

Академия ИКТ для лидеров Государственного управления

Модуль по
Осуществлению государственного управления на основе данных



UNITED NATIONS
APCICT - ESCAP

Академия ИКТ для лидеров государственного управления

**Осуществление государственного управления
на основе данных**

Академия ИКТ для лидеров государственного управления

Осуществление государственного управления на основе данных

Данная работа размещена в открытом доступе в соответствии с лицензией Creative Commons, созданной для межправительственных организаций, по адресу: <http://creativecommons.org/licenses/by/3.0/igo/>

Издатели должны удалить эмблему Организации Объединенных Наций из своего издания и создать новый дизайн обложки. Переводы должны иметь следующие оговорки: «Настоящая работа является неофициальным переводом, за которую издатель несет полную ответственность». Издателям необходимо направить файл своего издания на электронный адрес: apcict@un.org

Фотокопии и репродукции выдержек разрешены с соответствующими выражениями признательности.

Отказ от ответственности: взгляды, выраженные в настоящем документе, принадлежат авторам и не обязательно отражают взгляды Организации Объединенных Наций. Данная публикация была издана без официального редактирования, и используемые обозначения и представленные материалы не подразумевают выражения какого-либо мнения со стороны Секретариата Организации Объединенных Наций относительно статуса какой-либо страны, территории, города, района, или их администрации, или относительно делимитации каких-либо границ или разграничений.

Упоминание фирменных наименований и коммерческих продуктов не означает одобрения со стороны Организации Объединенных Наций.

Все вопросы по поводу данной работы должны быть адресованы на электронную почту: apcict@un.org

Контактные данные:

Asian and Pacific Training Centre for Information and Communication Technology for Development (APCICT/ESCAP)
5th Floor G-Tower, 175 Art Center Daero, Yeonsu-gu, Incheon, Republic of Korea
Tel +82 32 458 6650
Fax +82 32 458 6691/2

Email apcict@un.org
<http://www.unapcict.org>

Copyright © United Nations 2019

Все права защищены

О СЕРИИ УЧЕБНЫХ МОДУЛЕЙ

В нынешний «информационный век» легкий доступ к информации меняет наш образ жизни, работы и развлечений. «Цифровая экономика», также известная как «экономика знаний», «сетевая экономика» или «новая экономика», характеризуется переходом от производства товаров к созданию идей. Это подчеркивает растущую, если не центральную роль, которую играют информационные и коммуникационные технологии (ИКТ) в экономике, в частности, и в обществе, в целом.

Как следствие, правительства во всем мире все больше внимания уделяют ИКТ в целях развития (ИКТР). Для таких правительств ИКТР – это не только развитие отрасли ИКТ или сектора экономики, но также и использование ИКТ для стимулирования экономического роста, а также социального и политического развития.

Однако незнание быстро меняющегося технологического ландшафта и компетенций, необходимых для использования ИКТ для развития страны, вызывают трудности, с которыми сталкиваются правительства при разработке политики в области ИКТ. Поскольку нельзя регулировать то, чего не понимаешь, многие разработчики политики уклоняются от создания политики в области ИКТ. Но отдавать политику ИКТ на разработку «технарям» также неправильно, потому что зачастую «технари» не знают о социальных и политических последствиях технологий, которые они разрабатывают и используют.

Серия модулей Академии ИКТ для государственных лидеров была разработана Азиатско-тихоокеанским учебным центром информационно-коммуникационных технологий в целях развития (АТУЦ ИКТР) для:

1. Разработчиков политики на национальном и местном правительственном уровне, отвечающих за разработку политики в области ИКТ;
2. Государственных должностных лиц, ответственных за разработку и внедрение приложений на основе ИКТ;
3. Менеджеров в государственном секторе, стремящихся использовать инструменты ИКТ для управления проектами.

Целью данной серии модулей является рассмотрение основных вопросов, связанных с ИКТР как с политической, так и с технологической точек зрения. Мы не собираемся разрабатывать техническое руководство по ИКТ. Скорее, цель состоит в том, чтобы дать хорошее понимание того, на что способны современные цифровые технологии, куда они развиваются, и что это означает для разработки политики. Темы, охватываемые модулями, были определены путем анализа потребностей в обучении и обзора других учебных материалов по всему миру.

Модули разработаны таким образом, что они могут быть использованы отдельными лицами для самостоятельного изучения или в качестве материала для использования в учебном курсе или программе. Модули сами по себе самостоятельны, но в то же время

связаны между собой, и в каждом модуле были предприняты усилия для связи с темами и обсуждениями в других модулях серии. Долгосрочной целью является объединение модулей в цельный курс, который может пройти соответствующую сертификацию.

Каждый модуль начинается с формулировки целей модуля и целевых результатов обучения, по которым читатели могут оценить свой прогресс. Содержание модуля разделено на разделы, которые включают в себя тематические исследования и упражнения для углубленного изучения ключевых концепций. Упражнения могут выполняться как отдельными читателями, так группами участников тренинга. Рисунки и таблицы приведены для иллюстрации конкретных аспектов обсуждения. Также вниманию читателей представлены ссылки на литературные источники и Интернет-ресурсы, чтобы предоставить возможность получения дополнительной информации и знаний.

Использование ИКТР настолько разнообразно, что иногда тематические исследования и примеры внутри и между модулями могут показаться противоречивыми. Этого следует ожидать. Поскольку страны используют потенциал ИКТ в качестве инструментов для развития, то данная дисциплина сопровождается вдохновением, сложностями и перспективами.

Поддержка серии модулей Академии в печатном формате осуществляется на платформе дистанционного обучения в сети – Виртуальной Академией АТУЦ ИКТР (AVA – <http://www.unapcict.org/academy>) — в которой применяются виртуальные классы, показывающие выступления преподавателей в видео формате и презентации PowerPoint учебных модулей.

Кроме того, АТУЦ ИКТР разработал электронный хаб по вопросам ИКТР (E-co Hub - <http://www.unapcict.org/ecohub>), специализированный онлайн-сайт для практиков и разработчиков политики в области ИКТР, призванный расширить их опыт в области обучения и профессиональной подготовки. E-co Hub предоставляет доступ к ресурсам знаний по различным аспектам ИКТР и предоставляет пространство для обмена знаниями и опытом по развитию ИКТР.

Осуществление государственного управления на основе данных

Правительства продолжают испытывать трудности с использованием данных для улучшения управления. Даже в разгар происходящей революции данных некоторые из них сталкиваются с традиционными проблемами отсутствия данных, данных низкого качества и устаревших данных. Другие сталкиваются с более современной проблемой придания смысла множеству данных, которые им доступны. Как мы можем улучшить использование и анализ данных для поддержки управления и принятия решений на уровне правительства?

Данный модуль, предназначенный для государственных служащих в развивающихся странах, призван помочь им понять управление на основе данных и использовать управление на основе данных в своих юрисдикциях. Модуль состоит из семи разделов.

В Разделе 1 обсуждается революция данных, ее причины и социальные последствия.

В Разделе 2 рассматривается управление на основе данных.

Раздел 3 описывает традиционные и современные источники данных.

Раздел 4 посвящен управлению данными и вопросам администрирования данных.

Раздел 5 углубляется в аналитику данных, необъективность и интуицию.

В разделе 6 рассматривается благоприятная среда для управления данными.

Раздел 7 посвящен культуре данных в государственном секторе.

ЦЕЛИ МОДУЛЯ

Настоящий модуль преследует следующие цели:

1. Обсуждение социальных последствий революции данных в контексте управления данными;
2. Изучение управления, основанного на данных, путем рассмотрения решений, принятых на основе данных, научно-обоснованной политики, администрирования, ориентированного на конкретные результаты, и задач в области мониторинга и реализации целей Организации Объединенных Наций в области устойчивого развития;
3. Описание традиционных и современных источников данных;
4. Изучение роли управления и администрирования данными в обеспечении доступности, надежности, качества и своевременности данных для пользователей;
5. Обсуждение вопросов аналитики данных – процесса изучения данных для вывода заключений, опасность алгоритмической необъективности и роль интуиции в принятии решений;
6. Изучение политики и других вопросов, связанных с разработкой и внедрением управления на основе данных в развивающихся странах;
7. Рассмотрение формирования культуры данных в государственном секторе.

ИТОГИ ОБУЧЕНИЯ

После завершения изучения модуля читатели должны уметь:

1. Обсуждать социальные последствия революции данных;
2. Понимать концепцию управления на основе данных;
3. Оценивать и объяснять прогрессивную роль управления данными, администрирования данными и анализа данных в управлении на основе данных;
4. Определять и реализовывать ключевые стимулирующие политики, программы и мероприятия, способствующих управлению на основе данных;
5. Оценивать важность культуры данных в государственном секторе.

БЛАГОДАРНОСТИ

Модуль был подготовлен Эммануэлем К. Лалланой под общим руководством Киюнга Ко, директора Азиатско-тихоокеанского учебного центра по информационно-коммуникационным технологиям в целях развития (АТУЦ ИКТР).

Важные комментарии были предоставлены партнерами АТУЦ ИКТР, которые участвовали в нескольких обзорных мероприятиях. В модуле использовались комментарии во время совещания Группы экспертов по управлению на основе данных, которое состоялось 6-7 декабря 2017 года в Маниле, и консультативного совещания, состоявшегося 29-31 августа 2018 года в Бангкоке. Ценные предложения были также получены от Тицианы Бонапаче, Ацуко Окуда, Сиопе Вакатаки Офа и Мэтью Перкинса из подразделения ЭСКАТО по вопросам ИКТ и сокращению рисков стихийных бедствий, а также от Джеммы Ван Халдерен и Рикке Мунк Хансена из Статистического отдела ЭСКАТО.

Разработка модуля координировалась Робертом де Хесусом и Nuankaе Wongthawatchai. Редакторскую поддержку оказала Кристина Апикуль. Byeongjo Kong, Кевин Друин, Майкл Сантьяго и Юньцзинь Ли помогли вычитать учебный модуль. Joo-Eun Chung и Ho-Din Ligaу оказали административную поддержку и другую необходимую помощь для выпуска данного модуля.

СОДЕРЖАНИЕ

О СЕРИИ УЧЕБНЫХ МОДУЛЕЙ	4
Осуществление государственного управления на основе данных	6
ЦЕЛИ МОДУЛЯ	7
ИТОГИ ОБУЧЕНИЯ	8
БЛАГОДАРНОСТИ	9
СОДЕРЖАНИЕ	10
Список вставок	12
Список таблиц	13
Сокращения	13
1. Революция данных и справедливость данных	14
1.1 Данные и технологии	15
1.2 От оцифровки до датафикации	18
1.3 Справедливость на основе данных	25
2. Данные и управление	30
2.1 Управление	30
2.2 Управление на основе данных	33
2.2.1 Принятие решений на основе данных	35
2.2.2 Разработка политики на основе фактов	39
2.2.3 Администрирование, ориентированное на результаты	42
2.3 Проблемы в управлении на основе данных: пример целей ООН в области устойчивого развития	45
3. Мир переполнен данными	49
3.1 Официальная статистика	49
3.2 Большие данные	55
3.3 Данные в режиме реального времени	61
3.4 Малые данные	63
3.5 Данные, генерируемые гражданами	65
4. Управление и администрирование данными	68
4.1 Управление данными	68
4.2 Администрирование данными	72
4.2.1 Архитектура данных	74
4.2.2 Моделирование и дизайн данных	75
4.2.3 Хранение и операции с данными	77
4.2.4 Безопасность данных	78
4.2.5 Интеграция и совместимость данных	80
4.2.6 Управление документацией и контентом	82
4.2.7 Справочные и базовые данные	84
4.2.8 Хранение данных и бизнес-аналитика	86
4.2.9 Управление метаданными	88
4.2.10 Управление качеством данных	89
4.2.11 Оценка прогресса	91
5. Аналитика данных, алгоритмическая необъективность и интуиция	93

5.1 Аналитика данных	93
5.1.1 Описательная аналитика.....	93
5.1.2 Диагностическая аналитика	94
5.1.3 Прогностическая аналитика	96
5.1.4 Предписывающая аналитика	97
5.1.5 Аналитика данных в правительстве	98
5.2 Алгоритмическая необъективность.....	100
5.3 Интуиция и принятие решений.....	105
6. Благоприятная политика.....	108
6.1 Укрепление процесса сбора данных.....	108
6.1.1 Совершенствование национальной системы статистики	108
6.1.2 Улучшение гендерной статистики	112
6.1.3 Стимулирование генерации данных гражданами	114
6.2 Институционализация управления данными.....	116
6.3 Повышение общественного доступа и доверия к данным.....	119
6.3.1 Открытие государственных данных.....	119
6.3.2 Укрепление конфиденциальности данных	125
6.4 Готовимся к большим данным	130
7. Создание культуры данных в государственном секторе	133
ЗАМЕТКИ ДЛЯ ИНСТРУКТОРА.....	136

Список вставок

- Вставка 1. Зетабайт в числах
- Вставка 2. Интернет вещей и водная безопасность
- Вставка 3. Пути, которыми мы оцифровали наш мир
- Вставка 4. Запись протокола встречи
- Вставка 5. Преимущества цифровизации для правительства
- Вставка 6. Последствия цифровизации
- Вставка 7. Цифровая трансформация и технологии
- Вставка 8. Датафикация повседневной жизни
- Вставка 9. Как данные могут влиять на правительство, ориентированное на интересы граждан: случай Австралии
- Вставка 10. Основные принципы официальной статистики
- Вставка 11. Проблемы количественной оценки женского труда
- Вставка 12. Большие данные в бизнесе
- Вставка 13. Как Таиланд использует большие данные для власти правительства
- Вставка 14. Аналитика данных в режиме реального времени в правительстве
- Вставка 15. Почему малые данные?
- Вставка 16. Картографирование случаев сексуальных домогательств в Египте
- Вставка 17. Satu Data Индонезия
- Вставка 18. Шесть принципов современной архитектуры данных
- Вставка 19. Что такое моделирование данных?
- Вставка 20. Типы данных для хранения
- Вставка 21. Предотвращение угроз
- Вставка 22. Совместимость данных и стихийные бедствия
- Вставка 23. Управление документацией среди администраций американских штатов
- Вставка 24. Что такое базовые данные и справочные данные?
- Вставка 25. Что такое хранилище данных?
- Вставка 26. Основы метаданных
- Вставка 27. Шесть составляющих качества данных
- Вставка 28. Описательная аналитика и бизнес-аналитика
- Вставка 29. Как выполнять диагностическую аналитику
- Вставка 30. Розничная торговля использует прогностическую аналитику для обслуживания беременных женщин
- Вставка 31. Совершенствование сектора здравоохранения с помощью предписывающей аналитики
- Вставка 32. Аналитика и наблюдение за заболеваниями
- Вставка 33. Алгоритмическая необъективность
- Вставка 34. Принципы подотчетности алгоритмов: справедливость, подотчетность и прозрачность в машинном обучении

Список таблиц

Таблица 1. Влияние датафикации на отрасли экономики

Таблица 2. Проблемы управления на местном, национальном и глобальном уровнях

Сокращения

APCICT	Asian and Pacific Training Centre for Information and Communication Technology for Development
APEC	Asia-Pacific Economic Cooperation
API	Application Programming Interface
CD	Compact Disc
CEO	Chief Executive Officer
COO	Chief Operating Officer
DAMA-DMBOK 2	Data Management Association International's Guide to the Data Management Body of Knowledge, Second Edition
ESCAP	Economic and Social Commission for Asia and the Pacific (United Nations)
EU	European Union
GDPR	General Data Protection Regulation
GPHIN	Global Public Health Intelligence Network
GPS	Global Positioning System
HCFAC	Health Care Fraud and Abuse Control
HHS-OIG	Office of Inspector General for the Department of Health and Human Services (United States of America)
ICT	Information and Communication Technology
IoT	Internet of Things
MfDR	Monitoring for Development Results
OECD	Organisation for Economic Co-operation and Development
SDG	Sustainable Development Goal
UNDG	United Nations Development Group
UNDP	United Nations Development Programme
UNESCO	United Nations Educational, Scientific and Cultural Organization

1. Революция данных и справедливость данных

Задачи данного раздела:

- Описать информационно-коммуникационные технологии (ИКТ), ставшие основой революции данных;
 - Объяснить оцифровку, оцифровку, цифровое преобразование и передачу данных;
 - Обсудить понятие справедливости данных.
-

Данные – «факты и статистика, собранные вместе для справочного материала или анализа»¹ - это новая нефть. Согласно Economist:

Данные для этого столетия – то, чем нефть была до последнего: драйвер роста и перемен. Потоки данных создали новую инфраструктуру, новые предприятия, новые монополии, новую политику и, что особенно важно, новую экономику. ... Будет много сражений за то, кто должен обладать данными и извлекать из них выгоду.²

В развивающихся странах более широкий доступ и использование данных могут помочь в:

- **Улучшении управления** за счет повышения прозрачности и подотчетности, повышения эффективности предоставления услуг и расширения обмена информацией в государственных ведомствах;
- **Расширении прав и возможностей граждан** путем улучшения их способности принимать решения и расширения их выбора, а также в качестве катализатора социальной мобилизации;
- **Создании экономических возможностей**, предоставляя возможности для построения бизнеса, создания рабочих мест и новых форм инноваций, а в более общем плане - стимулируя экономический рост;
- **Решении сложных общественных проблем**, улучшая ситуационную осведомленность, предлагая более широкий спектр знаний и опыта для решения общественных проблем и предоставляя политикам, гражданскому обществу и гражданам возможности лучшего реагирования на вмешательства и отслеживания воздействий.³

1 Oxford Living Dictionaries, "Data". Доступно по <https://en.oxforddictionaries.com/definition/data> (accessed on 8 January 2019).

