

АЛЬТЕРНАТИВНЫЕ ИСТОЧНИКИ СТАТИСТИЧЕСКОЙ ИНФОРМАЦИИ КАК ОСНОВА ПРИНЯТИЯ ПОЛИТИЧЕСКИХ РЕШЕНИЙ

Морозов А.Н.¹

Аннотация

Традиционно в качестве источника информации для принятия решений в государственном управлении выступают данные статистических наблюдений, однако они не всегда актуальны и не всегда соответствуют требованиям исследователей. В качестве альтернативы статистическим наблюдениям могут быть рассмотрены административные данные, а также иные источники информации. В статье представлен обзор международного опыта использования административных данных и альтернативных источников информации в процессе принятия политических решений. Показано, что статистические ведомства многих стран используют с целью сокращения расходов на проведение собственных наблюдений административные данные и централизованные системы доступа к ним. Также в качестве альтернативных источников в целях городского планирования, в экологии или сельском хозяйстве могут быть использованы геопространственные данные. Что касается «больших данных», то они пока не находят широкого применения в государственном управлении, тем не менее рассматриваются как перспективный источник информации. Проведенное исследование позволило сделать вывод, что перспективными направлениями реформирования деятельности статистических ведомств являются: предоставление пользователям открытого доступа к дезагрегированным данным микроуровня; замена данных статистических наблюдений административными данными; предоставление доступа к пространственным данным, информации спутниковых наблюдений, а также к обезличенным данным пользователей государственных услуг.

Ключевые слова: административные данные; «большие данные»; доказательная политика; официальная статистика; политические решения.

¹ Морозов Антон Николаевич – младший научный сотрудник, Российской академия народного хозяйства и государственной службы при Президенте РФ; младший научный сотрудник кафедры конкурентной и промышленной политики экономического факультета МГУ. Адрес: Московский государственный университет имени М.В. Ломоносова. 119991, г. Москва, ГСП-1, Ленинские горы. E-mail: antmorozov88@gmail.com

Введение

Подход к принятию политических решений, основанный на идеологии, общественном мнении, а также ценностях и традициях общества, часто критикуется в литературе за необъективность (Marchi, Lucertini и Tsoukiàs, 2016; Holt, 2008; Scott, 2005; Sabatier, 1991), поскольку принимаемые в рамках такого подхода решения во многом зависят от политической конъюнктуры и позиции элит, а статистическая и иная информация используется преимущественно на заключительном этапе мониторинга уже реализованных мер.

Для обоснования экономической политики во многих странах все шире применяется доказательный подход (evidence based policy), при котором решения формулируются и принимаются на основе исследований, имеющих под собой эмпирическую базу. Такой подход минимизирует риск несоответствия ожидаемых эффектов управления его результатам.

Доказательный подход позволяет ослабить (хотя и не устраняет полностью) негативные влияния политической конъюнктуры. Необходимость фактических данных и данных результатов научных исследований при таком подходе предполагает возможность доступа для исследователей и лиц, принимающих политические решения, к достоверной и актуальной информации. Как правило, за предоставление такой информации для целей государственного управления отвечает статистическая служба. Соответственно, препятствиями в применении доказательного подхода являются некорректность статистических данных, ограничение доступа к ним либо их полное отсутствие. Решение данной проблемы видится, с одной стороны, в модернизации методик проведения статистических наблюдений и, возможно, функций самого статистического ведомства, чтобы собираемая им информация соответствовала требованиям исследователей и политиков; с другой – в поиске альтернативных источников информации, которые могут быть использованы в процессе принятия решений. Положительный аспект такого подхода состоит в независимости исследователя от статистического ведомства. Альтернативная информация может оказаться более доступной и своевременной. В то же время такая информация зачастую сложно формализуема (например, тексты сообщений из социальных сетей), либо ее обработка требует серьезных затрат ресурсов (Big Data), либо ее сложно получить в частном порядке (административные данные).

Поскольку издержки использования некоторых видов данных для индивидуального исследователя могут оказаться высоки, во многих странах существуют государственные программы, обеспечивающие доступ к различного рода альтернативным источникам.

Доказательный подход к принятию политических решений

В соответствии с классической моделью политического цикла, разработанной еще в середине 1950-х гг. (см., например: Jann & Wegrich, 2007), выделяются пять основных этапов принятия политических решений:

1. Постановка вопроса.
2. Формулирование политики.
3. Принятие решения (выбор из альтернативных вариантов).
4. Реализация политики.
5. Оценка политики.

Субъект принятия политических решений может совершить серьезную ошибку на каждом из этапов, например, не учесть все альтернативы на этапе принятия решения. Как отмечает М. Хоулетт (Howlett, 2009), такие ошибки можно свести к минимуму, если в процессе принятия решения опираться на некоторую достоверную информацию, которая бы проясняла состав доступных альтернатив или ресурсные возможности государства. Подобную информацию политик может получить, например, используя выводы научных работ. Такой подход, при котором лицо, ответственное за принятие политических решений, опирается на результаты существующих исследований, называется *доказательным* (evidence-based).

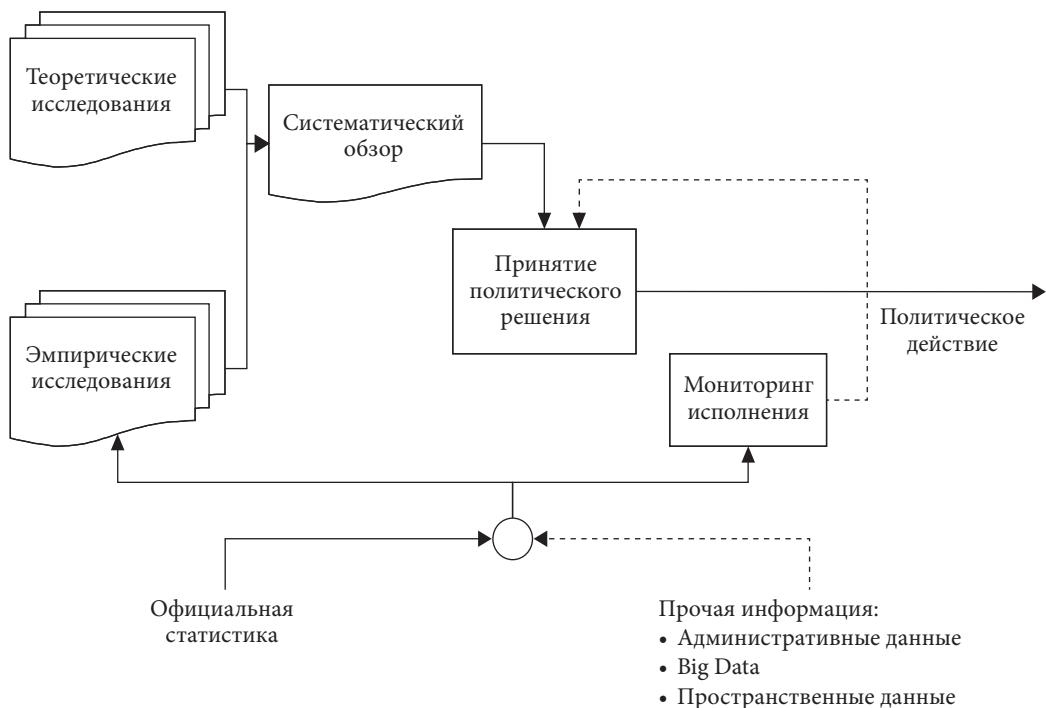
Идея использовать результаты научных исследований для проведения государственной политики отнюдь не нова. Рассматривая историю развития доказательного подхода, У. Бартлетт (Bartlett, 2013) упоминает опыт применения экономических исследований в США в 1960-х гг., в частности, это касалось таких отраслей, как оборона и развитие городских территорий. Деятельность Госплана в СССР, по мнению Б.А. Райзберга (2008), также основывалась на результатах исследований подведомственных научных учреждений. Современное название этого явления (evidence-based policy), по всей видимости, возникло на волне успехов «доказательной медицины» в начале 1990-х гг. Первые попытки внедрения «доказательного подхода» в политическом процессе как раз коснулись отрасли здравоохранения (Bartlett, 2013).

