa, K. & Senik, C. (2019) Well-being through the lens of the Internet. *PLoS ONE*. 14 (1). Art. No. e0209562.
57. Collins, S., Sun, Y., Kosinski, M., Stillwell, D. & Markuzon, N. (2015) Are you satisfied with life? Predicting satisfaction with life from Facebook. *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*. Springer. pp. 24–33.

58. Kosinski, M., Stillwell, D. & Graepel, T. (2013) Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*. 110 (15). pp. 5802–5805.
59. Liu, P., Tov, W., Kosinski, M., Stillwell, D.J. & Qiu, L. (2015) Do Facebook status updates reflect subjective well-being. *Cyberpsychology, Behavior, and Social Networking*. 18 (7). pp. 373–379.
60. Algan, Y. (2016) Big Data Measures of Well-Being: Evidence from a Google Well-Being Index in the United States. *OECD Statistics Working Papers*. [Online] Available from: https://www.oecd-ilibrary.org/economics/big-data-measures-of-well-being_5jlz9hpg0rd1-en
61. Hills, T., Proto, E. & Sgroi, D. (2019) Historical analysis of national subjective well-being using millions of digitized books. *Nature: Human Behaviour*. 3 (12). pp. 1271–1275.
62. Leyva, G. (2018) *The mood of twitterers in Mexico*. [Online] Available from: <https://www.cepal.org/sites/default/files/presentations/presentation-the-mood-of-twitterers-mexico-inegi.pdf>
63. Voskolovich, N.A. (2019) Measuring the Impact of Digital Transformation of Services on the Quality of Life. *Gosudarstvennoe upravlenie. Elektronnyy vestnik – Public Administration. E-Journal*. 75. pp. 6–23. (In Russian).
64. Litvintseva, G.P. & Petrov, S.P. (2019) Theoretical Foundations of Digital Transformation of Economy and People's Quality of Life. *Zhurnal ekonomicheskoy teorii*. 16 (3). pp. 414–427. (In Russian).
65. Shchekotin, E., Baryshev, A. & Mundrievskaya, Y. (2018) Images of Well-Being in Everyday Communication. *Proceedings of the 32nd International Business Information Management Association Conference (IBIMA 2018): Vision 2020: Sustainable Economic Development and Application of Innovation Management from Regional Expansion to Global Growth*. 15–16 November. Seville, Spain. pp. 2663–2673.
66. Shchekotin, E.V. et al. (2020) Subjective Measurement of Population Well-Being/Well-Being in the Russian Regions Based on Social Media Data. *Monitoring obshchestvennogo mneniya: Ekonomicheskie i sotsial'nye peremeny – Monitoring of Public Opinion: Economic and Social Changes Journal (Public Opinion Monitoring)*. 1. pp. 78–116. (In Russian). DOI: 10.14515/monitoring.2020.1.05
67. Shchekotin, E.V. et al. (2020) Assessment of the Population Life Quality in Russian Regions Based on Digital Data: Methodological Aspects. *Vektory blagopoluchiya: ekonomika i sotsium – Journal of Wellbeing Technologies*. 3 (38). pp. 138–156. (In Russian).

Received: 25 May 2021