2 The Economist, "Data is giving rise to a new economy", 6 May 2017. Доступно по <https://www.economist.com/briefing/2017/05/06/data-is-giving-rise-to-a-new-economy>.

3 Stefaan Verhulst and Andrew Young, "The Evidence that Open Government Data Improves Developing Economies", *ICTworks*, 18 June 2018. Доступно по <https://www.ictworks.org/the-evidence-that-open-government-data-improves-developing-economies/#.W43755Mzau4>.

Всемирный банк полагает, что «как и нефть, необработанные данные имеют относительно небольшую ценность и должны быть собраны, переработаны, сохранены и проданы для создания стоимости».⁴ Таким образом, современная задача для правительств состоит в том, чтобы «извлекать ценность из данных для улучшения предоставления услуг так же, как частные компании научились делать это с целью получения прибыли».⁵

Но в ключевых моментах данные не похожи на нефть.⁶ Нефть - это ограниченный ресурс, а данные «эффективно бесконечно долговечны и могут использоваться повторно».⁷ Добывать нефть дорого и сложно, в то время как получение данных становится все дешевле и проще. Перемещение нефти обходится дорого, в то время как данные «могут бесконечно копироваться и перемещаться по всему миру со скоростью света, с очень низкими затратами».⁸

1.1 Данные и технологии

Нет сомнений в том, что мы находимся на пороге революции данных: «Взрыв объема данных, скорость, с которой производятся данные, количество производителей данных, распространение данных и количество предметов, на которых существуют данные».⁹

В 2002 году исследователи из Калифорнийского университета в Беркли подсчитали, что мир произвел «5 экзабайт новой информации».¹⁰ Это эквивалентно 37 000 новым библиотекам с более чем 162 миллионами книг, рукописей, карт, фотографий и т. д. (это примерно размер самой большой в мире библиотеки - Библиотеки Конгресса США). Пять экзабайт – это также вдвое больше информации, произведенной в 1999 году.

В 2013 году во всем мире было создано 4,4 зетабайта данных - «столько же бит, сколько звезд в физической вселенной».¹¹ В том же году было подсчитано, что к 2020 году объем данных вырастет до 44 зетабайт. Также прогнозировалось, что в 2017 году развивающиеся страны обгонят развитые страны в создании данных.

4 World Bank, *Information and Communications for Development 2018: Data-Driven Development* (Washington D.C., 2019), p. 1. Доступно по <https://www.worldbank.org/en/topic/digitaldevelopment/publication/data-driven-development.print>.

5 Там же

6 Amol Rajan, “Data is not the new oil”, *BBC*, 9 October 2017. Доступно по <https://www.bbc.com/news/entertainment-arts-41559076>.

7 Bernard Marr “Here’s Why Data Is Not The New Oil”, *Forbes*, 5 March 2018. Доступно по <https://www.forbes.com/sites/bernardmarr/2018/03/05/heres-why-data-is-not-the-new-oil/#59bc073f3aa9>.

8 Там же

9 United Nations Secretary-General’s Independent Expert Advisory Group on the Data Revolution for Sustainable Development, “A World that Counts: Mobilizing the Data Revolution for Sustainable Development”, November 2014, p. 6. Доступно по <http://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf>.

10 Regents of the University of California, “How Much Information? 2003: Executive Summary”, 27 October 2003. Доступно по <http://groups.ischool.berkeley.edu/archive/how-much-info-2003/execsum.htm>.

11 EMC Digital Universe, “The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things”, Info Brief, April 2014. Доступно по <https://www.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf>.

В 2017 году прогнозировалось, что «в 2025 году мир создаст и скопирует 163 зетабайта данных, что в десять раз больше, чем объем данных, созданных в 2016 году».¹²

Вставка 1. Зетабайт в числах

1 килобайт	1,000
1 мегабайт	1,000,000
1 гигабайт	1,000,000,000
1 терабайт	1,000,000,000,000
1 петабайт	1,000,000,000,000,000
1 эксабайт	1,000,000,000,000,000,000
1 зетабайт	1,000,000,000,000,000,000,000

Взято из: Carascojames, "Era Zeta Bytes: How many bytes in Zeta bytes, how will it affect us?" steemKR, <https://steemkr.com/technology/@carascojames/era-zeta-bytes-how-many-bytes-in-zeta-bytes-how-will-it-affect-us> (accessed on 8 January 2019).

Наблюдатели частично связывают революцию в области данных с технологией – увеличением вычислительной мощности, ускорением широкополосных соединений, недорогими датчиками и повсеместным распространением мобильных телефонов.¹³ Развитие систем хранения (таких как облачные вычисления) и систем баз данных также позволило собирать, хранить и обрабатывать очень большие объемы данных.

Широкое использование технологий позволило нам эволюционировать от потребителей информации к производителям информации.¹⁴ Мы привыкли смотреть телевизор, слушать радио и читать газеты. В 2016 году **каждую минуту** мы твитнули 456 000 раз, опубликовали 46 740 фотографий в Instagram, провели поиск в Гугле 3,6 млн. раз, отправили 103 447 520 электронных писем со спамом и публиковали 600 новых правок в Википедии.¹⁵

Как производители информации, мы создаем **цифровые следы (footprints)** - сумму всех данных, которые мы производим в результате нашей онлайн-деятельности.¹⁶ Это включает в себя данные, которые мы создаем, когда мы используем социальные сети (такие как Facebook и Twitter), а также когда мы ищем в Google.

В то же время мы совершаем **производство (exhaust) данных** - побочный продукт

12 David Reinsel, John Gantz and John Rydning, "Data Age 2025: The Evolution of Data to Life-Critical", IDC White Paper, April 2017.

13 United Nations Global Pulse, "Big Data for Development: Challenges and Opportunities", May 2012, p. 9. Доступно по <http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseMay2012.pdf>.

14 Kirsty, "Where Did the 'Data Explosion' Come From?" *Bime Blog*. Доступно по <https://blog.bimeanalytics.com/english/where-did-the-data-explosion-come-from> (accessed on 8 January 2019).

15 Tom Hale, "How Much Data Does the World Generate Every Minute?" *IFL Science*, 26 July 2017. Доступно по <http://www.iflscience.com/technology/how-much-data-does-the-world-generate-every-minute/>.

16 Margaret Rouse, "Data Exhaust", *WhatIs.com*, April 2015. Доступно по <https://whatIs.techtarget.com/definition/data-exhaust>.

нашей онлайн-деятельности.¹⁷ Например, когда мы звоним с помощью наших мобильных телефонов, создается журнал вызовов, который состоит из нашего номера телефона и номера абонента, которому мы звоним, времени и продолжительности вызова, а также информации о взаимодействии нашего устройства с сотовыми вышками.¹⁸ Мы также совершаем производство данных, когда мы используем наш веб-браузер для поиска или онлайн-покупки. Наши произведенные данные предоставляют значительную информацию о нашем поведении в Интернете, и при обработке могут привести к ценной информации о наших привычках и предпочтениях вне сети.

Хотя мы и стали производителями информации, мы больше не являемся единственными производителями данных. Наши устройства, подключенные к Интернету, также создают данные.¹⁹ Это развитие называется Интернетом вещей (IoT): «Система взаимосвязанных вычислительных устройств, механических и цифровых машин, объектов, животных или людей, которым предоставляются уникальные идентификаторы, и возможность передачи данных по сети без необходимости взаимодействия человека с человеком или человека с компьютером».²⁰ Если кратко, то «Интернет-связь между повседневными объектами».²¹

В 2017 году более 20 миллиардов устройств были подключены к Интернету.²² Ожидается, что это число вырастет почти до 31 миллиарда в 2020 году, а затем до 75 миллиардов устройств к 2025 году. В результате к 2025 году «IoT будет генерировать более 2 зеттабайтов данных, в основном, генерируемых устройствами бытовой электроники».²³

Люди будут гипер-связаны Интернетом вещей. Прогнозируется, что к 2025 году «среднестатистический подключенный человек в любой точке мира будет взаимодействовать с подключенными устройствами почти 4800 раз в день или по **одному взаимодействию каждые 18 секунд**».²⁴

IoT не только будет предоставлять важную информацию об индивидуальном поведении, но и поможет решать проблемы общества.

17 Там же

18 Jessica Leber, "Mobile Call Logs Can Reveal a Lot to the NSA", *MIT Technology Review*, 18 June 2013. Доступно по <https://www.technologyreview.com/s/516181/mobile-call-logs-can-reveal-a-lot-to-the-nsa/>.

19 Steve Ranger, "What is the IoT? Everything you need to know about the Internet of Things right now", *ZDNet*, 21 August 2018. Доступно по <http://www.zdnet.com/article/what-is-the-internet-of-things-everything-you-need-to-know-about-the-iot-right-now/>.

20 Margaret Rouse, "Internet of Things (IoT)", *TechTarget IoT Agenda*, June 2018. Доступно по <https://internetofthingsagenda.techtarget.com/definition/Internet-of-Things-IoT>.

21 Hyea Won Lee, "Agriculture 2.0: how the Internet of Things can revolutionize the farming sector", *World Bank Information and Communications for Development Blog*, 17 August 2017. Доступно по <http://blogs.worldbank.org/ic4d/agriculture-20-how-internet-things-can-revolutionize-farming-sector>.

22 Statista, "Internet of Things (IoT) connected devices installed base worldwide from 2015 to 2025 (in billions)". Доступно по <https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/> (accessed on 8 January 2019).

23 Ian Scales, "IoT 2025: 27 billion devices spewing 2 zettabytes of data and generating \$3 trillion", *TelecomTV*, 5 August 2016. Доступно по <http://www.telecomtv.com/articles/iot/iot-2025-27-billion-devices-spewing-2-zettabytes-of-data-and-generating-3-trillion-13872/>.

24 Andrew Cave, "What Will We Do When the World's Data Hits 163 Zettabytes in 2025?" *Forbes*, 13 April 2017. Доступно по <https://www.forbes.com/sites/andrewcave/2017/04/13/what-will-we-do-when-the-worlds-data-hits-163-zettabytes-in-2025/#550efad8349a>.

Вставка 2. Интернет вещей и водная безопасность

Технологии IoT могут обеспечить лучшее понимание сложных проблем, связанных с водной безопасностью, предоставляя возможность правительствам лучше определять приоритеты по вопросам водоснабжения, потребительского спроса и управления. Кроме того, приложения IoT могут помочь государственным органам лучше координировать практические действия среди заинтересованных сторон, фиксируя конкретные воздействия каждой политики. Это может быть достигнуто с помощью прогнозных моделей или измерений в режиме реального времени, которые позволяют проводить A/B-тестирование (то есть сравнение двух версий A и B).

Увеличение водоснабжения часто является первым вариантом, рассматриваемым в качестве сокращения запасов воды, и традиционно компании вкладывают значительные средства в поиск новых источников воды. По мере того как новые источники иссякают, коммунальные предприятия могут вместо этого сосредоточиться на повышении качества доставки – поскольку инфраструктура часто устаревшая, и системы водоснабжения теряют в среднем 16 процентов воды во время доставки. Одна из проблем, которую может решить IoT, заключается в том, чтобы точно определить, где проводить ремонт, чтобы улучшить доходность, и будет ли объем, сэкономленный для этой области, компенсировать капитальные затраты на ремонт. Датчики могут обеспечить более точное понимание потоков воды и помочь определить приоритеты улучшений, даже на уровне отдельных домовладельцев, которые обычно не связаны с состоянием водной инфраструктуры. Остановка или замедление утечек в домашних условиях, которые могут составить до 10 000 галлонов в год, могут еще больше повысить доходность от доставки очищенной воды. Такие продукты, как, например, LeakSmart, сочетают в себе простой датчик и практический механизм для определения повреждения трубы и необходимости перекрытия воды.

Взято (с изменениями) из: Max Meyers, Claire Niech and William D. Eggers, "Anticipate, sense, and respond: Connected government and the Internet of Things", Deloitte Insights, 28 August 2015. Доступно по <https://www2.deloitte.com/insights/us/en/focus/internet-of-things/iot-in-government.html>.

1.2 От оцифровки до датафикации

Понятие революции данных также может быть рассмотрена с точки зрения «оцифровки», «цифровизации», «цифрового преобразования» и «датафикации».

Оцифровка (digitization) - это процесс преобразования аналогового формата в цифровой.²⁵ Наши телефоны раньше были аналоговыми, теперь они цифровые. Если раньше наши данные хранились на бумаге и в ящиках шкафов, то теперь мы храним их в двоичном коде (единицах и нулях) и в базах данных.²⁶

Вставка 3. Пути, которыми мы оцифровали наш мир

От обычной почты до электронных сообщений: Почта перешла на цифровой формат. Большинство компаний предлагают ту или иную форму онлайн-биллинга, поэтому клиенты

²⁵ Margaret Rouse, "Digitization", *WhatIs.com*, April 2007. Доступно по <http://whatIs.techtarget.com/definition/digitization>.

²⁶ Gil Press, "A Very Short History of Digitization", *Forbes*, 27 December 2015. Доступно по <https://www.forbes.com/sites/gilpress/2015/12/27/a-very-short-history-of-digitization/#bf522f349ac2>.

могут никогда не увидеть бумажный счет, и мы можем получить доступ к каталогам в онлайн, а не получать их в наших почтовых ящиках. От деловой переписки до писем друзей, большая часть нашей ежедневной почты теперь находится на наших компьютерах, планшетах или смартфонах.

От компакт-дисков к MP3-файлам: В основном, компакт-диски (CD) были заменены на MP3-файлы. Большинство людей загружают свою музыку для воспроизведения на свои телефоны или MP3-плееры, а не складывают огромную коллекцию компакт-дисков. Превращение музыки в медиа-файлы, а не в реальные записи становится нормой.

От дорожных карт к GPS: Когда вы в последний раз покупали карту? Теперь легко получить в свое распоряжение устройство глобальной системы позиционирования (GPS), такое как Garmin или TomTom, или просто использовать функции навигации на смартфонах. Или мы можем залезть в интернет, чтобы получить направления от Google. Все наши карты теперь доступны через компьютеры и смартфоны с быстро обновляемой информацией, а не через бумажные источники, которые быстро устаревают.