Особый интерес к доказательному подходу приходится на конец 1990-х гг., когда в Великобритании к власти приходит лейбористское правительство во главе с Т. Блэром. Новое правительство провозглашает необходимость применять только те методы, которые «действительно работают» и фактически провозглашают доказательный подход новой идеологией (Marais & Matebesi, 2013). В США аналогичную политику начал проводить президент Обама. Доказательный подход применяется также в Австралии, Канаде, Новой Зеландии, ЮАР, его рекомендуют к применению такие международные организации, как Международный банк, ОЭСР и ЮНЕСКО (Davies, 2012). Стоит также упомянуть «Кокрейн» (Cochrane Collaboration) – организацию, осуществляющую методическую поддержку систематических обзоров, применяемых в целях доказательного подхода к принятию политических решений.

На рисунке 1 представлена схема осуществления доказательного подхода. Результаты существующих исследований, проведенных на базе статистических данных и иной доступной информации, агрегируются в виде обзоров. Субъект принятия политических решений, с одной стороны, опирается на эти обзоры, с другой – изучает результат мониторинга предыдущих решений. Это должно обеспечить его информацией о доступной альтернативе и ожидаемом эффекте принятия политического решения.

Рисунок 1

Схема осуществления доказательного подхода



Следует понимать, что «доказательство» само по себе не способно подсказать лицу, ответственному за решение, что ему стоит делать. Отдельно взятое доказательство несет в себе риск неопределенности, случайной ошибки, которая могла произойти при проведении анализа (Davies, 2012). Не исключаются также преднамеренные фальсификации результатов. Л. Розенсток и Л.Дж. Ли (Rosenstock & Lee, 2002) в своей работе приводят ряд примеров, когда на ученых оказывалось финансовое или моральное давление с целью повлиять на результаты их работы². Следовательно, опрометчиво было бы основывать политику на одном-единственном исследовании, как бы мы ни доверяли его автору.

Чтобы агрегировать информацию по всей предметной области, проводимые в рамках осуществления доказательного подхода исследования опираются на *систематические обзоры* литературы. Такие обзоры содержат анализ источников, отобранных на основе заранее определенной поисковой стратегии, что позволяет выделить наиболее релевантные. Нередко выражается мнение, что такой подход может быть излишне затратным. Кроме того, составители обзоров не всегда имеют достаточный опыт в конкретной предметной области (см. об этом: Anderson, 2010; Haux, 2013). Однако

² Так, с целью дискредитации правительенного отчета о вреде пассивного курения представители американской табачной промышленности профинансировали ряд критических статей в ведущих научных журналах (Rosenstock & Lee, 2002).

применение систематических обзоров позволяет снизить ошибки, связанные с недостоверными результатами отдельного исследования, поскольку в таком обзоре существует возможность сопоставить выводы всех работ по заданной тематике.

Так как проблемы методологии и фальсификации результатов в некоторой мере решаются за счет применения механизма систематических обзоров, на первый план выходит проблема качества входных данных. Насколько бы ни была верна методология работы, ее результат заведомо будет некорректен, если исследователь использует искаженные данные. Источником информации для политических исследований могла бы быть национальная статистическая служба. Однако официальная статистика отнюдь не всегда представляет актуальную и корректную информацию, состав которой отвечал бы требованиям исследовательских задач.

Проблемы использования данных статистических ведомств

Анализ опыта различных стран показывает, что случаи, когда государственное статистическое ведомство не справляется с задачей информационного обеспечения политического процесса, отнюдь не редки. Зачастую проблема кроется в недостатке внимания государства к проблеме сбора и обработки статистической информации (Scott, 2005). Если руководство страны принимает политические решения, опираясь на идеологический базис, а не на обоснованные доказательства, то такое государство не заинтересовано в развитии официальной статистики. Напротив, возможны ситуации искажения фактов со стороны статистических служб для легитимизации принятых решений. В частности, подобные обвинения довольно часто звучат в адрес Китая, где можно обнаружить определенные искажения в данных о рабочей силе, потреблении или затратах на образование. В частности, в своей работе Г. Чау (Chow, 2006) указывает, что результаты его исследования подтверждают наличие подобных искажений, хотя, по его мнению, они вряд ли носят преднамеренный характер.

Отсутствие должного внимания к проблемам официальной статистики со стороны государства может находить отражение и в применении устаревших технологий обработки данных (Tambuwala et al., 2012), а также в использовании некорректных методик сбора данных и расчета показателей (Webster, 2002). Так, А.К. Соловьев и С.А. Донцова (2015) отмечают, что показатель среднесписочной численности, рассчитываемый по методике Росстата, не учитывает некоторые категории занятого населения. Н.В. Мкртчян (2012) указывает, что в России во время переписи 2010 г. имели место двойной учет пенсионеров и, напротив, недоучет детей до 6 лет, а также некоторые искажения, связанные с учетом внутренней и внешней миграции. В российской статистике существует и ряд других проблем: несопоставимость временных рядов, отсутствие доступа к дезагрегированным данным (Бессонов, 2014).

Для развитых стран характерна проблема, когда официальная статистика испытывает влияние этических требований. Так, например, в соот-

ветствии с британским законодательством персонализированные данные не должны носить дискриминационный характер, т.е. не должны содержать информацию о национальности или расовой принадлежности респондента. Как отмечает С.Форсон (Forson, 2006), подобное ограничение затрудняет исследования культурных, гендерных или этнических особенностей, поскольку такую информацию оказывается невозможно получить из официальных источников.

Проблемы с устаревшими или некорректными методами сбора и обработки информации могут быть решены путем подготовки квалифицированных кадров для статистических ведомств. В частности, Дж. Лейн (Lane, 2016) считает, что необходимо внести изменения в образовательные программы, чтобы обучить специалистов в области data science, способных реализовать эту задачу. Однако такой подход не может быть реализован в короткие сроки и требует длинного горизонта планирования.

В краткосрочной перспективе нивелировать негативные тенденции в официальной статистике позволило бы активное использование низкоуровневых данных государственных органов, бюджетных и частных организаций – административных данных. В частности, государственные органы в процессе исполнения своих основных функций так или иначе накапливают массивы данных, содержащие информацию о гражданах и компаниях, которые могли бы быть использованы как в целях принятия политических решений, так и для экспертных исследований. При этом такие данные остаются лишь в пользовании одного ведомства, хотя содержащаяся в них информация могла бы быть полезна в работе других ведомств и при принятии политических решений.

Использование административных данных меняет роль статистического ведомства. Оно начинает выполнять координационную функцию: объединяет разрозненные базы данных в единую платформу, проводит стандартизацию справочников и классификаторов, предоставляет интерфейс доступа для пользователей. Тогда исследователь окажется способен самостоятельно выполнить проверку таких данных, провести их очистку и при необходимости агрегировать в показатели, которые окажутся удобными для его расчетов.

Административные данные

Административные данные могут быть использованы как статистическими службами, так и непосредственно исследователями. Статистические службы используют такие данные как основу для официальной статистики, заменяя ими стандартную процедуру статистического наблюдения. Исследования, в том числе проводимые с целью поддержания процесса принятия политических решений, используют административные данные и как альтернативу официальной статистике, и вместе со статистической информацией, чтобы контролировать ее точность либо восполнить нехватку в полноте.

Статистические ведомства западных стран активно используют административные данные для подготовки отчетов, как это происходит, напри-

мер, в Финляндии, Швеции или Дании (Nordbotten, 2008). Это позволяет, с одной стороны, экономить средства на проведении собственных наблюдений, как это указано в регламенте канадской статистической службы (Statistics Canada, 2017). С другой стороны, такая информация в большей степени соответствует отраслевым стандартам и требованиям пользователей, поскольку собирается специализированными ведомствами. Кроме того, такой подход требует организации единой системы хранения информации, когда все данные микроуровня собраны на одном сервере. Как отмечается в докладе П. Менезес и Л. Д’Агиар (Menezes & D’Aguiar, 2013), единая система хранения позволяет при анализе агрегированных данных «проваливаться» в их структуру, рассматривать их на необходимом уровне детализации.

Примером использования административных данных в целях официальной статистики может быть начавшаяся в 1994 г. реформа шведского статистического ведомства. До этого всю статистическую информацию в Швеции собирал централизованный орган. Однако на волне критики качества официальной статистики парламент Швеции принял решение о реформировании статистической службы. В результате функции сбора, обработки и предоставления статистической информации по определенным отраслям были переданы соответствующим государственным органам и министерствам. 24 различных государственных органа подготавливают статистические отчеты на основе доступных им административных данных. Центральное статистическое ведомство продолжило функционировать, но его роль теперь скорее координационная: методологическое обеспечение, стандартизация и классификация, а также подготовка агрегированных отчетов (Statistics Sweden, 2007).

Конечно, исследователь может запрашивать административные данные в частном порядке у руководства каждого обладающего ими учреждения, однако данный процесс связан со значительными трансакционными издержками, особенно если исследование проводится на уровне региона или страны. Поэтому некоторые государства с целью поддержания открытости информации стараются обеспечивать доступность таких данных для исследователей, формируя единые базы данных с универсальным режимом доступа. Подобная система существует с 2012 г. в Великобритании – «The Administrative Data Research Network» (ADRN). Система содержит информацию о доступных регистрах административных данных. Желающий использовать такие данные в своем исследовании высылает запрос, содержащий информацию о планируемой работе. ADRN проверяет запрос на предмет этичности, законности, научной ценности предполагаемого исследования и принимает решение о предоставлении доступа к данным (Administrative Data Research Network, 2017). Аналогичная система обеспечения доступа к данным по социальной защите и занятости населения существует в Германии³ (Research Data Centre, 2018).

³ Центр исследования данных Федеральной службы занятости при Институте исследований рынка труда и занятости (Das Forschungsdatenzentrum der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung).

Как правило, подобные системы предоставляют преимущество в доступе к административным данным для национальных пользователей, что связано с правовыми ограничениями на использование низкоуровневых данных. Снять такие ограничения был призван проект «Данные без границ» (The Data without Boundaries, DwB). Этот проект объединил 10 европейских национальных статистических ведомств, 11 архивов и 7 образовательных и коммерческих организаций. Проект был призван обеспечить равный доступ к административным данным для различных категорий европейских пользователей. Однако в 2015 г. финансирование проекта было прекращено в связи с истечением срока предоставления гранта на его реализацию (Data without Boundaries, 2015). Несмотря на то что проект уже завершен, он имеет активную страницу в сети Интернет, где содержится информация о всех разработках в области безопасного международного доступа к административным данным.

Помимо универсальных и централизованных площадок доступа к административным данным существуют отраслевые проекты. Так, П. Менезес и Л. Д'Агиар (Menezes & D'Aguiar 2013) приводят пример Банка Португалии (Banco de Portugal), использующего в своей работе доступные ему административные данные о деятельности экономических субъектов как альтернативу официальной статистике.

В сфере медицины подобные проекты активно применяются как руководством учреждений, так и на уровне региональных властей. Проект SOCRATES (Systems Outcomes and Clinical Resources Administrative Efficiency Software), объединяющий базы данных лечебных учреждений, входящих в систему Университетских больниц Кливленда (150 отделений), содержит информацию о пациентах и лечащих врачах и позволяет вести учет в соответствии с единой классификацией (Lawrence & Delaney, 2013). На уровне руководства региона или штата административные данные могут быть применены для обеспечения процесса диспансерного наблюдения, выделяя граждан (потенциальных пациентов), находящихся в группе риска (Stein et al., 2014). Наконец, такие данные могут быть использованы для мониторинга и контроля качества оказания медицинских услуг со стороны государства (Simonet & Alkafaji, 2017).

Примером использования административных данных для целей государственного управления в сфере здравоохранения является система, направленная на повышение качества диспансерного наблюдения HARP-CDM (Hospital Admission Risk Program – Chronic Disease Management), действующая в Мельбурне (Berlowitz & Graco, 2010). Данная система на основе определенных характеристик, связанных с повышенной вероятностью необходимости госпитализации, отирает граждан, с тем чтобы пригласить их на осмотр.

До внедрения этой системы аналогичные функции выполняли 14 программ, которые курировали четыре отдельные организации. Каждая из программ рассматривала отдельные виды заболеваний. После внедрения HARP-CDM данные этих 14 программ были объединены в одну общую базу. В результате лечебные учреждения стали способны эффективно назначать

плановое обследование гражданам. Отметим, что система также предоставляет органам власти возможность оценивать проводимые мероприятия и определять те группы заболеваний, которым следует уделить большее внимание.

Аналогичная система, ориентированная на оказание психиатрической помощи, действует в США, в штате Нью-Йорк (Smith et al., 2011). Она позволяет своевременно определять пациентов, которые не осуществляют предписанные амбулаторные процедуры (необоснованно прерывают лечение, вовремя не получают рецепты). Система использует административные данные лечебных учреждений и анализирует информацию о лицах, которые обязаны пройти лечение по решению суда, а также о тех, кому была предоставлена психиатрическая помощь амбулаторно или в стационаре в течение предыдущих 12 месяцев. Если система выдает предупреждение, то ответственный исполнитель получает информацию о пациенте, которого следует посетить. В результате внедрения системы отмечается значительное снижение числа отказов от лечения или невыполнения рецептов (Stein et al., 2014).

Тем не менее для административных данных, также как и для централизованной статистики, актуальна проблема полноты первичных данных. Она связана, с одной стороны, с проблемой морального риска, когда лица, ответственные за сбор данных, оказываются заинтересованы в завышении или занижении некоторых показателей⁴ (Heinrich, 2008); с другой – с тем, что часть данных может в принципе не отражаться в административных регистрах (Forson, 2006).

Отсутствующие либо неполные данные могут быть дополнены из альтернативных источников, наиболее перспективными из которых являются «большие данные».