Взято (с изменениями) из: Jaymi Heimbuch, "7 major ways we're digitizing our world, and 3 reasons we still want hardcopies", Treehugger, 11 October 2010. Доступно по <https://www.treehugger.com/clean-technology/7-major-ways-were-digitizing-our-world-and-3-reasons-we-still-want-hardcopies.html>.

Цифровизация (digitalization) означает «превращение взаимодействий, коммуникаций, бизнес-функций и бизнес-моделей в (более) цифровые».²⁷

На практике, это означает улучшение бизнес-операций с использованием цифровых технологий.

Вставка 4. Запись протокола встречи

Представьте себе компанию, в которой проводятся утренние встречи топ-менеджмента. Поскольку встречи важны, все согласны с тем, что помощник главного исполнительного директора (СЕО) должен вести протокол. Тем не менее, участники разделились во мнениях, как они должны ознакамливаться с протоколами.

Генеральный директор предлагает, чтобы помощник написал протокол, а затем скопировал его. Генеральный директор утверждает, что это будет быстро и легко.

Главный исполнительный директор (СОО) предлагает, чтобы помощник набрал протокол, а затем отправил участникам по электронной почте копию в формате PDF. Потерять сообщение электронной почты сложнее, чем фотокопию, говорит СОО.

Главный технический директор предлагает создать веб-страницу утреннего собрания топ-менеджмента и шаблон в веб-приложении (SaaS - «ПО как услуга»), которое будет обслуживать данную потребность. Таким образом, каждый знает, где сохранены предыдущие протоколы,

²⁷ I-Scoop, "Digitization, digitalization and digital transformation: the differences". Доступно по <https://www.i-scoop.eu/digitization-digitalization-digital-transformation-disruption/> (accessed on 8 January 2019).

может легко получить доступ к ним и поделиться ими, может добавлять элементы на веб-страницу собраний и создавать ссылки на собрания.

Взято (с изменениями) из: David Burkett, "Digitisation And Digitalisation: What Means What?" WorkingMouse, 19 December 2017. Доступно по <https://workingmouse.com.au/innovation/digitisation-digitalisation-digital-transformation>.

Для правительства цифровизация может помочь повысить эффективность и прозрачность, а также улучшить оказание социальных услуг.

Опрос американских федеральных служащих, проведенный в 2018 году, показал, что «82% государственных учреждений планируют проекты по цифровизации».²⁸ Межведомственный обмен (34 процента), отслеживание пособий (31 процент) и заявок на получение пособий (29 процентов) являются одними из ключевых функций, которые будут переведены в цифровую форму.

Вставка 5. Преимущества цифровизации для правительства

Эффективность и экономия:

Цифровизация обычно приводит к повышению эффективности. Различные учреждения имеют разные побуждающие причины для цифровизации, например, доступность, сокращение расходов, возможность отслеживания или даже просто ради расширения самого процесса цифрового внедрения. Все это, в определенной степени, приведет к эффективности. В самом простом смысле, правительственное ведомство, которое хочет перейти на безбумажные процессы (переместив все свои печатные материалы в облако), сможет сэкономить и совместно использовать пространство с другими государственными учреждениями, а также уменьшить свой «углеродный след»²⁹ и объемы перевозок (за счет сокращения необходимости доставки и утилизации материалов). Кроме того, государственные служащие могут экономить время при поиске конкретных документов благодаря цифровой индексации, что может еще больше повысить производительность. В Эстонии, например, утверждает, что благодаря своей цифровой кампании она сэкономила 800 лет рабочего времени в год.

Улучшение предоставления социальных услуг:

Цифровизация может потенциально улучшить качество жизни граждан. Примером может служить продление водительских прав, которое может включать поездку в национальное транспортное агентство, заполнение форм и ожидание в длинных очередях. Весь процесс может занять несколько минут, а не часов, если государственное транспортное агентство использует цифровой подход. Многие процессы могут быть выполнены в онлайн.

Продвижение прозрачности:

Учитывая, как большинство правительств продвигают программы честности и прозрачности в госуправлении, цифровизация транзакций может помочь обеспечить лучшую наглядность и ясность. Коррупция может возникнуть во время кассовых операций между гражданами и

²⁸ Mathew Chase, "The state of data management in the public sector in 2018", *Experian*, 8 February 2018. Доступно по <https://www.edg.com/blog/the-state-of-data-management-in-the-public-sector-in-2018/>.

²⁹ Углеродный след – это «совокупность всех выбросов парниковых газов, произведенных прямо и косвенно отдельным человеком, организацией, мероприятием или продуктом»

государственными органами. Внедрение цифровых транзакций не только поможет государственным аудиторам контролировать денежные потоки, но также будет стимулировать граждан и государственные органы придерживаться этических норм. С помощью цифровых платформ можно эффективно отслеживать и контролировать каждую транзакцию, одновременно снижая бюрократию и коррупцию.

Взято (с изменениями) из: Christian Lauron and Irsyad Stamboel, "Digitalization of Government: Suits The C-Suite", BusinessWorld, 18 February 2018. Доступно по <http://www.bworldonline.com/digitalization-of-government/>.

Некоторые используют цифровизацию для описания более крупного явления - «интеграции цифровых технологий в повседневную жизнь».³⁰

Вставка 6. Последствия цифровизации

Процесс цифровизации предоставляет пользователям сети новые и инновационные способы увеличения и совместного использования своего социального, культурного и экономического капитала. Возможности подключения предлагают людям богатую гамму решений, основанных на коллективном интеллекте. Но цифровизация нашей повседневной жизни также становится все более агрессивной. Ученые и исследователи изучают темную сторону сети, чтобы выявить негативные внешние эффекты процесса цифровизации. На заре этой новой цифровой эпохи создается видение, связанное с обществом цифрового риска, в котором выделяются новые формы силы убеждения, риски социальной аномии, коллективные навязчивые идеи, антиутопические настроения и формы социального отчуждения из-за чрезмерной зависимости от цифровых инструментов.

Переработано из:

Michele Bonazzi, "For a Critical Theory of the Digitalization of Everyday Life", 14 July 2016. Доступно по <https://isaconf.confex.com/isaconf/forum2016/webprogram/Paper77619.html>.

Цифровая трансформация – это использование цифровых технологий во всех аспектах деятельности предприятия для фундаментального изменения способов создания и доставки добавленной стоимости. Это также «культурное изменение, которое требует от организаций постоянно бросать вызов существующему положению вещей, экспериментировать и быть готовым к неудачам».³¹

30 Heikki Otsolampi, "Digitalization – the first true revolution in business history". Доступно по <https://www.avaus.fi/en/blog/digitalization-the-first-true-revolution-in-business-history/> (accessed on 8 January 2019).

31 Enterprisers Project, "What is Digital Transformation?" Доступно по <https://enterpriseproject.com/what-is-digital-transformation> (accessed on 8 January 2019).

Вставка 7. Цифровая трансформация и технологии

Правительствам необходимо выйти за рамки оцифровки существующих процессов и услуг. Им необходимо использовать возможности цифровых данных и технологий для фундаментального переосмысления и преобразования бизнес-моделей управления.

Несмотря на то, что многие правительства мира достигли значительного прогресса, полный потенциал концепции цифрового правительства остается в значительной степени неиспользованным. Многие транзакционные и платежные услуги по-прежнему недоступны в онлайн. Существующие цифровые сервисы часто не оптимизированы для мобильных устройств. Функциональность и опыт использования онлайн-сервисов, разработанных и управляемых правительствами, обычно оставляют желать лучшего по сравнению с лучшими практиками коммерческих организаций.

Почему так происходит? Устаревшие системы мешают, законодательство и нормативно-правовые акты трудно изменить, а проблемы обеспечения безопасности и конфиденциальности являются сложными вопросами. Но, несмотря на эти проблемы, правительства признают потенциал социальных, мобильных, облачных и облачных технологий для стимулирования преобразований в государственном секторе.

Успешное цифровое преобразование требует: сильного лидерства на самом высоком уровне; инвестиции в науку, технологии, развитие инженерных и математических навыков; а также культурные и поведенческие изменения. Аналогичным образом, правительства должны максимизировать свои цифровые инвестиции путем стратегического использования новых технологий, а также данных и передовой аналитики для оптимизации политики, программ, платежных и информационных систем.

Взято (с изменениями) из: Boston Consulting Group, "Digital Transformation and Technology: How Governments are Upping their Game in Digital". Доступно по <https://www.bcg.com/industries/public-sector/digital-transformation-technology.aspx> (accessed on 8 January 2019).

Джейсон Блумберг предлагает хороший способ различать данные тенденции: «Мы оцифровываем информацию, мы оцифровываем процессы и роли, которые составляют бизнес-операции, и мы преобразуем в цифровой форме бизнес и его стратегию».³² В государственном секторе общегосударственные инициативы электронного правительства, такие как Gov 2.0, эквивалентны цифровой трансформации.

Датафикация (datafication) превращает ранее невидимый процесс/деятельность в данные, которые можно мониторить, отслеживать, анализировать и оптимизировать.³³ Под «datafy» понимается «превращение в данные многих аспектов окружающего мира,

32 Jason Bloomberg "Digitization, Digitalization, and Digital Transformation: Confuse Them At Your Peril", *Forbes*, 29 April 2018. Доступно по <https://www.forbes.com/sites/jasonbloomberg/2018/04/29/digitization-digitalization-and-digital-transformation-confuse-them-at-your-peril/#5c2560f82f2c>.

33 Margarita Shilova, "The Concept Of Datafication; Definition & Examples", *Apiumhub*, 15 June 2017. Доступно по <https://apiumhub.com/tech-blog-barcelona/datafication-examples/>.

которые никогда не были определены количественно».³⁴

Возьмите эмоции – гнев, страх, радость, изумление и т. д. До широкого использования социальных сетей эмоции было очень трудно отслеживать и анализировать. Теперь мы создаем цифровой след наших эмоций каждый раз, когда любим, смеемся или злимся в сообщении в Facebook. Как отмечают Майер-Шонбергер и Кукье: «Платформы социальных сетей не просто предлагают нам возможность находить и поддерживать связь с друзьями и коллегами, они берут нематериальные элементы нашей повседневной жизни и превращают их в данные, которые можно использовать для производства новых вещей».³⁵

Вставка 8. Датафикация повседневной жизни

Тимо Эллиот написал в своем блоге:

Мои упражнения теперь датафицируются (datafied – переводятся в данные). Этим утром я пробежался, и мое устройство Fitbit One записало, сколько именно я пробежал, сколько шагов я сделал и сколько калорий я сжег в процессе. Впервые мне очень легко отслеживать и контролировать ход моих упражнений.

И это только один маленький пример. Большая часть моей ежедневной активности теперь автоматически отслеживается. Моя сеть друзей теперь датафицируется в Facebook. Моя сеть профессиональных связей датафицируется в LinkedIn. Мое местоположение датафицируется в Foursquare. Мои последние случайные мысли датафицируются в Twitter. Мои музыкальные предпочтения хранятся в Spotify.

Даже чтение книг в настоящее время датафицируется. Пока я читаю на своем устройстве Kindle, оно на самом деле следит за мной. Amazon отслеживает данные моего чтения и использует их для предоставления полезных услуг. Например, он знает, на какой странице я нахожусь, поэтому я могу легко переключаться между различными устройствами. Он использует мою скорость чтения, чтобы оценить, сколько времени мне понадобится, чтобы закончить книгу. И Amazon учитывает некоторые аспекты «мудрости толпы» - например, я могу выбрать для чтения те отрывки, которые были выделены другими людьми как наиболее интересные.

Эти данные также собираются и анализируются Amazon для оптимизации продаж книг. Например, когда я недавно закончил книгу в серии Кена Фоллета, на следующее утро я получил электронное письмо со специальным предложением для следующей книги из серии (что интересно, по цене, которая была выше текущей «нормальной» цены...).

Взято (с изменениями) из: Timo Elliott, "The Datification Of Our Daily Lives", Digital Business & Business Analytics, 9 July 2013. Доступно по <https://timoelliott.com/blog/2013/07/the-datification-of-our-daily-lives.html>.

34 Kenneth Cukier and Viktor Mayer-Schoenberger, "The Rise of Big Data: How It's Changing the Way We Think About the World", *Foreign Affairs*, vol. 92, no. 3 (May/June 2013), p. 29. Доступно по <https://www.foreignaffairs.com/system/files/pdf/articles/2013/92305.pdf>.

35 Viktor Mayer-Schoenberger and Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work and Think* (London, John Murray, 2013) p. 91.

Датафикация означает, что теперь все элементы могут быть проанализированы на наличие паттернов и корреляций. В случае оцифрованной книги можно анализировать не только слова или фразы, но и предложения и абзацы.

Датафикация уже меняет бизнес.³⁶ Отделы кадров используют данные с мобильных телефонов, приложений или социальных сетей для определения конкретных характеристик потенциальных сотрудников, таких как профиль принятия рискованных решений и особенности характера. Датафикация сделает тесты на проверку личности или аналитическое мышление устаревшими. Это также позволит предложить более персонализированные продукты и услуги.

Кроме того, датафикация оптимизирует и улучшает бизнес-процессы (набор задач и действий по оказанию услуги или доставке продукта клиенту). В таблице 1 приведены некоторые примеры.

Таблица 1. Влияние датафикации на отрасли экономики

Отрасль	Влияние датафикации
Микро- и «короткие» цепочки поставок	Облегчает создание коротких цепочек поставок, создавая бизнес-процессы микро-цепочек поставок, инкапсулированные с помощью недорогих технологий, таких как мобильные устройства
Сельское хозяйство и продовольственная промышленность	Увеличивает прослеживаемость и устранение посредников в цепочке отраслевых поставок
Производство	Отзывы об уже используемых продуктах обеспечивают возможность улучшения методов разработки продуктов. Совместное использование данных несколькими производителями позволяет повысить эффективность в цепочках поставок.
Управление коммерческой недвижимостью	Анализирует, как различные части города классифицируются для промышленного использования. Увеличивает уровень детализации для клиентов недвижимости, чтобы понять, где расположить их бизнес для лучшего результата и обеспечения потока людей.

Источник: Ericsson, *The Impact Of Datafication On Strategic Landscapes* (Stockholm, 2014). Доступно по <https://www.ericsson.com/assets/local/news/2014/4/the-impact-of-datafication-on-strategic-landscapes.pdf>.

Датафикация также стимулирует использование алгоритмов для принятия решений.³⁷ Некоторые люди уже выступают за то, чтобы «многие решения, суждения и прогнозы, выполняемые в настоящее время людьми, были бы переданы алгоритмам».³⁸

36 Rahul Zingre, "The Increasing Datafication Of Our Lives", *LinkedIn*, 22 February 2018. Доступно по <https://www.linkedin.com/pulse/increasing-datafication-our-lives-rahul-zingre>.

37 Marijn Janssen, Yannis Charalabidis and Helmut Krcmar, "Open Data, Information Processing and Datafication of Government", *Proceedings of the 50th Hawaii International Conference on System Sciences* (2017), p. 2670. Доступно по <https://scholarspace.manoa.hawaii.edu/bitstream/10125/41478/1/paper0329.pdf>.

38 Andrew McAfee and Erik Brynjolfsson, *Machine, Platform, Crowd: Harnessing Our Digital Future* (New York and London,

В то время как некоторые приветствуют такое развитие, другие отмечают, что датафикация и принятие решений на основе алгоритмов имеют как негативные, так и непреднамеренные последствия.³⁹

1.3 Справедливость на основе данных

Революция данных имеет глубокие социальные последствия.

Для некоторых это открыло новые экономические возможности и большую свободу. Для других это означает более пристальное и нежелательное наблюдение.

Теперь у нас есть новый термин – наблюдение на основе данных (*dataveillance*) – наблюдение путем анализа объемов данных (*data footprints*). Он был придуман в середине 1980-х годов, чтобы «привлечь внимание к происходящему существенному изменению... от (дорогого) физического и электронного наблюдения за людьми к (дешевому) надзору за поведением людей через все более интенсивные «цифровые следы», генерируемые их поведением».⁴⁰ Понятие *dataveillance* также «конкретно указывает на способность переориентировать или подталкивать будущее поведение людей посредством четырех классов действий: (1) записанное наблюдение; (2) идентификация и отслеживание; (3) аналитическое вмешательство; и (4) поведенческие манипуляции».⁴¹

Данные могут поддерживать существующую асимметрию власти или способствовать большей справедливости распределения, делая бедных «видимыми».

Г. Томас Кингсли из Института урбанизации утверждает, что «нам нужно использовать данные таким образом, чтобы происходили изменения, а не просто отслеживались данные».⁴² Получение правильных данных для лиц, принимающих решения, определит, «будет ли в мире будущее городов историей инклюзивности и процветания или трагедией (более миллиарда людей, живущих в крайней нищете в городских трущобах со скудным водоснабжением, антисанитарией или недостатком других услуг – и в высоком риске экологической катастрофы)».⁴³

W. W. Norton & Company, 2017), p. 64.

39 Olivera Marjanovic and Dubravka Cezec-Kecmanovic, “Understanding Datafication Effects of Open Government Information Systems – A Contemporary Systems Thinking Approach”, *Proceedings of the 50th Hawaii International Conference on System Sciences* (2017). Доступно по <https://pdfs.semanticscholar.org/2d72/1ffcba1a30d3e259a06cbd2905f0dbb41419.pdf>.

40 Roger Clarke, “Dataveillance and Information Privacy”. Доступно по <http://www.rogerclarke.com/DV/#SurvD> (accessed on 8 January 2019).

41 Sara Degli-Esposti, “When big data meets dataveillance: The hidden side of analytics”, *Surveillance & Society*, vol. 12, no. 2 (May 2014), p. 210. Доступно по https://www.researchgate.net/publication/262493771_When_big_data_meets_dataveillance_The_hidden_side_of_analytics.

42 G. Thomas Kingsley, “Global development demands a data revolution that will make change happen, not just track it”, *Urban Institute*, 4 May 2017. Доступно по <https://www.urban.org/urban-wire/global-development-demands-data-revolution-will-make-change-happen-not-just-track-it>.

43 Там же

Линнет Тейлор определяет принцип справедливости на основе данных как: «Справедливость заключается в том, что в результате производства своих цифровых данных люди становятся видимыми, представленными и обслуживаемыми». ⁴⁴ Здесь главная задача заключается в том, «как сбалансировать и интегрировать необходимость быть замеченными и представленными соответствующим образом с потребностями в независимости и целостности». ⁴⁵

Справедливость на основе данных опирается на три принципа: (1) видимость; (2) взаимодействие с технологиями; и (3) недискриминация. ⁴⁶

Принцип видимости справедливости на основе данных касается как вопросов обеспечения конфиденциальности, так и представления тех, кто живет за пределами социально принятых норм. Сколько населения невидимо? И когда они становятся видимыми, рассматриваются ли они как субъекты / клиенты, подлежащие контролю, или как граждане с независимой властью?

Во всем мире более 1,1 миллиарда человек являются «невидимками», и более трети из них - дети. ⁴⁷ В развивающихся странах отсутствие данных о детях может привести к двум вещам. Во-первых, это может сохранить навсегда существующие модели бедности и неравенства из-за плохой информации, используемой при принятии решений. ⁴⁸ Во-вторых, их «более трудно отслеживать и они становятся легкими жертвами торговцев людьми, сексуальной эксплуатации в коммерческих целях, незаконного усыновления, детоубийства и других форм насилия». ⁴⁹

Те, кто живет за пределами социально принятых норм, также невидимы. Переписи/опросы населения на национальном уровне не включают вопросы, касающиеся сексуальной ориентации и гендерной идентичности или самовыражения. ⁵⁰ Следствием этого является «нехватка понимания об опыте сообщества лесбиянок, геев, бисексуалов, трансгендеров, и о том, как гендерная идентичность и сексуальность влияют на другие демографические данные, такие как раса, способности или геополитика». ⁵¹

Второй принцип справедливости на основе данных – это взаимодействие с технологиями. Этот принцип касается свободы не использовать конкретные цифровые технологии и, что более важно, того, как не стать частью коммерческих баз данных (как

44 Linnet Taylor, "What is data justice? The case for connecting digital rights and freedoms globally", 16 February 2017, p. 1. Доступно по https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2918779.

45 Там же, стр.18.

46 Там же, стр. 15.

47 Agence France-Presse, "More than 1 billion 'invisible people' worldwide have no proof of identity", 22 October 2017.

Доступно по

<https://www.pri.org/stories/2017-10-22/more-1-billion-invisible-people-worldwide-have-no-proof-identity>.

48 Caroline Ford, "Making Invisible Populations Count – The development agenda of the 21st century", *Consortium for Street Children*, 12 October 2018. Доступно по <https://www.streetchildren.org/news-and-updates/making-invisible-populations-count-the-development-agenda-of-the-21st-century/>.

49 Lucia Hanmer and Marina Elefante, *The Role of Identification in Ending Child Marriage: Identification for Development (ID4D)* (Washington D.C., World Bank, 2016), p. 7. Доступно по <http://documents.worldbank.org/curated/en/130281472492551732/pdf/107932-WP-P156810-OUO-9-Child-Marriage.pdf>.

50 Charlie Whittington, "Invisible In Data: The Lack of LGBT Data Collection", *Georgetown Public Policy Review*, 17 July 2018. Доступно по <http://gppreview.com/2018/07/17/invisible-data-lack-lgbtq-data-collection/>.

51 Там же

побочный продукт активности в целях развития). Это также включает свободу контролировать условия взаимодействия кого-либо с рынками данных.

Из 3,2 миллиарда людей, имеющих интернет, 2 миллиарда из развивающихся стран. Преодолев проблему доступа, теперь они сталкиваются с другой проблемой – надзором за сбором данных.

Сбор персональных данных стал нормой в нашем интернет-связанном мире. Бизнес регулярно собирает данные своих клиентов. Помимо продолжения сбора административных данных, правительства теперь контролируют учетные записи своих граждан в социальных сетях.⁵² Брокеры (посредники) данных «получают личные данные от государственных и частных лиц для объединения в целях перепродажи бизнес-предприятиям».⁵³

Подобно цифровому разрыву, социально-экономическое положение, пол, этническая принадлежность и место происхождения определяют, кто подвергается слежению. Эти факторы определяют, в каких базах данных состоит отдельный человек, как эти системы используют его данные и какое влияние они могут оказывать на отдельных людей.⁵⁴ Например:

Подросток из семьи иммигрантов, проживающий в районе с низкими доходами, чьи родители бедны и принадлежат к этнической группе и религиозному меньшинству, экспоненциально чаще подвергается слежению со стороны как защитных (социальные службы), так и профилактических (правоохранительные органы) ведомств, и, вероятно, также будет иметь меньше возможностей противостоять этому наблюдению или вмешательству, чем ее подруга, которая живет в районе с высоким уровнем дохода и принадлежит к этнической группе большинства.⁵⁵

Кроме того, последние тенденции свидетельствуют о снижении способности контролировать условия взаимодействия кого-то с рынками данных. Европейская инспекция по обеспечению конфиденциальности данных отмечает, что «постоянно увеличивающиеся объемы личной информации собираются и обрабатываются все более непрозрачными и сложными способами».⁵⁶ Что еще хуже, «следующее поколение персональных данных, вероятно, будет еще менее доступно для лиц, к которым они имеют отношение».⁵⁷ Европейская инспекция по обеспечению конфиденциальности данных признает, что такое развитие имеет важные последствия для защиты человеческого достоинства, свободы личности и функционирования общества.⁵⁸

52 Heidi Swart, "Government surveillance of social media is rife. Guess who's selling your data?" *Daily Maverick*, 25 April 2018. Доступно по <https://www.dailymaverick.co.za/article/2018-04-25-government-surveillance-of-social-media-is-rife-guess-whos-selling-your-data/>.

53 World Bank, *Information and Communications for Development 2018: Data-Driven Development* (Washington D.C., 2019), p. 54. Доступно по <https://www.worldbank.org/en/topic/digitaldevelopment/publication/data-driven-development.print>.

54 Linnet Taylor, "What is data justice? The case for connecting digital rights and freedoms globally", 16 February 2017, p. 3. Доступно по https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2918779.

55 Там же

56 European Data Privacy Supervisor, "Towards a new digital ethics: Data, dignity and technology", Opinion 4, 2015, p. 6. Доступно по https://edps.europa.eu/sites/edp/files/publication/15-09-11_data_ethics_en.pdf.

57 Там же, стр. 9.

58 Там же

Недискриминация, третий принцип справедливости на основе данных, состоит из способности выявлять и оспаривать предвзятость в использовании данных, а также свободу от предвзятого обращения.

Случай коренного полинезийского народа Новой Зеландии является показательным. Новозеландских форум по будущему данных сообщает, что:

Многие представители маори не считают, что они извлекли большую пользу из сбора и использования данных. Они чувствуют реальный и непосредственный риск от большей доступности данных, используемых для этнического профилирования, к возможности вреда. Несмотря на широко распространенные требования к ним в отношении данных в прошлом, данные, по-видимому, редко используются способами, которые могли бы им помочь... Сбор, хранение и использование данных часто происходит таким образом, который не уважает законов «*тиканга*» маори.⁵⁹

Принцип недискриминации подразумевает, что:

Должны быть разработаны методы, которые могут позволить управлять алгоритмическими процессами и принятием решений, и ответственность за борьбу с дискриминацией по отношению отдельных лиц должна сопровождаться способностью выявлять и налагать штрафы за нее по отношению к правительству.⁶⁰

Подходом, имеющим связь со справедливостью на основе данных, является структурная справедливость данных. Данный подход направлен на: «Уровень, при котором общество содержит и поддерживает связанные с данными институты, отношения и системы знаний, необходимые для реализации ценностей, лежащих в основе хорошей жизни».⁶¹

Подход структурной справедливости данных утверждает, что недостаточно законодательно закрепить открытый доступ к политике в отношении данных. Доступ к данным выходит за рамки законов, экономических затрат или физического расстояния. Чтобы действительно обеспечить доступ к данным, нам необходимо изучить, как социальные структуры (структурная позиция, ресурс, институциональный и эпистемологический контроль) обеспечивают или ограничивают доступ. Например, цифровизация государственных услуг в развивающихся странах, в первую очередь, принесет пользу тем, у кого есть доступ к Интернету, и которые, вероятно, будут состоятельными мужчинами, живущими в городских районах.

59 Elizabeth Stuart, Emma Samman, William Avis and Tom Berliner, "The data revolution: Finding the missing millions", Overseas Development Institute Research Report 3, April 2015, p. 43. Доступно по <https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9604.pdf>.

60 Там же, стр. 9.

61 Richard Heeks, "A Structural Model and Manifesto for Data Justice for International Development", Development Informatics Working Paper No. 69, University of Manchester, 2017. Доступно по http://hummedia.manchester.ac.uk/institutes/gdi/publications/workingpapers/di/di_wp69.pdf.

Манифест «Справедливость данных для развития» - это первый шаг в продвижении справедливости на основе данных в «датафицированном» мире. Манифест призывает к следующему:

1. Требование справедливого и законного использования данных в целях развития;
2. Получение данных с согласия граждан, которые действительно информированы;
3. Нарращивание восходящих и нисходящих возможностей, связанных с данными, среди тех, кому их не хватает в развивающихся странах;
4. Продвижение прав доступа к данным, конфиденциальности данных, владения данными и представления данных;
5. Содействие результатам системы данных, которые касаются целей и приоритетов международного развития, включая цели и приоритеты субъектов данных;
6. Поддержка использования «небольших данных» отдельными лицами и сообществами в развивающихся странах;
7. Пропаганда устойчивого использования данных и систем данных;
8. Создание общественного движения за «данные из периферии» глобального Юга;
9. Поощрение альтернативных суждений вокруг интенсивного развития данных, ставящего во главу угла вопросы справедливости;
10. Разработка новых организационных форм, таких как кооперативы развития с интенсивным использованием данных;
11. Лоббирование новых законов и политики, основанных на справедливости в отношении данных, в развивающихся странах (включая действия в отношении монополистов данных);
12. Создание, оспаривание и предложение альтернатив техническим структурам (код, алгоритмы, стандарты и т. д.), связанным с данными, которые все больше контролируют международное развитие.⁶²

Справедливость данных необходима для инклюзивного развития и управления на основе данных.

2. Данные и управление

Задачи данного раздела:

- Дать определение управлению;
 - Определить управление на основе данных и управляемое данными правительство;
 - Обсудить принятие решений на основе данных, разработку политики на основе фактических данных и администрирования на основе результатов как формы управления на основе данных;
 - Обсудить достижение Целей в области устойчивого развития в качестве примера управления на основе данных на глобальном уровне.
-

2.1 Управление

Управление – это то, как общества организуются для определения и достижения общего будущего.

Это определение согласуется с определением управления Петерса и Пьера как «определения и достижения коллективных интересов», а также с точкой зрения Пракаша и Харта об управлении как «организации коллективных действий».⁶³

Управление – это «постоянный процесс, посредством которого могут быть учтены противоречивые или различные интересы и могут быть приняты совместные меры» и включает «официальные институты и властные структуры, уполномоченные обеспечивать соблюдение законов, а также неформальные договоренности, с которыми люди и учреждения либо согласились, либо восприняли их в своих интересах».⁶⁴ Управление «не является статичным и представляет собой постоянную модель адаптации к возможностям и обстоятельствам».⁶⁵

В управлении участвуют как государственные, так и частные субъекты. «Новая теория управления», по словам Джона Рагги, «основывается на предпосылке, что государство само по себе не может выполнить всю тяжелую работу, необходимую для решения самых насущных социальных проблем, и поэтому ему необходимо привлекать других участников для использования своих возможностей».⁶⁶

63 B. Guy Peters and Jon Pierre, *Comparative Governance: Rediscovering the Functional Dimension of Governing* (Cambridge, Cambridge University Press, 2016), p. 6; and Aseem Prakash and Jeffrey Hart, "Globalization and Governance: an Introduction" in *Globalization and Governance*, Aseem Prakash and Jeffrey Hart, eds. (London and New York, Routledge, 1999), p. 2.

64 Commission on Global Governance, "Our Global Neighborhood: Chapter One – A New World". Доступно по <https://www.gdrc.org/u-gov/global-neighborhood/chap1.htm> (accessed on 8 January 2019).

65 B. Guy Peters and Jon Pierre, *Comparative Governance: Rediscovering the Functional Dimension of Governing* (Cambridge, Cambridge University Press, 2016), p. 16.

66 John Gerard Ruggie, "Global Governance and 'New Governance Theory': Lessons from Business and Human Rights", *Global Governance*, vol. 20 (2014), pp. 8-9. Доступно по <https://www.hbs.edu/faculty/conferences/2014-business-beyond-the-private-sphere/Documents/Global%20Governance%20and%20%27New%20Governance%20Theory%27.pdf>.

Вопросы управления можно классифицировать в качестве «политики спроса» и «политики предложения».⁶⁷

Политика спроса - это выражение общественных ожиданий, обычно направленных на правительство. Включает в себя проявления желания совместных действий. Политика предложения заключается в «способности правительства удовлетворять (гражданские) ожидания и решать социальные проблемы».⁶⁸ Основное внимание уделяется способности гражданского общества (например, бизнес-сектора, неправительственных организаций и других групп) предоставлять меры и решения проблем общества.