Большие данные

С развитием цифровых технологий все более актуальным становится вопрос использования «больших данных» (Big Data). Это вся информация, персональная и обезличенная, собираемая производителями программного обеспечения, поисковыми сервисами, крупными ритейлерами и т.д. о своих пользователях и покупателях. Фактически этот термин обозначает гигантские массивы неструктурированных данных и методы работы с ними. Существует подход к определению «больших данных» через их свойства – так называемые 3V: Volume, Velocity, Variety (**объем, скорость, разнородность**) (Malomo & Sena, 2017; Abdulkadri, Evans & Ash, 2016; Severo, Feredj & Romele, 2016; Kim Trimi & Chung, 2014).

Для хранения «больших данных» требуется колоссальный **объем** дискового пространства, который сегодня исчисляется зеттабайтами (10^{21} байт) и иоттабайтами (10^{24} байт).

⁴ Классический пример манипуляции данными касается ситуации, когда заработок учителей зависит от оценок школьников, что впоследствии приводит к завышению средних показателей успеваемости.

Свойство **скорости** отражает тот факт, что новые данные появляются все быстрее. В ответ на это возникает проблема поддержания скорости обработки и передачи таких данных.

Разнообразие связано с тем, что данные приходят из различных источников: структурированных (например, база данных сотового оператора); с частичной структурой (текст статьи, содержащий выделенные ключевые слова); не структурированных совсем (простой текст, не имеющий жесткой структуры).

Чтобы быть отнесенными к категории Big Data, данные должны отвечать *всем* перечисленным свойствам. Например, данные переписи облашают большим объемом, но не могут быть отнесены к Big Data, поскольку для них не характерны прочие свойства: скорость обновления и разнообразие форм.

Среди источников Big Data растет число неструктурированных данных, поэтому некоторые авторы (см., например: Abdulkadri et al., 2016; Severo et al., 2016) также выделяют четвертое V – Veracity (**точность**). Данные должны быть четко определены, не допускать двусмысленности или случайного характера их формирования. Важное следствие этого свойства – возможность проверки: один и тот же запрос к определенному источнику всегда должен возвращать один и тот же набор данных.

Применение «больших данных» позволяет агрегировать информацию по большому числу наблюдений, при этом всегда остается возможность раскрыть агрегированные данные до уровня отдельного индивида (или отдельного факта) – это делает их незаменимыми при анализе предпочтений потребителей. Причем с развитием технологий их становится все дешевле собирать и обрабатывать (Rieder & Simon, 2016). Ф. Маломо и В. Сена (Malomo & Sena, 2017) предполагают, что методы обработки «больших данных» могут быть применены к крупным массивам административных данных, содержащих информацию по отдельным индивидам. «Большие данные» в перспективе могут составить альтернативу переписи. В работе Д. Флореску и др. (Florescu et al., 2014) отмечается, что уже сейчас применение этой технологии позволяет вести учет населения небольших регионов, проводить оценку миграции.

Учитывая перспективность применения «больших данных» для процесса принятия политических решений, некоторые страны на государственном уровне организуют проекты, использующие технологию Big Data: США – «Return in Review Project» (RRP), Сингапур – «Risk Assessment and Horizon Scanning» (RAHS), Великобритания – «Horizon Scanning Center» (HSC). Отметим, что некоторые подобные проекты⁵ являются сугубо научными и не связаны с потенциальным применением их результатов в области государственного управления. Однако они развиваются методологию обработки больших данных. Другие же, формально заявленные как использующие «большие данные»⁶, фактически осуществляют обра-

⁵ Например: в США – «NASA: GEOSS», в ЕС – «DOME».

⁶ США – «Syracuse: Smarter City Project», «Michigan: MSDW», Корея – «KOSTAT: EPS», «ACRC: CIAS».

ботку просто объемных массивов данных, которые не могут быть классифицированы как Big Data: не соответствуют свойствам Velocity и Variety (используемые данные структурированы, но нет высокой скорости появления новых данных).

Международные организации также уделяют внимание этому вопросу. На уровне ООН существует программа по развитию использования «больших данных» для целей официальной статистики (United Nations, 2018). ОЭСР периодически публикует отчеты по тематике «больших данных», а также выпускает серию работ «Digital Economy Papers», посвященную среди прочего теме «больших данных». В отчетах ОЭСР отмечается, что такие данные могут найти применение в сфере коммунального хозяйства, при оптимизации транспортных потоков и в иных сферах государственного управления (OECD, 2015).

Интернет на данный момент является основным источником «больших данных». Анализ сообщений в социальных сетях, выполненный с использованием теории графов, позволяет понять структуру общества, например, выделить группы населения, поддерживающие определенные политические решения (Jung, Park, Wu & Park, 2015). В другом исследовании (Panagiotopoulos, Bowen & Brooker, 2017) на основе анализа сообщений в сети Twitter ученым удалось выявить, что британские производители молочной продукции интересуются возобновляемой энергетикой, а представители пахотного земледелия – финансовой поддержкой со стороны государства и местных фермерских объединений. Учет различных интересов этих групп фермеров позволит разрабатывать адресные меры поддержки сельхозпроизводителей. Анализ сообщений оказывается дешевле, чем стандартные методы получения информации от населения, например опросы или фокус-группы. Кроме того, данные можно обновлять с необходимой частотой.

Источником данных могут быть также сообщения в СМИ. Так, например, Л. Шер, Дж. Мюллер и З. Фокс (Sherr, Mueller & Fox, 2009) предлагают подход к поиску информации о брошенных детях, используя анализ новостных сообщений. Помимо количественных данных о числе брошенных детей авторам удалось получить дополнительные сведения: пол, возраст и национальность детей; были ли при них какие-то дополнительные вещи (одежда, игрушки, записки); была ли впоследствии найдена мать.

Такую информацию из официальных источников получить невозможно. В результате анализа данных СМИ авторы смогли выделить факторы, влияющие на выживаемость брошенных детей. Такой набор данных позволяет получить наиболее полное представление о проблеме брошенных детей и больше, чем официальная статистика, подходит для анализа этого феномена.

Как и любой источник информации, «большие данные» имеют свои ограничения. Несмотря на развитие механизмов защиты информации, в последние годы растут опасения пользователей за безопасность своих персональных данных (OECD, 2017). Граждане утрачивают доверие к ме-

ханизму хранения данных. В целях решения проблемы доверия аналитики ОЭСР рекомендуют осуществлять политику по защите персональных данных на принципах открытости, когда каждый пользователь имеет возможность ознакомиться с той информацией, которая передается в сеть, и выбрать (или сменить) организацию, ответственную за хранение данных (OECD, 2015).

Существует риск того, что применяемые аналитические алгоритмы могут привести к неожиданным отрицательным результатам. Так, в 2012 г. компания Knight Capital Group менее чем за час потеряла 440 млн долл. только потому, что за часть операций отвечала автоматизированная торговая система (Metha, 2012).

Непредсказуемость алгоритмов обработки «больших данных» также подрывает доверие к ним, но уже со стороны исследователей и государственных органов. В такой ситуации необходимы механизм легализации Big Data, выделение лучших практик и источников данных, которые могут быть использованы в целях государственного управления. Функция по определению таких практик и источников данных может быть возложена на статистическое ведомство (Reimsbach-Kounatze, 2015). При этом ведомство может как самостоятельно заниматься обработкой больших данных, так и предоставлять консультации заинтересованным сторонам, разъясняя методики и предоставляя каталог доверенных источников данных.

Указанные ограничения являются причиной того, что практическое применение «больших данных» для целей государственного управления пока не нашло широкого применения. Однако с развитием технологий информационной безопасности и обработки «больших данных» такой источник информации окажется способен дополнять (или даже заменять) существующие статистические и административные данные.

Пространственные данные

Пространственные модели, построенные на основе спутниковых данных и информации о геолокации из мобильных приложений, могут быть применены в целях государственной политики совокупно с имеющимися административными или «большими данными» (United Nations, 2018). В частности, информация, полученная в результате анализа спутниковых снимков, может быть использована в целях принятия политических решений в сельском хозяйстве (Scott, 2005) или при организации сбора данных об объектах земельного рынка (Tambuwala et al., 2012).

Методы пространственного анализа могут найти применение для мониторинга экологической ситуации. С. Клер и И. Крид (Clare & Creed, 2014) в своем исследовании выясняют, как изменились площади заболоченных территорий. Для этого на основе анализа цифровой карты высот и климатических данных они строят пространственную модель болотистой местности Канады за 1999 и 2009 гг. Авторам удалось выявить, что в 80% случаев осушение болот было осуществлено без соответствующего согласия со стороны контролирующего органа.

А. Хосцило с соавторами (Hoscilo et al., 2015) предлагают методику для мониторинга изменений экосистемы, которая уже успешно применяется Еврокомиссией, рядом африканских национальных экологических служб, а также межправительственными организациями, например Африканским союзом. Авторы используют данные спутниковых и метеорологических наблюдений для отслеживания изменений в зеленом покрове на Африканском континенте. Выделены изменения климатического и неклиматического характера (влияние человека или экологических катастроф) за период 2001–2010 гг. Так, среди изменений неклиматического характера авторы отмечают большой очаг исчезновения зеленого покрова в Эфиопии.

Пространственный анализ может быть применен при оценке доступности для населения различных товаров или услуг, чтобы улучшать планировку города или региона. Одни авторы (Grubesic et al., 2016) предлагают применять метод «гравитационного потенциала» для осуществления политики контроля за оборотом алкогольной продукции. При таком подходе рассчитывается агрегированный показатель, учитывающий число точек продаж алкоголя в близлежащих районах, численность населения района и расстояние до каждой из точек продаж. Предложенная ими методика позволяет учесть не только постоянно проживающее население, но и ежедневную миграцию (на работу, учебу и назад) и, соответственно, более точно определить доступность алкогольной продукции. Другие авторы (Huerta, Munoz & Källestål, 2012) изучили вопрос доступности услуг здравоохранения для жителей Западной провинции Руанды. Они объединили как официальные данные о расположении медицинских учреждений и численности населения, так и географическую информацию о местности (высота, транспортная доступность, дорожное покрытие и т.д.). Учитывая указанные географические факторы, была построена пространственная модель провинции. На основе модели составлены карты доступности медицинских учреждений: число людей, которые фактически смогут до них добраться в течение часа. Выяснилось, что пешком до ближайшего учреждения могут дойти лишь 27% населения страны, на велосипеде – 58%, общественным транспортом – 34%. Полученные результаты позволяют улучшить распределение ресурсов.

Проблема доступности для населения медицинской помощи и иных социальных услуг актуальна для многих стран, в том числе и России, так что заимствование методов, примененных в указанных работах (Huerta Munoz & Källestål, 2012; Grubesic et al., 2016), представляется целесообразным для развития подходов к городскому и региональному планированию. Методы, использованные в вышеприведенных исследованиях (Clare & Creed, 2014; Hoscilo et al., 2015), могли бы быть использованы для наблюдения за вырубкой леса, добычей полезных ископаемых, а также для контроля незаконного строительства – за теми видами активности, результаты которых заметны на спутниковых снимках. Таким образом, пространственные данные оказываются важным дополнением к существующим статистическим и административным данным.

Заключение

С распространением цифровых технологий все более доступными становятся различные альтернативные источники информации, которые способны стать успешной заменой официальной статистики. В связи с этим, данные статистических ведомств, получаемые на базе традиционных методов наблюдений, теряют свою ценность для исследователей. Государственные органы ежегодно накапливают колоссальные массивы данных о гражданах и компаниях, которые могут являться источником актуальной информации.

Современное статистическое агентство не занимается непосредственным сбором и обработкой информации, а скорее выполняет координационную и методологическую функцию, поддерживая систему хранения информации, обеспечивая сопоставимость данных и предоставляя удобный интерфейс доступа для заинтересованных лиц.

Что касается будущего как статистических служб, так и в целом процесса информационного обеспечения процесса государственного управления, то здесь возрастает актуальность использования альтернативных источников информации и перспективных технологий, таких как Big Data и геоинформационные системы (ГИС). Приведенный обзор показывает, что некоторые подобные источники данных активно используются государственными органами уже в настоящее время, в частности, это касается пространственных данных. Важность применения геоинформационных данных и пространственного анализа отмечается даже на уровне ООН и ОЭСР. Подобные системы могут быть использованы в целях развития сельского хозяйства, экологии или планирования городской инфраструктуры.

Что касается российского статистического ведомства, то, на наш взгляд, реформирование работы Росстата могло бы осуществляться в четыре этапа – по степени возрастания сложности реализации:

- 1) предоставление доступа к дезагрегированным данным микроуровня;
- 2) использование административных данных, причем не только для текущих периодов, но и для ретроспективного анализа;
- 3) предоставление доступа к «большим данным», полученным в результате сбора обезличенной информации о пользователях государственных услуг, в том числе данные мобильных приложений государственных органов;
- 4) предоставление доступа к пространственным данным, полученным в результате спутникового наблюдения, и данным геолокации пользователей государственных услуг.

Примеры реформ в северных европейских странах, таких как Швеция или Дания, показывают, что статистическая служба может быть эффективно модернизирована. Кроме реформирования статистической службы государство могло бы обеспечить унифицированный механизм доступа к административным данным, как это, например, организовано в Великобритании или Германии. Полученная таким образом информация может как выступать в роли заменителя официальной статистики, так и использоваться совместно с ней.