Управление происходит на разных уровнях - местном, национальном и глобальном.⁶⁹

Местное управление осуществляется на уровне сообщества, в пределах небольшой территории или на самом низком уровне политических организаций страны. Для Программы развития Организации Объединенных Наций (ПРООН) местное управление «является самым близким к гражданам каналом для доступа к базовым услугам, для участия в общественных решениях, которые влияют на их жизнь, и для осуществления их прав и обязанностей».⁷⁰

Управление на национальном уровне происходит на всей территории страны. Оно представляет собой «способ использования власти при управлении экономическими и социальными ресурсами страны в целях развития».⁷¹

Глобальное управление – это «способ, с помощью которого субъекты – отдельные лица, учреждения (опять-таки как государственные, так и частные) - пытаются учитывать конфликтующие интересы посредством процессов принятия коллективных действий в ряде областей, действующих за пределами (национальных) государственных границ».⁷² Это также «международный процесс формирования консенсуса, который порождает руководящие принципы и соглашения».⁷³

Ниже в таблице 2 ниже перечислены проблемы управления на разных уровнях.

Таблица 2. Проблемы управления на местном, национальном и глобальном уровнях

	Местный	Национальный	Глобальный
--	---------	--------------	------------

67 B. Guy Peters and Jon Pierre, *Comparative Governance: Rediscovering the Functional Dimension of Governing* (Cambridge, Cambridge University Press, 2016), p. 8.

68 Там же

69 John Pierre and B. Guy Peters, *Governance, Politics, and the State* (London, McMillan Press Ltd., 2000), pp. 75-93.

70 UNDP, "Local Governance and Local Development". Доступно по <http://www.undp.org/content/undp/en/home/democratic-governance-and-peacebuilding/responsive-and-accountable-institutions/local-governance-and-local-development.html> (accessed on 8 January 2019).

71 World Bank, *Governance and Development* (Washington D.C., World Bank, 1992), p. 1. Доступно по <http://documents.worldbank.org/curated/en/604951468739447676/pdf/multi-page.pdf>.

72 Richard Higgott, "The Theory and Practice of Global and Regional Governance: Accommodating American Exceptionalism and European Pluralism", University of Warwick GARNET Working Paper No. 01/05, pp. 4-5. Доступно по <https://warwick.ac.uk/fac/soc/pais/research/researchcentres/csgr/garnet/workingpapers/0105.pdf>.

73 InternationalRelations.org, "Global Governance Definition". Доступно по <http://internationalrelations.org/global-governance/> (accessed on 8 January 2019).

Политика спроса	Антикоррупционные кампании	Выборы, лоббирование, общественные движения	Протесты в Сиэтле против Всемирной торговой организации, борьба с торговлей детьми
Политика предложения	Социальные службы, управление общественным достоянием	Государственные законы и правила, оборона и безопасность	Цели развития тысячелетия, цели устойчивого развития

Управление тесно связано с правительством или «формальными структурами государственного сектора и группой субъектов, осуществляющих государственную власть».⁷⁴ Оран Янг определяет правительство как «материальное образование, специализирующееся на обеспечении управления».⁷⁵

Также общепризнанно, что управление «не ограничивается правительством, поскольку другие социальные институты также могут предоставлять услуги по управлению».⁷⁶

Оран Янг утверждает о пяти различных системах управления:

1. Управление правительством;
2. Управление по межправительственному соглашению;
3. Частное управление через отраслевое саморегулирование и кодексы поведения;
4. Гражданское управление или управление со стороны гражданского общества;
5. Гибридные механизмы, в которых различные системы управления объединяются для решения конкретных задач.⁷⁷

Лемос и Агравал расширили гибридный механизм управления, определив его четыре формы:

1. Совместное управление (между государственными органами и сообществами);
2. Государственно-частные партнерства (между государственными органами и бизнес-участниками);
3. Частно-социальные партнерства (между бизнес-участниками и сообществами);
4. Управление несколькими партнерами (с участием всех типов представителей).⁷⁸

Несмотря на присутствие частных субъектов, правительство остается ключевым представителем управления. Основное различие между государственными и частными

⁷⁴ B. Guy Peters and Jon Pierre, *Comparative Governance: Rediscovering the Functional Dimension of Governing* (Cambridge, Cambridge University Press, 2016), p. 5.

⁷⁵ Oran Young, "Governance for Sustainable Development in a World of Rising Interdependencies", in *Governance for the Environment*, Magali A. Delmas and Oran Young, eds. (Cambridge, Cambridge University Press, 2009), p. 20.

⁷⁶ Aseem Prakash and Jeffrey Hart, "Globalization and Governance: an Introduction" in *Globalization and Governance*, Aseem Prakash and Jeffrey Hart, eds. (London and New York, Routledge, 1999), p. 2.

⁷⁷ Oran Young, "Governance for Sustainable Development in a World of Rising Interdependencies", in *Governance for the Environment*, Magali A. Delmas and Oran Young, eds. (Cambridge, Cambridge University Press, 2009), pp. 24-30.

⁷⁸ Maria Camen Lemos and Arun Agrawal, "Environmental Governance and Political Science", in *Governance for the Environment*, Magali A. Delmas and Oran Young, eds. (Cambridge, Cambridge University Press, 2009), p. 79.

субъектами заключается в том, что правительство «является властью, способной принуждать, заставлять».⁷⁹ Правительство также имеет право собирать налоги. Как отмечается Комиссией по глобальному управлению, «любая адекватная система управления должна обладать способностью контролировать и использовать ресурсы, необходимые для достижения ее основных целей».⁸⁰

Хотя есть те, кто относится к правительству как к одному из многих действующих лиц управления, их недостаток заключается в том, что они «недооценивают постоянство и в некоторых случаях распространение форм политической власти в отношении безопасности, суверенитета, а также применения и угрозы физической силы и насилия».⁸¹

В данном модуле есть второе значение управления: «Управление как руководство и координация».⁸² Это определение используется в контексте управления данными - «совокупность методов и процессов, которые помогают обеспечить формальное администрирование активами данных в организации».⁸³

2.2 Управление на основе данных

Идея использования данных для поддержки управления и принятия решений правительства не нова. История статистики тесно связана с появлением модернизирующегося государства XIX века.⁸⁴ Отличительная черта нынешних усилий заключается в том, что революция в области данных сделала идеал более достижимым.

Сегодня мы используем словосочетание «управляемый данными» для описания деятельности, которая «обусловлена данными, а не интуицией или личным опытом».⁸⁵ Управляемый данными также означает принятие «стратегических решений, основанных на анализе и интерпретации данных».⁸⁶

Управление на основе данных определяется как интенсивное и широкое использование данных в деятельности общества по обозначению и достижению общего будущего. Оно охватывает повсеместное использование данных как в политике спроса, так и в

79 Ann Florini, "Global Governance and What It Means", *Brookings*, 16 February 2009. Доступно по <https://www.brookings.edu/on-the-record/global-governance-and-what-it-means/>.

80 Commission on Global Governance, "Our Global Neighborhood: Chapter One – A New World". Доступно по <https://www.gdrc.org/u-gov/global-neighborhood/chap1.htm> (accessed on 8 January 2019).

81 Mitchell Dean, *Governing Societies: Political Perspectives on Domestic and International rule* (Berkshire and New York, Open University Press, McGraw Hill Education, 2007) p. 58.

82 B. Guy Peters and Jon Pierre, *Comparative Governance: Rediscovering the Functional Dimension of Governing* (Cambridge, Cambridge University Press, 2016), p. 4.

83 Charles Roe, "What is Data Governance?" *Dataversity*, 18 December 2017. Доступно по <http://www.dataversity.net/what-is-data-governance/>.

84 Stuart Woolf, "Statistics and the Modern State", *Comparative Studies in Society and History*, vol. 31, no. 3 (July 1989), pp. 588-604. Доступно по https://www.jstor.org/stable/178772?read-now=1&refreqid=excelsior%3Ae9117494e9c81b206098b366401bfc44&seq=1#page_scan_tab_contents.

85 Wikipedia, "Data-driven". Доступно по <https://en.wikipedia.org/wiki/Data-driven> (accessed on 8 January 2019).

86 AT Internet, "Glossary: Data-Driven". Available at <https://www.atinternet.com/en/glossary/data-driven/> (accessed on 8 January 2019).

политике предложения.

Правительство, осуществляющее управление на основе данных – это «то место, где для всех критических решений доступна, когда и где это необходимо, действенная информация».⁸⁷ Это может привести к:

- Более эффективному управлению и контролю;
- Оптимизированному обнаружению мошенничества и ошибок, уменьшению и предотвращению;
- Улучшенному предоставлению услуг, основанному на информации, полученной от тех, кто обслуживается;
- Повышению эффективности благодаря интеллектуальным сетям, что может привести к снижению затрат;
- Улучшению общественного восприятия работы ведомства.⁸⁸

Вставка 9. Как данные могут влиять на правительство, ориентированное на интересы граждан: случай Австралии

Благодаря взаимодействию с гражданами государственные органы собирают огромное количество данных.

Данные могут помочь правительству улучшить оказание государственных услуг и разработку политики, предоставляя широкие социальные льготы гражданам и улучшая ту деятельность, которая осуществляется правительством.

Данные могут открыть способы помощи Австралии по активному планированию и реагированию на некоторые из самых больших проблем, стоящих перед правительством в таких областях, как здравоохранение, социальное обеспечение и вопросы коренных народов.

Помимо разработки более эффективных политик, данные могут выступать в качестве ключевого показателя эффективности для измерения результатов реализации политики. Данные также могут предоставить возможность юрисдикциям и неправительственным организациям добиться большего без дополнительного финансирования.

Австралийцы уже ощущают преимущества более эффективного использования данных при предоставлении услуг. Внедрение биометрических технологий в аэропортах позволило сократить очереди на иммиграцию, а электронные медицинские карты, контролируемые лично, помогли улучшить планы борьбы с заболеваниями.

Таким образом, данные могут помочь правительству Австралии:

- Разрабатывать более разумную, обоснованную политику;
- Предоставлять более улучшенные услуги, ориентированные на интересы граждан;
- Повысить эффективность государственного сектора и сократить накладные расходы;
- Оценить эффективность программ и сторонних поставщиков услуг.

⁸⁷ Terence Lutes, “Data-driven government: Challenges and a path forward”, IBM Analytics White Paper, 2015, p. 3. Доступно по <https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=GQW03008USEN>.

⁸⁸ Там же

Взято (с изменениями) из: Commonwealth of Australia, Department of the Prime Minister and Cabinet, "Public Sector Data Management", July 2015, p. 12. Доступно по https://www.pmc.gov.au/sites/default/files/publications/public_sector_data_mgt_project.pdf.

Расширенное использование данных может улучшить управление на местном, национальном и глобальном уровнях. Данные могут помочь сделать услуги, предоставляемые населению, более целенаправленными и эффективными, а также привести к лучшему распределению ограниченных ресурсов и пониманию причин и решений дорогостоящих социальных проблем. Данные также могут способствовать информированному участию граждан в управлении.

Опрос британских органов местного самоуправления 2016 года выявил следующие 11 ключевых вариантов использования данных:

1. Оптимизация управления площадями и инфраструктурой;
2. Тестирование на эффективность;
3. Интеллектуальное управление делами;
4. Управление производительностью на основе результатов;
5. Раннее выявление неблагоприятных событий и воздействий в сфере услуг в будущем;
6. Понимание и реагирование на потребности граждан;
7. Информирование о преобразовании государственных услуг;
8. Оптимизация процессов совещательной работы;
9. Обеспечение открытости правительства;
10. Поддержка экономики на местах, бизнеса и инноваций;
11. Выявление мошенничества и ошибок.⁸⁹

В социальном секторе «данные сделают работу агентов социальных изменений более эффективной и послужат основанием для поддержки лучших программ и предприятий».⁹⁰ Данные могут также улучшить процесс принятия решений.

2.2.1 Принятие решений на основе данных

Принятие решений на основе данных «включает сбор данных, извлечение закономерностей и фактов из этих данных, и использование этих фактов для вывода заключения, влияющее на принятие решений».⁹¹

Внедренное в повседневную деятельность принятие решений на основе данных может трансформировать местные и национальные органы власти от простого реагирования

89 Tom Symons, "Wise Council: Insights From the Cutting Edge of Data-Driven Local Government", NESTA, November 2016, p. 7. Доступно по https://media.nesta.org.uk/documents/wise_council.pdf.

90 Jim Fruchterman, "Using Data for Action and for Impact", *Stanford Social Innovation Review* (Summer 2016). Доступно по https://ssir.org/articles/entry/using_data_for_action_and_for_impact.

91 Joel Schwartz, "Data-Driven Decision Making: A Primer For Beginners", *Northeastern University Blogs*, 14 December 2017. Доступно по <https://www.northeastern.edu/graduate/blog/data-driven-decision-making/>.

на потребности и беспокойства граждан к упреждающему прогнозированию проблем, наиболее важных для граждан.

Как тогда правительства могут перейти к принятию решений на основе данных?

Хорошее начало – это определение уровня использования данных в организации.

Модель зрелости данных, созданная британской благотворительной организацией по инновациям Nesta, предоставляет полезный и удобный способ для оценки состояния принятия решений на основе данных в правительствах.⁹² В этой модели зрелость данных определяется по пятибалльной шкале от «зарождающегося» до «датавора» следующим образом:

- **Зарождающийся** – Богат данными, обделен интеллектом. Данные не являются ключевой частью процессов принятия решений.
- **Базовый** – Данные используются в отчетах, но, как правило, поверхностно и практически без ссылок на решения, которые необходимо принять.
- **Средний** – Анализ данных обычно запрашивается для принятия решения, но может быть не отвечающим требованиям, потому что анализ не высокого качества, не нацелен на решение, которое должно быть принято, или нужных данных нет.
- **Продвинутый** – Некоторые решения основываются на данных как с низового, так и старшего уровней, но они не согласованы в рамках всей организации.
- **Датавар (Datavore)** – Богатый интеллект данных и понимание. Данные анализируются специально для принятия ключевых решений последовательно во всей организации. Данные доступны своевременно для поддержки принятия решений.⁹³

Следующим шагом является разработка плана, который выведет организацию на более высокий уровень зрелости.

Десять шагов Бернарда Марра по улучшению использования данных при принятии решений для бизнеса могут быть адаптированы к местным и национальным органам власти.⁹⁴ Шаги такие:

1. **Начните со стратегии** – Вместо того, чтобы начинать с того, к каким данным вы могли или должны получить доступ, начните с выяснения того, чего ваша организация хочет достичь.
2. **Сосредоточьтесь «бизнес»-областях** – Теперь вам нужно определить, какие области наиболее важны для достижения вашей общей стратегии. Если бы вы могли работать только над улучшением одной или двух областей, что бы вы выбрали?

92 Tom Symons, "Wise Council: Insights From the Cutting Edge of Data-Driven Local Government", *NESTA*, November 2016, p. 105. Доступно по https://media.nesta.org.uk/documents/wise_council.pdf.

93 Там же

94 Bernard Marr, "Data-Driven Decision Making: 10 Simple Steps for Any Business", *Forbes*, 14 June 2016. Доступно по <https://www.forbes.com/sites/bernardmarr/2016/06/14/data-driven-decision-making-10-simple-steps-for-any-business/>.