ЛИТЕРАТУРА

1. Бессонов В.А. Что сохранит для истории современная российская статистика? [Электронный ресурс] : препринт WP2/2014/02 / В.А. Бессонов ; Нац. исслед. ун-т «Высшая школа экономики». (Серия WP2 «Количественный анализ в экономике»). – М.: Изд. дом Высшей школы экономики, 2014.
2. Мкртчян Н.В. Проблемы учета населения отдельных возрастных групп в ходе переписи населения 2010 г.: причины отклонений полученных данных от ожидаемых // Демографические аспекты социально-экономического развития.– М.: МАКС Пресс, 2012. – С. 197–214.
3. Райзберг Б.А. Госплан и государственное стратегическое планирование // Экономическая наука современной России. – 2008. – № 1 (40). – С. 116–124.
4. Соловьев А.К., Донцова С.А. Методологические проблемы пенсионной статистики // Народонаселение. – 2015. – № 4. – С. 79–92.
5. Abdulkadri A., Evans A., Ash T. An Assessment of Big Data for Official Statistics in the Caribbean: Challenges and Opportunities. – Naciones Unidas Comisión Económica para América Latina y el Caribe (CEPAL). 2016. N. 48.
6. Administrative Data Research Network. What We Do. – URL: <https://www.adrn.ac.uk/about/background/what-we-do/> (дата обращения: 04.01.2018).
7. Bartlett W. Obstacles to Evidence-based Policy-making in the EU Enlargement Countries: The Case of Skills Policies // Social Policy & Administration. 2013. Vol. 47. N. 4. P. 451–467.
8. Berlowitz D.J., Graco M. The Development of a Streamlined, Coordinated and Sustainable Evaluation Methodology for a Diverse Chronic Disease Management Program // Australian Health Review. 2010. Vol. 34. N. 2. P. 148–151.
9. Chow G. Are Chinese Official Statistics Reliable? // CESifo Economic Studies. 2006. Vol. 52. N. 2. P. 396–414.
10. Clare S., Creed I. Tracking Wetland Loss to Improve Evidence-Based Wetland Policy Learning and Decision Making // Wetlands Ecology & Management. 2014. Vol. 22. N. 3. P. 235–245.
11. Data without Boundaries. Data without Boundaries. – URL: <http://www.dwbproject.org/> (дата обращения: 04.01.2018).
12. Davies P. The State of Evidence-Based Policy Evaluation and its Role in Policy Formation // National Institute Economic Review. 2012. N. 219. P. R41–R52.
13. Florescu D., Karlberg M., Reis F., Rey P., Castillo D., Skaliotis M., & Wirthmann, A. Will ‘Big Data’ Transform Official Statistics? // European Conference on the Quality of Official Statistics. Vienna. Austria. 2014.
14. Forson C. The Strategic Framework for Women’s Enterprise: BME Women at the Margins // Equal Opportunities International. 2006. Vol. 25. N. 6. P. 418–432.
15. Grubesic T.H., Wei R., Murray A.T., Pridemore W.A. Comparative Approaches for Assessing Access to Alcohol Outlets: Exploring the Utility of a Gravity Potential Approach // Population Health Metrics. 2016. Vol. 14. N. 1. P. 25–42.
16. Heinrich C.J. Advancing Public Sector Performance Analysis // Applied Stochastic Models in Business and Industry. 2008. Vol. 24. N. 5. P. 373–389.

17. Holt D.T. Official Statistics, Public Policy and Public Trust // Journal of the Royal Statistical Society: Series A (Statistics in Society). 2008. Vol. 171. N. 2. P. 323–346.
18. Hoscilo A., Balzter H., Bartholomé E., Boschetti M., Brivio P.A., Brink A., Clerici, M., Pekel J. A Conceptual Model for Assessing Rainfall and Vegetation Trends in Sub-Saharan Africa from Satellite Data // International Journal of Climatology. 2015. Vol. 35. N. 12. P. 3582–3592.
19. Howlett M. Policy Analytical Capacity and Evidence-Based Policy-Making: Lessons from Canada // Canadian Public Administration. 2009. Vol. 52. N. 2. P. 153–175.
20. Huerta Munoz U., Källestål C. Geographical Accessibility and Spatial Coverage Modeling of the Primary Health Care Network in the Western Province of Rwanda. // International Journal of Health Geographics. 2012. Vol. 11. N. 1. P. 40–51.
21. Jann W., Wegrich K. Theories of the Policy Cycle // Handbook of Public Policy Analysis: Theory, Politics And Methods. New York. CRC Press. 2007. P. 43–62.
22. Jung K., Park S.J., Wu W.-N., Park H.W. A Webometric Approach to Policy Analysis and Management Using Exponential Random Graph Models // Quality & Quantity. 2015. Vol. 49. N. 2. P. 581–598.
23. Kim G.-H., Trimi S., Chung J.-H. Big-Data Applications in the Government Sector // Communications of the ACM. 2014. Vol. 57. N. 3. P. 78–85.
24. Lane J. Big Data: the Role of Education and Training // Journal of Policy Analysis & Management. 2016. Vol. 35. N 3. P. 722–724.
25. Lawrence J., Delaney C.P. Integrating Hospital Administrative Data to Improve Health Care Efficiency and Outcomes : “ The Socrates Story ” // Clinics in Colon and Rectal Surgery. 2013. Vol. 26. N. 1. P. 56–62.
26. Malomo F., Sena V. Data Intelligence for Local Government? Assessing the Benefits and Barriers to Use of Big Data in the Public Sector // Policy and Internet. 2017. Vol. 9. N. 1. P. 7–27.
27. Marais L., Matebesi Z. Evidence-Based Policy Development in South Africa: the Case of Provincial Growth and Development Strategies // Urban Forum. 2013. Vol. 24. N. 3. P. 357–371.
28. Marchi G., Lucertini G., Tsoukiàs A. From Evidence-Based Policy Making to Policy Analytics // Annals of Operations Research. 2016. Vol. 236. N. 1. P. 15–38.
29. Menezes P., D’Aguiar L. Impact and Benefits of Micro-Databases’ Integration on the Statistics of the Banco de Portugal // Proceedings of the 59th World Statistics Congress. Hong-Kong. 2013.
30. Metha N. Knight \$440 Million Loss Sealed by Rules on Canceling Trades. – URL: www.bloomberg.com/news/2012-08-14/knight-440-million-loss-sealed-bynew-rules-on-cancelling-trades.html (дата обращения: 17.01.2018).
31. Nordbotten S. The Use of Administrative Data in Official Statistics – Past, Present and Future – With Special Reference to the Nordic Countries In: Michael Carlson, Hans Nyquist, Mattias Villani (red.) // Official statistics: Methodology and Applications in Honour of Daniel Thorburn. 2010. P. 205–223.
32. OECD. Data-Driven Innovation. Paris : OECD Publishing. 2015.
33. OECD. Digital Economy Outlook 2017. Paris : OECD Publishing 2017.
34. Panagiotopoulos P., Bowen F., Brooker P. The Value of Social Media Data: Integrating Crowd Capabilities in Evidence-Based Policy // Government Information Quarterly. 2017. N. 34. P. 601–612.

35. Reimsbach-Kounatze C. The Proliferation of “Big Data” and Implications for Official Statistics and Statistical Agencies: A Preliminary Analysis. Working paper OECD Digital Economy Papers. Paris: OECD. 2015. N. 245.
36. Research Data Centre. Wir über uns. 2018. – URL: http://fdz.iab.de/de/FDZ_Scope_of_Services.aspx (дата обращения: 04.01.2018).
37. Rieder G., Simon J. Big Data: A New Empiricism and its Epistemic and Socio-Political Consequences. In: Pietsch W., Wernecke J., Ott M. (eds) Berechenbarkeit der Welt? Springer VS, Wiesbaden. 2017. – URL: DOI: https://doi.org/10.1007/978-3-658-12153-2_4 (дата обращения: 08.05.2018).
38. Rosenstock L., Lee L.J. Attacks on Science: The Risks to Evidence-Based Policy // American Journal of Public Health. 2002. Vol. 92. N. 1. P. 14–18.
39. Sabatier P.A. Toward Better Theories of the Policy Process // Political Science. 1991. Vol. 24. N. 2. P. 147–156.
40. Scott C. Measuring Up to the Measurement Problem: The Role of Statistics in Evidence-Based Policymaking // Proceedings of the 2005 CBMS Network Meeting. 2005. P. 35–93.
41. Severo M., Feredj A., Romele A. Soft Data and Public Policy: Can Social Media Offer Alternatives to Official Statistics in Urban Policymaking? // Policy and Internet. 2016. Vol. 8. N. 3. P. 354–372.
42. Sherr L., Mueller J., Fox Z. Abandoned Babies in the UK – A Review Utilizing Media Reports. // Child: Care, Health & Development. 2009. Vol. 35. N. 3. P. 419–430.
43. Simonet D., Alkafaji Y. Critical evaluations of the French health care accounting indicators: the use of DRGs // Public Adm. Q. 2017. T. 41. N. 3. C. 569–609.
44. Smith T.E., Appel A., Essock S.M., Jackson C.T., Karpati A., Marsik T., Myers R.W., Tom L., Sederer L.I. Public-Academic Partnerships. Using Medicaid Claims Data to Identify Service Gaps for High-Need Clients: The NYC Mental Health Care Monitoring Initiative. // Psychiatric Services. 2011. Vol. 62. N. 1. P. 9–11.
45. Statistics Canada. Use of Administrative Data. 2017. – URL: <https://www.statcan.gc.ca/pub/12-539-x/2009001/administrative-administratives-eng.htm> (дата обращения: 04.01.2018).
46. Statistics Sweden. Official statistics of Sweden. 2007. – URL: https://millenniumindicators.un.org/unsd/statcom/statcom_seminar/sweden.pdf (дата обращения: 04.01.2018).
47. Stein B.D., Pangilinan M., Sorbero M.J., Marcus S.M., Donahue S.A., Xu Y., Smith T.E., Essock S.M. Using Claims Data to Generate Clinical Flags Predicting Short-Term Risk of Continued Psychiatric Hospitalizations // Psychiatric Services. 2014. Vol. 65. N. 11. P. 1341–1346.
48. Tambuwala N., Rajabifard A., Bennett R., Williamson I., Wallace J. Inter-governmental Land Information Asymmetries in Australia // Journal of Spatial Science. 2012. Vol. 57. N. 1. P. 83–100.
49. United Nations. GWG for Big Data. 2018. – URL: <https://unstats.un.org/bigdata/> (дата обращения: 04.01.2018).
50. Webster D. Unemployment: How Official Statistics Distort Analysis and Policy, and Why // Radical Statistics. 2002. Vol. 79. P. 96–127.