3. **Определите ваши вопросы без ответа** – Анализируя именно то, что вам нужно знать, вы можете сосредоточиться на данных, которые вам действительно нужны. Ваши требования к данным, уровень затрат и уровень стресса значительно снижаются, когда вы переходите от «собирать все на всякий случай» к «собирать и измерять x и y для ответа на вопрос z».
4. **Найдите данные, чтобы ответить на ваши вопросы** – Важно понимать, что ни один тип данных по своей природе не является лучшим или более ценным, чем любой другой тип. Сосредоточьтесь на определении идеальных данных для вашей организации - данных, которые могут помочь вам ответить на ваши самые насущные вопросы и выполнить ваши стратегические задачи.
5. **Определите, какие данные у вас уже есть** – Проведите инвентаризацию данных, которые ваша организация имеет или может получить к ним доступ. Если данные еще не существуют, найдите способы их сбора, либо внедрив системы сбора данных, либо приобретая или получая доступ к данным извне.
6. **Приступайте к работе, если затраты и усилия оправданы** – Как только вы узнаете стоимость сбора необходимых данных, вы сможете решить, перевесит ли ощутимые выгоды эти затраты. Вы должны относиться к данным как к любым другим ключевым инвестициям – вам необходимо четко обосновать инвестиции, которые отражают долгосрочную ценность данных для вашей стратегии.
7. **Собирайте данные** – Большая часть этого шага сводится к настройке процессов и подбору людей, которые будут собирать и администрировать ваши данные. Возможно, вы приобретете доступ к готовому для анализа набору данных, и в этом случае нет необходимости собирать данные как таковые. Но, на самом деле, многие проекты, связанные с данными, требуют определенных действий по сбору данных.
8. **Проведите анализ данных** – вам необходимо проанализировать данные, чтобы извлечь значимые и полезные бизнес-выводы.
9. **Представьте и распространите результаты** – Если результаты не будут представлены нужным людям в нужное время должным образом, то размер наборов данных или сложность аналитических инструментов не будут иметь большого значения. Вы должны убедиться, что полученные из ваших данных сведения используются для принятия решений и, в конечном итоге, для повышения производительности.
10. **Внедрите результаты в организацию** – Наконец, вам необходимо применить результаты, полученные из анализа данных, для принятия решений, которые преобразят вашу организацию.

В идеале правительства должны стать «датаварами».

Определение и измерение успеха отождествляется в качестве барьера для более быстрого внедрения принятия решений на основе данных в правительстве⁹⁵, а именно:

95 Mahesh Kelkar and others, "Mission analytics: Data-driven decision making in government", *Deloitte Insights*, 26 September 2016. Доступно по <https://www2.deloitte.com/insights/us/en/industry/public-sector/data-driven-decision-making-in-government.html>.

- **Препятствие 1** – Трудно понять, что важно. Меры успеха может быть трудно определить. В правительстве проблемы определения успеха и определения мер, которые способствуют этому, могут затруднить разработку ключевых показателей эффективности.
- **Препятствие 2** – Цели ведомства могут не совпадать с общей миссией правительства.
- **Препятствие 3** – Самооценки неоднозначны и субъективны. Неразумно ожидать, что должностные лица ведомства будут давать объективную самооценку своих собственных результатов, так как желание подчеркнуть успех может внести неоднозначность и субъективность в данные.
- **Препятствие 4** - Форматы хранения могут препятствовать использованию оперативных данных. Некоторые данные, имеющие большое значение для оптимизации распределения ресурсов, например, важные сведения о том, почему действия ведомства выполняются успешно или неудачно, хранятся в форматах, затрудняющих доступ к ним. Документы могут быть отсканированными копиями рукописных заметок или полей с произвольным текстом в неструктурированных форматах.⁹⁶

Эти препятствия могут быть преодолены следующими способами:

- **Решение 1** – Определите наиболее полезные меры. Время, необходимое для определения того, какие меры наиболее информативны и наиболее тесно связаны с успехом миссии, является не затраченным зря.
- **Решение 2** – Создайте «линию прямой видимости» от каждого сотрудника по крайней мере до одной ведомственной цели высшего уровня. Сотрудникам очень полезно понять, как их работа способствует общей миссии ведомства. Для поддержки такой связи правительственные учреждения могут распространять информацию о «линии прямой видимости», связывающей роль каждого сотрудника с одной или несколькими целями высокого уровня.
- **Решение 3** – Улучшайте управление и анализ данных о производительности. Ведомства, возможно, пожелают передать ответственность за сбор и отчетность по показателям эффективности независимому субъекту, потенциально подчиненному главному сотруднику по данным. В качестве альтернативы можно рассмотреть вопрос об увеличении инвестиций в технологические решения для обеспечения качества данных о производительности.
- **Решение 4** – Используйте когнитивные технологии для расширения и углубления данных о производительности. Когнитивные технологии – это методы обработки информации, которые могут выполнять определенные задачи, требовавшие до недавнего времени человеческого труда. Некоторые из этих технологий, такие как обработка естественного языка, распознавание речи и робототехника, получают широкое распространение и достигают быстрого прогресса.⁹⁷

⁹⁶ Там же

⁹⁷ Там же

2.2.2 Разработка политики на основе фактов

Разработка политики на основе фактов определяется следующим образом: «Использование наилучших доступных исследований и информации о результатах исполнения программ для принятия решений на всех этапах политического процесса и в каждом подразделении правительства».⁹⁸

Четыре принципа формирования политики на основе фактов:

1. Составьте и соберите строгие доказательства того, что работает, включая затраты и выгоды;
2. Проводите мониторинг реализации программы и используйте оценки воздействия для измерения эффективности программы;
3. Используйте строгие доказательства для улучшения программ, расширяйте то, что работает, и перенаправляйте средства от постоянно неэффективных программ;
4. Поощряйте инновации и пробуйте новые подходы.⁹⁹

Разработка политики на основе фактов может использоваться при оценке программ, разработке бюджета, контроле за реализацией, мониторинге результатов и целенаправленной оценке.¹⁰⁰

Ниже приведен стилизованный процесс разработки политики на основе фактов:

1. Определите, ЧТО представляет собой реальное доказательство.
 - a. Методология – Позволяет ли аналитический подход должным образом рассмотреть проблемы?
 - b. Потенциал - Достаточно ли исследовательских навыков для проведения анализа?
2. Определите, КОГДА имеются достаточные доказательства для принятия решений.
 - a. Время - Когда собирать данные, проводить сбор новых данных и проверять анализ?
 - b. Хорошие данные – Поддерживают ли качественные базы данных своевременный анализ?
3. Подумайте, КАК обеспечивается достоверное доказательство.
 - a. Прозрачность – Есть ли открытые дебаты и дискуссии для проверки и просвещения общественности?

98 The Pew Charitable Trusts and MacArthur Foundation, "Evidence-Based Policymaking: A guide for effective government", November 2014, p. 2. Доступно по

<http://www.pewtrusts.org/~media/assets/2014/11/evidencebasedpolicymakingaguideforeffectivegovernment.pdf>.

99 Evidence-Based Policymaking Collaborative, "Principles of Evidence-Based Policymaking", September 2016. Доступно по <http://www.evidencecollaborative.org/principles-evidence-based-policymaking>.

100 The Pew Charitable Trusts and MacArthur Foundation, "Evidence-Based Policymaking: A guide for effective government", November 2014, p. 5. Доступно по

<http://www.pewtrusts.org/~media/assets/2014/11/evidencebasedpolicymakingaguideforeffectivegovernment.pdf>.

- b. Независимость – Существуют ли стимулы для предоставления советов в общественных интересах (а не для корпоративных / личных интересов)?
4. Создайте восприимчивую политическую среду.
- a. Есть ли желание протестировать варианты политики, и есть ли структуры и ресурсы для этого?¹⁰¹

Для сторонников разработки политики, основанной на фактах, «большая приверженность научной передовой практике, сокращение манипуляций или неправомерного использования доказательств и более широкое применение науки приведет к улучшению результатов социальной политики».¹⁰² У них также есть предпочтительная методология в определении лучших доказательств.

С другой стороны, критики утверждают, что основанное на фактах формирование политики «затмевает политическую природу решений и при этом «смещает» решения в сторону конкретных результатов – то, что можно назвать деполитизацией политики».¹⁰³ Критики разработки политики на основе фактов считают, что политический процесс лучше всего рассматривать в качестве арены для конкуренции между соперничающими группами и / или интересами.

Джастин Пархерст, признавая важную роль доказательств и политическую природу процесса формулирования политики, разработал систему принципов, которую он называет «надлежащим управлением доказательствами». Элементами данной системы принципов являются:

- **Уместность** – Должны быть выбраны доказательства для решения (множественных) соответствующих социальных проблем. Доказательства необходимо рассмотреть с точки зрения создания способами, которые полезны для достижения целей политики. Применимость доказательств к местному контексту должна быть четко рассмотрена.
- **Качество** – Используемые доказательства должны оцениваться по их качеству, но критерии качества должны отражать методологические принципы, относящиеся к форме используемых исследований (например, качественные интервью в сравнении с клиническими испытаниями) и характеру полученных данных (например, описания в сравнении с измерениями или в сравнении с оценками).
- **Строгость** – Доказательства, представленные на рассмотрение политики, следует тщательно (всесторонне) собирать или обобщать, избегая выборочной практики «снятия сливок».
- **Управление** – Субъект, устанавливающий правила и форму официальных консультативных систем по доказательствам, должен иметь официальный мандат.
- **Представительство** – Власть, принимающая окончательное решение, основанное

101 Gary Banks, "Challenges of evidence-based policy-making", *Australian Public Service Commission*. Доступно по <https://www.apsc.gov.au/challenges-evidence-based-policy-making> (accessed on 8 January 2019).

102 Justin Parkhurst, *The Politics of evidence: From evidence-based policy to the good governance of evidence* (Oxford, Routledge, 2017) p. 27.

103 Там же, стр. 42-43.

на доказательствах, в отношении политики должна принадлежать к демократически представительным и публично ответственным должностным лицам.

- **Прозрачность** – Открытая информация и ясные способы для общественности увидеть, как выявляются и используются доказательные базы, определяющие принятие решения.
- **Обсуждение** – Общественность должна быть вовлечена таким образом, чтобы в процессе разработки политики учитывались различные коллизии ценностей и проблем. Это предполагает рассмотрение различных точек зрения, даже если не все проблемы могут быть учтены в окончательных политических решениях.
- **Состязательность** – Технические доказательства и научные исследования, используемые в политических решениях, должны быть открыты для критических вопросов и апелляций. Это включает в себя оспаривание конкретных научных результатов, а также позволяет рассматривать доводы в отношении решений о том, какие доказательства использовать (например, ставить под сомнение уместность доказательств для конкретного случая).¹⁰⁴

Может ли разработка политики на основе фактов помочь решить «злободневные проблемы», такие как изменение климата, бедность, недоедание, неграмотность или бездомность?

По мнению Австралийской комиссии по государственной службе, злободневные проблемы имеют следующие характеристики:

- Трудно четко определить;
- Имеют много взаимозависимостей и часто имеют многоаспектный характер;
- Попытки решить злободневные проблемы часто приводят к непредвиденным последствиям;
- Часто не стабильная (движущаяся цель);
- Обычно не имеют четкого решения;
- Социально-сложные (не просто технически сложные);
- Вряд ли когда-либо находятся в рамках ответственности какой-либо одной организации;
- Предполагает изменение моделей поведения;
- Некоторые из них характеризуются постоянным провалом реализации политики.¹⁰⁵

Многие полагают, что правительства сталкиваются с политическими проблемами, которые, в основном, носят злободневный характер.

Джошуа Ньюман и Брайан Хед утверждают, что «фокусирование внимания только на сборе данных и трансфере знаний не может быть эффективным методом решения злободневных проблем», потому что злободневные проблемы имеют как технические, так и социальные особенности.¹⁰⁶ Больше данных и знаний может помочь решить

104 Там же, стр. 161-163.

105 Australian Public Service Commission, “Tackling wicked problems: A public policy perspective”. Доступно по <https://www.apsc.gov.au/tackling-wicked-problems-public-policy-perspective> (accessed on 8 January 2019).

106 Joshua Newman and Brian W. Head, “Wicked tendencies in policy problems: Rethinking the distinction between social and technical problems”, *Policy and Society*, vol. 36, no. 3 (2017), pp. 414-429.

технические, но не социальные особенности злободневных проблем. По словам Брайана Хэда, для «укрощения» злободневных проблем требуются не только более глубокие знания, но и улучшение процесса консультаций и более эффективное взаимодействие со сторонними партнерами.¹⁰⁷

Комиссия по государственной службе Австралии подчеркивает важность повышения квалификации и потенциала должностных лиц для решения злободневных проблем. Эти навыки включают умение работать за пределами организационных рамок, привлечение заинтересованных сторон и влияние на поведение граждан, в дополнение, к аналитическим, концептуальным компетенциям и знаниям управления проектами.¹⁰⁸

2.2.3 Администрирование, ориентированное на результаты

Администрирование, ориентированное на результаты – это еще один способ, с помощью которого данные могут улучшить управление на национальном уровне.

Группа Организации Объединенных Наций по вопросам развития (ГООНРВ) определяет администрирование, ориентированное на результаты, как: «Стратегия управления, с помощью которой все участники, прямо или косвенно способствуя достижению ряда результатов, обеспечивают условия того, чтобы их процессы, продукты и услуги способствовали достижению желаемых результатов (результаты, итоги и цели или воздействие более высокого уровня)».¹⁰⁹ В администрировании, ориентированном на результаты, участники «используют информацию и доказательства по фактическим результатам для информирования лиц, принимающих решения о разработке, выделении ресурсов и реализации программ и мероприятий, а также для учета и отчетности».¹¹⁰

Для Организации Объединенных Наций по вопросам образования, науки и культуры (ЮНЕСКО) основная цель подхода к администрированию, ориентированного на конкретные результаты, заключается в следующем: «Создание и использование информации об эффективности для отчетности перед заинтересованными сторонами извне и для принятия решений».¹¹¹ Кроме того, администрирование, основанное на результатах, позволяет лицам, принимающим решения, и исполнителям «принимать обоснованные решения, иметь возможность учиться на своих успехах или неудачах и делиться этим опытом со своими коллегами и всеми другими заинтересованными сторонами».¹¹²

107 Brian Head, "Wicked Problems in Public Policy", *Public Policy*, vol. 3, no. 2 (January 2008), p. 114. Доступно по https://www.researchgate.net/publication/43502862_Wicked_Problems_in_Public_Policy.

108 Australian Public Service Commission, "Tackling wicked problems: A public policy perspective". Доступно по <https://www.apsc.gov.au/tackling-wicked-problems-public-policy-perspective> (accessed on 8 January 2019).

109 UNDG, "Results-Based Management Handbook: Harmonizing RBM concepts and approaches for improved development results at country level", October 2011, p. 2. Доступно по <https://undg.org/wp-content/uploads/2016/10/UNDG-RBM-Handbook-2012.pdf>.

110 Там же

111 UNESCO, "Results-Based Programming, Management, Monitoring and Reporting (RBM) approach as applied at UNESCO: Guiding Principles", September 2015, p. 8. Доступно по <http://unesdoc.unesco.org/images/0017/001775/177568E.pdf>.

112 Там же

Несмотря на различие в направленности, как ГОООНВР, так и ЮНЕСКО (а также другие учреждения Организации Объединенных Наций) считают администрирование, ориентированное на результаты, «основанной на объективной информации и, таким образом, делают упор на критическое расследование, мониторинг и оценку и принятие решений на основе фактов».¹¹³

Ключевые этапы администрирования, ориентированного на конкретные результаты, включают:

- Определение четких и измеримых целей (результатов) с помощью логической структурной основы;
- Выбор показателей, которые будут использоваться для измерения прогресса в достижении каждой цели;
- Установка конкретных целей для каждого показателя, используемого для оценки эффективности;
- Разработка систем мониторинга производительности для регулярного сбора данных о реальных результатах;
- Просмотр, анализ и представление фактических результатов по отношению к целям;
- Интеграция оценок для предоставления дополнительной информации о производительности, которая не всегда доступна в системах мониторинга производительности;
- Использование информации о производительности для внутренней подотчетности руководства, процессов обучения и принятия решений, а также для отчетности о результатах деятельности перед заинтересованными сторонами и партнерами.¹¹⁴

В контексте данной дискуссии важно помнить, что ключевой особенностью администрирования, ориентированного на результаты, является измерение эффективности – процесс объективного измерения того, насколько хорошо ведомство выполняет свои заявленные цели или задачи.

Администрирование, ориентированное на результаты, является одним из пяти программных принципов, используемых Организацией Объединенных Наций в качестве «отправной точки и руководства для анализа на всех этапах Рамочной программы Организации Объединенных Наций по оказанию помощи в целях развития».¹¹⁵ Оно также используется всеми учреждениями Организации Объединенных Наций для повышения эффективности своей работы. Донорские агентства, такие как USAID (Соединенные Штаты Америки), DFID (Великобритания), AustralianAID (Австралия), CIDA (Канада), Данида (Дания), SIDA (Швеция) и Всемирный банк, продвигают использование администрирования, ориентированного на результаты.¹¹⁶ Азиатский банк

113 United Nations Joint Inspection Unit, *Results-Based Management in the United Nations System: High-Impact Model For Results-Based Management – Benchmarking framework, stages of development and outcomes* (Geneva, United Nations, 2017). Доступно по https://www.unjiu.org/sites/www.unjiu.org/files/jiu_note_2017_1_english_0.pdf.

114 OECD, “Results Based Management in the Development Co-Operation Agencies: A Review of Experience (Executive Summary)”, February 2000, p. 4. Доступно по <https://www.oecd.org/development/evaluation/dcdndep/31950681.pdf>.

115 UNDG, “Guidance Note: Application of the Programming Principles to the UNDAF”, January 2010, p. 1. Доступно по <https://undg.org/wp-content/uploads/2016/09/Five-Programming-Principles.pdf>.

116 Annette Binnendijk, “Results Based Management in the Development Co-Operation Agencies: A Review of Experience

развития поддерживает институционализацию администрирования, ориентированного на результаты, в рамках филиппинского бюрократического аппарата.¹¹⁷

Администрирование, ориентированное на результаты, тесно связано с **управлением, ориентированном на достижение результатов в области развития** (MfDR – Managing for Development Results) – «стратегия управления, которая фокусируется на использовании информации о производительности для улучшения процесса принятия решений».¹¹⁸ MfDR смещает акцент управления с деятельности на измеримые результаты также, как и администрирование, ориентированное на результаты.

Принципы MfDR:

- На всех этапах – от стратегического планирования до реализации и после завершения – акцентируйте диалог на результатах для стран-партнеров, агентств по развитию и других заинтересованных сторон;
- Согласуйте фактическую деятельность по исполнению программ, мониторингу и оценке с согласованными ожидаемыми результатами;
- Сохраняйте систему отчетов о результатах как можно более простой, экономичной и удобной для пользователя;
- Управляйте не по результатам, а для достижения результатов путем организации ресурсов;
- Используйте информацию о результатах для обучения управлению и принятию решений, а также для отчетности и учета.¹¹⁹

Цикл MfDR включает в себя пять основных этапов:

1. Постановка целей и согласование задач и стратегий;
2. Выделение доступных ресурсов на мероприятия, которые будут способствовать достижению желаемых результатов;
3. Мониторинг и оценка на эффективность выделенных ресурсов для получения намеченной разницы;
4. Подготовка отчета о результатах деятельности для общественности;
5. Подача информации обратно в процесс принятия решений.¹²⁰

Чем MfDR похоже на администрирование, ориентированное на результаты?

По данным ПРООН, MfDR применяет те же базовые концепции администрирования, ориентированного на результаты – «хорошее планирование, мониторинг, оценка,

(Executive Summary)”, *Development Assistance Committee Working Party on Aid Evaluation*, February 2000. Доступно по <https://www.oecd.org/development/evaluation/dcdndep/31950681.pdf>.

117 Asian Development Bank, *Results-Based Management Framework in the Philippines: A Guidebook* (Mandaluyong City, 2013). Доступно по <https://www.adb.org/publications/results-based-management-framework-philippines-guidebook>.

118 Managing for Development Results. Доступно по <http://www.mfdr.org/1about.html> (accessed on 8 January 2019).

119 OECD and World Bank, “Part 1. MfDR Concepts, Tools and Principles”, in *Emerging Good Practice in Managing for Development Results: Sourcebook*, first edition (2006), pp. 10-14. Доступно по <http://www.mfdr.org/Sourcebook/1stEdition/4-MfDRPrinciples.pdf>.

120 OECD, “Managing for Development Results: Information Sheet”, September 2008. Доступно по <http://www.mfdr.org/About/Final-MfDR-information-sheet.pdf>.

обучение и обратная связь с планированием, - но при этом стремится сосредоточить внимание на помощи в целях развития, демонстрируя реальные и значимые результаты». Таким образом, «MfDR - это управление, ориентированное на результаты, в действии, но оно больше ориентировано на внешнюю среду и результаты, которые важны для стран, в которых осуществляются программы, а не на внутреннюю работу ведомства».¹²¹

Хотя разработка политики на основе фактов и администрирование, ориентированное на результаты, способствуют управлению на основе данных на национальном уровне, эти подходы одинаково полезны для управления на местном и глобальном уровнях.

2.3 Проблемы в управлении на основе данных: пример целей ООН в области устойчивого развития

Повестка дня в области устойчивого развития на период до 2030 года, включающая цели в области устойчивого развития (ЦУР), является самой последней попыткой глобального управления, которая подчеркивает важность данных.

Это соглашение, принятое в сентябре 2015 года, направлено на то, чтобы покончить с бедностью, защитить планету и обеспечить благосостояние.¹²² Повестка дня в области устойчивого развития на период до 2030 года предусматривает 17 целей, 169 задач и 232 показателя, которые должны быть достигнуты к 2030 году. Целями являются:

- Цель 1. Нет бедности
- Цель 2. Нулевой показатель голода
- Цель 3. Хорошее здоровье и благополучие
- Цель 4. Качественное образование
- Цель 5. Гендерное равенство
- Цель 6. Чистая вода и санитария
- Цель 7. Доступная и экологически чистая энергия
- Цель 8. Достойный труд и экономический рост
- Цель 9. Промышленность, инновации и инфраструктура
- Цель 10. Сокращение неравенства
- Цель 11. Устойчивые города и сообщества
- Цель 12. Ответственное потребление и производство
- Цель 13. Действия в защиту климата
- Цель 14. Жизнь под водой

121 UNDP, *Handbook on Planning, Monitoring and Evaluating for Development Results* (New York, 2009), p. 6. Доступно по <http://web.undp.org/evaluation/handbook/documents/english/pme-handbook.pdf>.

122 UNDP, "World leaders adopt Sustainable Development Goals", 25 September 2015. Доступно по <http://www.undp.org/content/undp/en/home/presscenter/pressreleases/2015/09/24/undp-welcomes-adoption-of-sustainable-development-goals-by-world-leaders.html>.

- Цель 15. Жизнь на Земле
- Цель 16. Мир, справедливость и сильные институты
- Цель 17. Партнерство для достижения целей.¹²³

Важно отметить, что ЦУР подчеркивают важность мониторинга результатов развития, а также важность данных и статистики для устойчивого развития.¹²⁴ В частности, задача ЦУР № 17.18 направлена на «значительное увеличение доступности качественных, своевременных и надежных данных в разбивке по доходу, полу, возрасту, расе, этнической принадлежности, миграционному статусу, инвалидности, географическому положению и другим характеристикам, имеющим отношение к национальным условиям». Задача ЦУР № 17.19 призывает к новым показателям прогресса в области устойчивого развития помимо показателя валового внутреннего продукта.

Страны-члены ООН договорились следить за механизмами реализации ЦУР. Они согласились «быть точными, опираться на факты, основываться на данные страны, которые являются высококачественными, доступными, своевременными, надежными и дезагрегированными по доходу, полу, возрасту, расе, этнической принадлежности, миграционному статусу, инвалидности и географическому положению и другим характеристикам, соответствующим национальным условиям».¹²⁵

Кроме того, в целях обеспечения эффективной реализации Генеральная Ассамблея Организации Объединенных Наций в июле 2017 года приняла показатели ЦУР – «надежную систему показателей и статистических данных для мониторинга прогресса, разработки политики и обеспечения подотчетности всех заинтересованных сторон». ЦУР имеют 232 показателя для измерения прогресса.

Однако сбор данных и методология, используемая для мониторинга показателей ЦУР, являются основными проблемами.

Показатели ЦУР разбиты на категории с использованием трехуровневой системы, основанной на методологии и данных:

- Уровень 1 – Существует установленная методология, и данные широко доступны;
- Уровень 2 – Существует установленная методология, но данные не всегда доступны;
- Уровень 3 – Согласованная на международном уровне методология еще не разработана.¹²⁶

По состоянию на 2017 год только треть показателей имеют доступные данные, которые можно использовать для мониторинга ЦУР, около четверти имеют методологию, но не

123 UNDP, “Sustainable Development Goals”, n.d. Доступно по https://www.undp.org/content/dam/undp/library/corporate/brochure/SDGs_Booklet_Web_En.pdf.

124 UNDP, “Guidance Note: Data for Implementation and Monitoring of the 2030 Agenda for Sustainable Development”, September 2017, p. 3. Доступно по <http://www.undp.org/content/undp/en/home/librarypage/poverty-reduction/guidance-note--data-for-implementation-and-monitoring-of-the-203.html>.

125 Там же, стр. 6.

126 Там же, стр. 9.

имеют данных, а 38 процентов не имеют согласованной методологии и данных.¹²⁷

Как отмечалось в ПРООН: «Благодаря не менее 232 показателям и подсчетам, интегрированной и неделимой структуре и основополагающему принципу «никого не забыть» Повестка дня на период до 2030 года на порядок увеличивает масштаб и объем данных, необходимых для осуществления и мониторинга устойчивого развития, являющейся непростой задачей даже для стран с наилучшими статистическими возможностями».¹²⁸

Хорошей новостью является то, что развивающиеся страны «усилили свои возможности по сбору данных и их обработке».¹²⁹ На основе Статистического показателя Всемирного банка страны с низким и средним уровнем дохода улучшили показатели со среднего значения 65,3 в 2004 году до 68,8 в 2015 году.¹³⁰

В Азии проблемы с данными, с которыми сталкиваются правительства, включают следующее:

- Необходимость укрепления национальных статистических систем для отчетности по ЦУР и поощрения нововведений в сборе данных для упрощения достижения поставленных задач;
- Необходимость для стран выбора показателей, которые являются наиболее значимыми для их приоритетов реализации (все 232 будут актуальны не во всех случаях);
- Устранение мониторинга, которое обусловлено наличием данных, а не их актуальностью для национальных приоритетов.¹³¹

Данные могут помочь правительствам в достижении ЦУР. С революцией в области данных «выработка политики может быть улучшена за счет использования огромных потоков точных, своевременных и детальных данных».¹³²

Это также тот случай, когда:

Данные и зачастую хорошие данные имеются в подавляющем большинстве развивающихся стран, но их все еще необходимо улучшать многими способами. Необходимо обеспечить условия для своевременного сбора уместных данных, данных хорошего качества.¹³³

127 OECD, *Development Co-operation Report 2017: Data for Development* (Paris, 2017), p. 24. Доступно по https://read.oecd-ilibrary.org/development/development-co-operation-report-2017_dcr-2017-en#page25.

128 UNDP, "Guidance Note: Data for Implementation and Monitoring of the 2030 Agenda for Sustainable Development", September 2017, p. 4. Доступно по <http://www.undp.org/content/undp/en/home/librarypage/poverty-reduction/guidance-note--data-for-implementation-and-monitoring-of-the-203.html>.

129 Neil Webster and Helle Munk Ravnborg, "Monitoring the implementation of the Sustainable Development Goals – The role of the data revolution", *European Parliament Directorate-General for External Policies*, July 2016. Доступно по [http://www.europarl.europa.eu/RegData/etudes/STUD/2016/578020/EXPO_STU\(2016\)578020_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/2016/578020/EXPO_STU(2016)578020_EN.pdf).

130 Там же, стр. 11.

131 Asian Development Bank, "From Goals to Action: Implementing the Sustainable Development Goals – Seminar Background Note", 2017, p. 6. Доступно по <https://www.adb.org/sites/default/files/publication/301696/goals-action-sdgs.pdf>.

132 OECD, *Development Co-operation Report 2017: Data for Development* (Paris, 2017), p. 27. Доступно по https://read.oecd-ilibrary.org/development/development-co-operation-report-2017_dcr-2017-en#page27.

133 Neil Webster and Helle Munk Ravnborg, "Monitoring the implementation of the Sustainable Development Goals –

Несмотря на стремление к более широкому использованию данных в управлении, опрос государственных должностных лиц и специалистов-практиков в области развития из 126 стран с низким и средним уровнем дохода, проведенный в 2017 году, показал, что «лидеры больше используют данные или анализ для проведения ретроспективных оценок прошлых результатов, чем для разработки будущих политик и программ».¹³⁴

Медленное принятие управления на основе данных можно объяснить техническими и политическими причинами.

Технические проблемы включают в себя: отсутствие стандартов данных, плохое качество данных, отсутствие системы взаимодействия и плохое управление информацией.¹³⁵

Тем не менее, «бюрократическая политика» является одним из главных препятствий для ее широкого и быстрого внедрения. В мире, где информация – это сила, все еще есть чиновники, которые не хотят модернизировать сбор и производство данных и/или боятся делиться/публиковать данные, опасаясь их потери.

К счастью, крепкие семена для внедрения управления на основе данных уже посеяны. Просвещенные должностные лица, высокий спрос со стороны групп гражданского общества и давление со стороны международных агентств требуют совершенствования сбора данных, обработки и обмена публичными данными и их использования в государственном управлении.¹³⁶

The role of the data revolution”, *European Parliament Directorate-General for External Policies*, July 2016, p. 33.

Доступно по [http://www.europarl.europa.eu/RegData/etudes/STUD/2016/578020/EXPO_STU\(2016\)578020_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/2016/578020/EXPO_STU(2016)578020_EN.pdf).

134 Такааки Масакиси и другие, “Decoding Data Use: How do leaders source data and use it to accelerate development? – Executive Summary”, *AidData*, n.d., p.1. Доступно по http://docs.aiddata.org/ad4/pdfs/Decoding_data_use--Executive_summary.pdf.

135 Kathleen Hickey, “What’s really needed for data-driven government”, *GCN*, 22 December 2016. Доступно по <https://gcn.com/articles/2016/12/22/data-governance-challenges-solutions.aspx>.

136 Ben Rossi, “Data-driven government: Oxymoron or reality?” *Information Age*, 14 January 2016. Доступно по <http://www.information-age.com/data-driven-government-oxymoron-or-reality-123460782/>.

3. Мир переполнен данными

Задачи данного раздела:

- Дать определение официальной статистики;
 - Обсудить понятие больших данных;
 - Объяснить, что такое данные в реальном времени;
 - Рассмотреть небольшие данные;
 - Изучить данные, генерируемые гражданами.
-

Несмотря на поток данных, который испытывает мир, высококачественные данные для принятия решений остаются проблемой в развивающихся странах. Согласно отчету *«Мир, который считает»*: «Слишком много стран по-прежнему имеют плохие данные, данные поступают слишком поздно, и слишком много проблемных вопросов все еще едва охватываются существующими данными». ¹³⁷

В этом разделе мы рассмотрим различные источники данных для управления на основе данных.

3.1 Официальная статистика

Несмотря на растущий интерес к большим данным, традиционные источники данных, такие как официальная статистика, остаются важными.

Официальная статистика - это «числовые наборы данных, производимые официальными государственными органами, в основном, для административных целей». ¹³⁸ Это «подкатегория «научных данных», которая более точно помогает понять, как общества функционируют и развиваются». ¹³⁹

Целью официальной статистики, согласно Европейской экономической комиссии Организации Объединенных Наций, является: «Получение и распространение авторитетных результатов, предназначенных для надежного отражения экономически и социально значимых явлений сложной и динамичной реальности в определенной стране». ¹⁴⁰

137 United Nations Secretary-General's Independent Expert Advisory Group on the Data Revolution for Sustainable Development, "A World that Counts: Mobilizing the Data Revolution for Sustainable Development", November 2014, p. 11. Доступно по <http://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf>.