ALTERNATIVE SOURCES OF STATISTICAL INFORMATION AS THE BASIS FOR POLITICAL DECISION MAKING

Anton N. Morozov

Junior researcher, Russian Presidential Academy of National Economy and Public Administration; Junior researcher of the Chair of Competition Policy and Industrial Policy, Faculty of Economics, Lomonosov MSU.

Address: Lomonosov Moscow State University, GSP-1, Leninskie Gory, 119991, Moscow, Russian Federation.
E-mail: antmorozov88@gmail.com

Abstract

To ensure that the outcome of the policy measures is in line with the initial expectations, many countries apply evidence based approach, in which policy decisions are made according to existing research. However, there arises a concern whether the information is correct and timely. Traditionally, governments employ the official statistics, but such data is not always relevant and useful for researchers. As an alternative for statistical observations, the administrative data could be used as well as other alternative information sources, the most advanced of which is the Big data. Especially due to the technological development the availability of such alternative sources is constantly growing.

The subject of the present paper is the international experience in using administrative data and some alternative information sources in the political decision-making process. The analysis shows that statistical offices in many countries employ administrative data reducing the cost of their own observations. Moreover there are centralized systems for accessing administrative data for research purposes. When it comes to alternative sources the role of geospatial data is noted, which is useful for urban planning, ecology or agriculture. The big data is not widely presented in public management yet; nevertheless it is the most promising information source. The result of current study could be applied in order to reform the statistical agency.

Keywords: administrative data; big data; evidence based policy; official statistics; policy making.

Citation: Morozov, A. N. (2018). Alternative sources of statistical information as the basis for political decision making. *Public Administration Issue*, no 2, pp. 50–70 (in Russian).

REFERENCES

1. Abdulkadri, A., Evans, A. & Ash, T. (2016). An Assessment of Big Data for Official Statistics in the Caribbean: Challenges and Opportunities. *Studies and Perspectives – The Caribbean*, no 48.

2. *Administrative Data Research Network*. (2017). What We Do. Available at: <https://www.adrn.ac.uk/about/background/what-we-do/> (accessed: 4 January, 2018)
3. Bartlett, W. (2013). Obstacles to Evidence-based Policy-making in the EU Enlargement Countries: The Case of Skills Policies. *Social Policy & Administration*, vol. 47 no 4, pp. 451–467.
4. Berlowitz, D. J., & Graco, M. (2010). The Development of a Streamlined, Coordinated and Sustainable Evaluation Methodology for a Diverse Chronic Disease Management Program. *Australian Health Review*, vol. 34, no 2, pp. 148–151.
5. Bessonov, V.A. (2014). *Chto sokhranit dlya istorii sovremennaya rossiyskaya statistika?* [What the Modern Russian Statistics Will Keep for History?]. Working paper WP2/2014/02, Moscow: HSE.
6. Chow, G. (2006). Are Chinese Official Statistics Reliable? *CESifo Economic Studies*, vol. 52, no 2, pp. 396–414.
7. Clare, S. & Creed, I. (2014). Tracking Wetland Loss to Improve Evidence-Based Wetland Policy Learning and Decision Making. *Wetlands Ecology & Management*, vol. 22, no 3, pp. 235–245.
8. *Data without Boundaries* (2015). Data without Boundaries. Available at: <http://www.dwbproject.org/> (accessed: 4 January, 2018).
9. Davies, P. (2012). The State of Evidence-Based Policy Evaluation and its Role in Policy Formation. *National Institute Economic Review*, no 219, R41-R52.
10. Florescu, D., Karlberg, M., Reis, F., Rey, P., Castillo, D., Skaliotis, M. & Wirthmann, A. (2014). Will “Big Data” Transform Official Statistics? *Proceedings of the European Conference on the Quality of official Statistics*. Vienna, Austria.
11. Forson, C. (2006). The Strategic Framework for Women’s Enterprise: BME Women at the Margins. *Equal Opportunities International*, vol. 25, no 6, pp. 418–432.
12. Grubesic, T.H., Wei, R., Murray, A.T. & Pridemore, W.A. (2016). Comparative Approaches for Assessing Access to Alcohol Outlets: Exploring the Utility of a Gravity Potential Approach. *Population Health Metrics*, vol. 14, no 1, pp. 25–42.
13. Heinrich, C.J. (2008). Advancing Public Sector Performance Analysis. *Applied Stochastic Models in Business and Industry*, vol. 24, no 5, pp. 373–389.
14. Holt, D.T. (2008). Official Statistics, Public Policy and Public Trust. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 171, no 2, pp. 323–346.
15. Hoscilo, A., Balzter, H., Bartholomé, E., Boschetti, M., Brivio, P. A., Brink, A., Clerici, M. & Pekel, J. F. (2015). A Conceptual Model for Assessing Rainfall and Vegetation Trends in Sub-Saharan Africa from Satellite Data. *International Journal of Climatology*, vol. 35, no 12, pp. 3582–3592.
16. Howlett, M. (2009). Policy Analytical Capacity and Evidence-Based Policy-Making: Lessons From Canada. *Canadian Public Administration*, vol. 52, no 2, pp. 153–175.
17. Huerta Munoz, U. & Källestål, C. (2012). Geographical Accessibility and Spatial Coverage Modeling of the Primary Health Care Network in the Western Province of Rwanda. *International Journal of Health Geographics*, vol. 11, no 1, pp. 40–51.
18. Jann, W. & Wegrich, K. (2007). *Theories of the Policy Cycle*. In: *Handbook of Public Policy Analysis: Theory, Politics and Methods*. New York: CRC Pres, pp. 43–62.

19. Jung, K., Park, S.J., Wu, W.-N. & Park, H.W. (2015). A Webometric Approach to Policy Analysis and Management Using Exponential Random Graph Models. *Quality & Quantity*, vol. 49, no 2, pp. 581–598.
20. Kim, G.-H., Trimi, S. & Chung, J.-H. (2014). Big-Data Applications in the Government Sector. *Communications of the ACM*, vol. 57, no 3, pp. 78–85.
21. Lane, J. (2016). Big Data: The Role of Education and Training. *Journal of Policy Analysis & Management*, vol. 35, no 3, pp. 722–724.
22. Lawrence, J. & Delaney, C.P. (2013). Integrating Hospital Administrative Data to Improve Health Care Efficiency and Outcomes : “The Socrates Story”. *Clinics in Colon and Rectal Surgery*, vol. 26, no 1, pp. 56–62.
23. Malomo, F. & Sena, V. (2017). Data Intelligence for Local Government? Assessing the Benefits and Barriers to Use of Big Data in the Public Sector. *Policy and Internet*, vol. 9, no 1, pp. 7–27.
24. Marais, L. & Matebesi, Z. (2013). Evidence-Based Policy Development in South Africa: the Case of Provincial Growth and Development Strategies. *Urban Forum*, vol. 24, no 3, pp. 357–371.
25. Marchi, G., Lucertini, G. & Tsoukiàs, A. (2016). From Evidence-Based Policy Making to Policy Analytics. *Annals of Operations Research*, vol. 236, no 1, pp. 15–38.
26. Menezes, P. & D’Aguiar, L. (2013). *Impact and Benefits of Micro-Databases’ Integration on the Statistics of the Banco De Portugal*. Proceedings of the 59th World Statistics Congress, Hong-Kong.
27. Metha, N. (2012). *Knight \$440 Million Loss Sealed by Rules on Canceling Trades*. Available at: www.bloomberg.com/news/2012-08-14/knight-440-million-loss-sealed-bynew-rules-on-canceling-trades.html (accessed: 17 January, 2018).
28. Mkrtchyan, N. V. (2012). Problemy ucheta naseleniya otdel’nykh vozrastnykh grupp v khode perepisi naseleniya 2010 g.: Prichiny otkloneniya poluchennykh dannykh ot ozhidaemykh [Problems of Accounting for the Population of Certain Age Groups During the 2010 Population Census: The Reasons for the Received Data Deviations From the Expected]. In: *Demograficheskie aspekty sotsial’no-ekonomicheskogo razvitiya* [Demographic Aspects of Socio-Economic Development]. Moscow: MAKS Press, pp. 197–214.
29. Nordbotten, S. (2010). The Use of Administrative Data in Official Statistics - Past, Present And Future: With Special Reference to the Nordic Countries. In: Michael Carlson, Hans Nyquist, Mattias Villani (red.), *Official statistics: Methodology and Applications in Honour of Daniel Thorburn*, pp. 205–223, Stockholm.
30. OECD (2015). Data-Driven Innovation. Paris: OECD.
31. OECD (2017). Digital Economy Outlook 2017. Paris: OECD.
32. Panagiotopoulos, P., Bowen, F. & Brooker, P. (2017). The Value of Social Media Data: Integrating Crowd Capabilities in Evidence-Based Policy. *Government Information Quarterly*, no 34, pp. 601–612.
33. Raizberg, B.A. (2008). Gosplan i gosudarstvennoe strategicheskoe planirovanie [Gosplan and State Strategic Planning]. *Economics of Contemporary Russia*, no 1, pp. 116–124.
34. Reimbsbach-Kounatze, C. (2015). *The Proliferation of “Big Data” and Implications for Official Statistics and Statistical Agencies: A Preliminary Analysis*. Working paper OECD Digital Economy Papers, no 245, Paris: OECD.

35. *Research Data Centre* (2018). Wir über uns. Available at: http://fdz.iab.de/de/FDZ_Scope_of_Services.aspx (accessed: 4 January, 2018)
36. Rieder, G. & Simon, J. (2017). Big Data: A New Empiricism and its Epistemic and Socio-Political Consequences. In: Pietsch W., Wernecke J., Ott M. (eds) *Berechenbarkeit der Welt?*. Springer VS, Wiesbaden. Available at: DOI: https://doi.org/10.1007/978-3-658-12153-2_4 (accessed: 08 May, 2018).
37. Rosenstock, L. & Lee, L.J. (2002). Attacks on Science: The Risks to Evidence-Based Policy. *American Journal of Public Health*, vol. 92, no 1, pp. 14–18.
38. Sabatier, P.A. (1991). Toward Better Theories of the Policy Process. *Political Science*, vol. 24, no 2, pp. 147–156.
39. Scott, C. (2005). *Measuring Up to the Measurement Problem: The Role of Statistics in Evidence-Based Policymaking*. Proceedings of the 2005 CBMS Network Meeting, pp. 35–93.
40. Severo, M., Feredj, A. & Romele, A. (2016). Soft Data and Public Policy: Can Social Media Offer Alternatives to Official Statistics in Urban Policymaking? *Policy and Internet*, vol. 8, no 3, pp. 354–372.
41. Sherr, L., Mueller, J. & Fox, Z. (2009). Abandoned Babies in the UK – A Review Utilizing Media Reports. *Child: Care, Health & Development*, vol. 35, no 3, pp. 419–430.
42. Simonet, D. & Alkafaji, Y. (2017). Critical Evaluations of the French Health Care Accounting Indicators: the Use of DRGs. *Public Administration Quarterly*, vol. 41, no 3, pp. 569–609.
43. Smith, T.E., Appel, A., Essock, S.M., Jackson, C.T., Karpati, A., Marsik, T., Myers, R.W., Tom, L. & Sederer, L.I. (2011). Public-Academic Partnerships. Using Medicaid Claims Data to Identify Service Gaps for High-Need Clients: The NYC Mental Health Care Monitoring Initiative. *Psychiatric Services*, vol. 62, no 1, pp. 9–11.
44. Solov'ev, A.K. & Dontsova, S.A. (2015). Metodologicheskie problemy pensionnoi statistiki [Methodological Problems of Pension Statistics]. *Narodonaselenie*, no 4, pp. 79–92.
45. *Statistics Canada* (2017). Use of administrative data. Available at: <https://www.statcan.gc.ca/pub/12-539-x/2009001/administrative-administratives-eng.htm> (accessed: 4 January, 2018)
46. *Statistics Sweden* (2007). Official statistics of Sweden. Available at: https://millenniumindicators.un.org/unsd/statcom/statcom_seminar/sweden.pdf (accessed: 4 January, 2018)
47. Stein, B.D., Pangilinan, M., Sorbero, M.J., Marcus, S.M., Donahue, S.A., Xu, Y., Smith, T.E. & Essock, S.M. (2014). Using Claims Data to Generate Clinical Flags Predicting Short-Term Risk of Continued Psychiatric Hospitalizations. *Psychiatric Services*, vol. 65, no 11, pp. 1341–1346.
48. Tambuwala, N., Rajabifard, A., Bennett, R., Williamson, I. & Wallace, J. (2012). Inter-Governmental Land Information Asymmetries in Australia. *Journal of Spatial Science*, vol. 57, no 1, pp. 83–100.
49. *United Nations* (2018). GWG for Big Data. Available at: <https://unstats.un.org/bigdata/> (accessed: 4 January, 2018)
50. Webster, D. (2002). Unemployment: How Official Statistics Distort Analysis and Policy, and Why. *Radical Statistics*, no 79, pp. 96–127.