138 Sage Research Methods, "Official Statistics". Доступно по <http://methods.sagepub.com/book/key-concepts-in-social-research/n34.xml> (accessed 8 January 2019).

139 Walter J. Radermacher, "The Future Role of Official Statistics", Power from Statistics Outlook Report, p. 2. Доступно по https://www.researchgate.net/publication/320616460_The_Future_Role_of_Official_Statistics.

140 United Nations Economic Commission for Europe, "How Should a Modern National System of Official Statistics Look?" January 2008. Доступно по <https://www.unece.org/fileadmin/DAM/stats/documents/applyprinciples.e.pdf>.

Правительства используют официальную статистику при планировании, принятии решений и мониторинге или оценке политики.¹⁴¹ Частный сектор использует их для поддержки бизнес-решений. Граждане используют официальную статистику для оценки деятельности правительства и обеспечения его подотчетности.

В отличие от данных частного сектора, правительственные данные являются:

- **комплексными** – статистические ведомства стремятся охватить как можно большую часть населения;
- **последовательными** – деятельность статистических ведомств характеризуется долгосрочной направленностью, и они предоставляют согласованные определения ключевых показателей во времени, что необходимо для интерпретации самых последних оценок;
- **заслуживающими доверия** – статистические ведомства должны соответствовать требованиям прозрачности, установленным законодательством и международными организациями.¹⁴²

На последний пункт необходимо обратить внимание. Официальная статистика основана на общих принципах, стандартах, методологиях и технологиях, установленных в соответствии с профессиональным этическим кодексом.¹⁴³

Вставка 10. Основные принципы официальной статистики

Следующие десять принципов были приняты Экономическим и Социальным Советом Организации Объединенных Наций на его 46-м пленарном заседании 24 июля 2013 года:

1. Официальная статистика является необходимым элементом информационной системы демократического общества, обеспечивая правительство, экономические круги и общественность данными об экономическом, демографическом, социальном и экологическом положении. С этой целью официальные статистические данные, имеющие практическую ценность, подготавливаются и распространяются на объективной основе государственными статистическими ведомствами для обеспечения уважения права граждан на общественную информацию.
2. В целях сохранения доверия к официальной статистике статистические ведомства в соответствии со строго профессиональными соображениями, включая научные принципы и профессиональную этику, должны принимать решения в отношении методов и процедур сбора, обработки, хранения и представления статистических данных.
3. Для облегчения правильной интерпретации данных статистические ведомства должны

141 Statistics South Africa, "Purpose of official statistics, and statistical principles". Доступно по http://www.statssa.gov.za/?page_id=750 (accessed 8 January 2019).

142 Nicholas Eberstadt and others, "In Order That They Might Rest Their Arguments on Facts": The Vital Role of Government-Collected Data", *Hamilton Project and American Enterprise Institute*, March 2017, p. 4. Доступно по https://www.brookings.edu/wp-content/uploads/2017/02/thp_20170227_govt_collected_data_report.pdf.

143 Walter J. Radermacher, "The Future Role of Official Statistics", *Power from Statistics Outlook Report*, p. 2. Доступно по https://www.researchgate.net/publication/320616460_The_Future_Role_of_Official_Statistics.

предоставлять информацию в соответствии с научными стандартами в отношении источников, методов и процедур в области статистики.

4. Статистические ведомства имеют право комментировать неверную интерпретацию или неправильное использование статистических данных.
5. Данные для статистических целей могут собираться из всех видов источников, будь то статистические обследования или административная отчетность. Статистические ведомства должны выбирать источник с учетом качества, своевременности, затрат и нагрузки, которая ложится на респондентов.
6. Личные данные, собираемые статистическими ведомствами для подготовки статистической информации, независимо от того, относятся ли они к физическим или юридическим лицам, должны носить строго конфиденциальный характер и использоваться исключительно для статистических целей.
7. Законы, нормы и меры, в рамках которых функционируют статистические системы, должны предаваться гласности.
8. Для обеспечения согласованности и эффективности в статистической системе необходимо осуществлять координацию деятельности статистических ведомств на уровне стран.
9. Использование статистическими ведомствами в каждой стране международных концепций, классификаций и методов способствуют обеспечению согласованности и эффективности статистических систем на всех официальных уровнях.
10. Двустороннее и многостороннее сотрудничество в области статистики содействует улучшению систем официальной статистики во всех странах.

Взято (с изменениями) из: *Resolution adopted by the Economic and Social Council on 24 July 2013 (E/RES/2013/21)*. Доступно по <https://unstats.un.org/unsd/dnss/gp/FP-Rev2013-E.pdf>.

Традиционно официальная статистика состоит из данных переписи, данных опросов и административных данных.

Переписью является: «Подсчет для официальных целей, особенно для учета количества людей, живущих в стране, и для получения такой информации, как возраст, пол, раса и т. д.»¹⁴⁴ Переписью населения является: «Операция, которая через регулярные интервалы времени производит официальный подсчет (или оценку) населения на территории страны и в ее самых мелких географических районах вместе с информацией по ряду демографических и социальных характеристик всего населения».¹⁴⁵

Основные особенности, которые делают перепись уникальной:

¹⁴⁴Cambridge Dictionary, "Census". Доступно по <https://dictionary.cambridge.org/us/dictionary/english/census> (accessed on 8 January 2019).

¹⁴⁵ United Nations Economic Commission for Europe, *Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing* (New York and Geneva, 2015), p. 5. Доступно по https://www.unece.org/fileadmin/DAM/stats/publications/2015/ECECES41_EN.pdf.

- **Индивидуальный учет** – Информация о каждом учтенном лице регистрируется таким образом, что их характерные данные могут быть записаны отдельно;
- **Одновременность** – Полученная информация должна относиться к четко определенному и уникальному отчетному периоду (или конкретному моменту времени);
- **Универсальность** – Подсчет (или сравнительный анализ) населения должен включать каждого человека, проживающего и / или присутствующего на определенной территории страны в определенный особый момент времени;
- **Данные о малых районах** – В ходе переписи должны быть получены данные о количестве и особенностях населения, связанного с мелкими географическими районами страны, и небольших подгрупп населения в соответствии с первостепенным требованием сохранения конфиденциальности индивидуальной информации;
- **Определенная периодичность** – Перепись должна проводиться через регулярные промежутки времени, чтобы сравнительная информация предоставлялась в установленной последовательности. Рекомендуется проводить перепись по меньшей мере один раз каждые десять лет.¹⁴⁶

Следующее определяет содержание переписи:

- Потребность в данных на национальном и местном уровнях;
- Наличие данных из других статистических источников;
- Ограничения переписи для сбора данных, когда (по крайней мере, для традиционных переписей) можно задавать только ограниченное количество вопросов по отдельным темам, а деликатные или более сложные темы, требующие дополнительных разделов и специализированной подготовки интервьюеров, можно охватить только до ограниченной степени.¹⁴⁷

Данные опроса поступают от проведения выборочного обследования – «изучения подгруппы населения с целью оценки характеристик людей».¹⁴⁸ «Опрос обычно учитывает меньшую долю от общей численности населения»,¹⁴⁹ а перепись – это полный подсчет. Данные опросов обеспечивают субнациональную оценку определенной характеристики населения и собираются из-за быстрых изменений в численности и других свойствах населения и растущего спроса на дополнительные подробные данные о социальных, экономических и жилищных особенностях, которые не подходят для сбора данных во время переписи.¹⁵⁰

Административные данные могут быть определены как «данные, собранные из источников, не относящихся к статистическим управлениям»,¹⁵¹ и обычно собираются

146 Там же, стр. 6-7.

147 Там же, стр. 7-8.

148 Stat Trek, "Statistics Dictionary: Sample Survey". Доступно по <http://stattrek.com/statistics/dictionary.aspx?definition=sample%20survey> (accessed on 8 January 2019).

149 United Nations Department of Economic and Social Affairs, *Handbook on Population and Housing Census Editing: Revision 1* (New York, 2010) p. 1. Доступно по https://unstats.un.org/unsd/publication/SeriesF/seriesf_82rev1e.pdf.

150 United Nations Economic Commission for Europe, *Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing* (New York and Geneva, 2015), p. 11. Доступно по https://www.unece.org/fileadmin/DAM/stats/publications/2015/ECECES41_EN.pdf.

151 United Nations Economic Commission for Europe, *Using Administrative and Secondary Sources for Official*

на регулярной основе по всем лицам, участвующим в конкретной программе. Административные данные предлагают возможности долгосрочного исследования.¹⁵² По сравнению с результатами опроса в административных данных содержится меньше недостатков, связанных с убылью респондентов, неполучением ответов и занижением информации в данных опроса.

Источниками административных данных являются:

- Налоговые данные (налог на доходы физических лиц, налог на добавленную стоимость, налог на бизнес/прибыль, налоги на имущество, пошлины на ввоз/вывоз);
- Данные социального обеспечения (взносы, пособия, пенсии);
- Данные системы здравоохранения/образования;
- Системы регистрации лиц/предприятий/имущества/транспортных средств;
- Удостоверения личности/паспорта/водительские права;
- Списки избирателей;
- Реестр хозяйств;
- Реестры местных советов;
- Разрешения на строительство;
- Системы лицензирования (телевидение, продажа запрещенных предметов);
- Опубликованные бизнес-счета;
- Внутренние данные бухгалтерского учета, которыми располагают предприятия;
- Частные предприятия с массивами данных (кредитные агентства, бизнес-аналитики, коммунальные компании, телефонные справочники, предприятия розничной торговли с карточками покупателей).¹⁵³

Статистическое использование административных данных включает в себя:

- Использование выборочных совокупностей, непосредственно в качестве выборки или дополнения/обновления существующей выборки;
- Замещение процесса сбора данных (использование налоговых данных для малых предприятий вместо поиска материалов их обследования);
- Использование при редактировании и выполнении условных расчетов;
- Прямое составление таблиц;
- Косвенное использование в оценке (в качестве вспомогательной информации при проверке оценок, сравнительном анализе или процессе стандартизации отчетных периодов);
- Оценка обследования, включая сопоставление данных (сравнение оценок обследования с оценками соответствующей административной программы).¹⁵⁴

Преимущества использования административных данных: снижение затрат, снижение

Statistics: A Handbook of Principles and Practices (New York and Geneva, 2011) p. 2.

152 Robert Doar and Linda Gibbs, "Unleashing the Power of Administrative Data: A Guide for Federal, State, and Local Policymakers", *Results for America and American Enterprise Institute*, October 2017, p. 5. Доступно по <https://results4america.org/wp-content/uploads/2017/10/Unleashing-the-Power-of-Administrative-Data.pdf>.

153 United Nations Economic Commission for Europe, *Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices* (New York and Geneva, 2011) p. 4.

154 Statistics Canada, "Use of administrative data", 16 June 2017. Доступно по <http://www.statcan.gc.ca/pub/12-539-x/2009001/administrative-administratives-eng.htm>.

нагрузки на поставщиков данных, частота, охват, своевременность и улучшение информированности общественности.¹⁵⁵

Важно отметить, что административные данные не предназначены для замены данных опросов в официальной статистике. Оба имеют значение для управления на основе данных. Как отмечают Доар и Гиббс: «Возможность связывать административные наборы данных друг с другом и данными опросов дает значительный потенциал для ответа на важные вопросы, которые ни один тип данных не может ответить самостоятельно».¹⁵⁶

Официальная статистика остается играть важную роль даже в разгар революции в использовании данных. Исследование AidData 2017 года о том, как ведущие страны используют данные, показало, что «лидеры наиболее часто используют национальные статистические данные и данные оценки, а также считают их наиболее полезными источниками данных о состоянии развития».¹⁵⁷

Несмотря на попытки быть всеобъемлющими, отсутствуют официальные данные по таким важным вопросам, как гендер.

В отчете «*Мир, который считает*» отмечается, что «гендерное неравенство и недооценка деятельности и приоритетов женщин во всех сферах были воспроизведены в статистических данных».¹⁵⁸

Отсутствие статистических данных по гендерным вопросам вызывает сожаление, поскольку «именно благодаря сбору, производству, анализу и использованию гендерных статистических данных лица, определяющие политику, и специалисты по развитию могут начать должным образом решать конкретные проблемы женщин и мужчин».¹⁵⁹ Более того, как отмечает Эмили Коури Прайор: «Когда мы не в состоянии измерить критические аспекты жизни женщин и девушек – от их экономического вклада до степени доступа к жизненно важным услугам – мы недооцениваем их роль и опыт в обществе, укрепляя предубеждения и, в конечном счете, выпускаем из поля зрения проблемы женщин и девушек».¹⁶⁰

Вставка 11. Проблемы количественной оценки женского труда

155 United Nations Economic Commission for Europe, *Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices* (New York and Geneva, 2011), pp. 7-10.

156 Robert Doar and Linda Gibbs, “Unleashing the Power of Administrative Data: A Guide for Federal, State, and Local Policymakers”, *Results for America and American Enterprise Institute*, October 2017, p. 4. Доступно по <https://results4america.org/wp-content/uploads/2017/10/Unleashing-the-Power-of-Administrative-Data.pdf>.

157 Takaaki Masaki and others, “Decoding Data Use: How do leaders source data and use it to accelerate development? – Executive Summary”, *AidData*, n.d., p.1. Доступно по http://docs.aiddata.org/ad4/pdfs/Decoding_data_use--Executive_summary.pdf.

158 United Nations Secretary-General’s Independent Expert Advisory Group on the Data Revolution for Sustainable Development, “A World that Counts: Mobilizing the Data Revolution for Sustainable Development”, November 2014, p. 14. Доступно по <http://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf>.

159 United Nations Department of Economic and Social Affairs, “Using data to measure gender equality”, 4 November 2014. Доступно по <http://www.un.org/en/development/desa/news/gender/using-data-to-measure-gender-equality.html>.

160 Emily Courey Pryor, “A World That Counts’ Everyone, Including Women and Girls”, *United Nations World Data Forum*, 27 December 2016. Доступно по <https://undataforum.org/WorldDataForum/a-world-that-counts-everyone-including-women-and-girls>.

Женщинам часто не платят за работу, которую они выполняют. Каждый день за ними числится в среднем 4,5 часа бесплатного труда по дому и по уходу за детьми независимо от того, где они находятся в мире. В развивающихся странах женщины могут работать в 10 раз больше мужчин.

Если часы неоплачиваемого труда женщин во всем мире оплачивались бы по минимальной заработной ставке, их стоимость составила бы как минимум на 10 триллионов долларов США больше, чем валовой внутренний продукт Китая, согласно самой скромной оценке McKinsey.

Но это только оценки. По правде говоря, у нас нет достаточных данных о большей части работы, выполняемой женщинами во всем мире. Те же самые общество и структуры, которые собирают и применяют данные по всем аспектам нашей жизни, не могут эффективно отслеживать часы работы женщин и их оплаты за пределами официальных рабочих мест.

Папа Сек, главный статистик структуры «ООН-женщины», который возглавляет проект по цифрам и фактам для гендерного равенства, заявил в интервью, что обследования занятости часто основаны на старых экономических теориях, которые предполагают, что только муж может предоставить информацию о домашнем хозяйстве.

Когда задаются вопросы о собственности или работе, принимаются во внимание только его ответы, поэтому работа женщины не может быть учтена. Даже когда муж и жена отвечают на обследование о занятости, вопросы могут быть недостаточно детализированы, чтобы охватить работу, которая не приносит заработную плату. В других случаях данные агрегируются без учета пола.

Взято (с изменениями) из: Annalisa Merelli, "There's a mind-boggling amount of work women do that we literally can't quantify", Quartz, 18 May 2016. Доступно по [at https://qz.com/686075/we-still-have-literally-no-way-to-quantify-exactly-how-much-work-women-do/](https://qz.com/686075/we-still-have-literally-no-way-to-quantify-exactly-how-much-work-women-do/).

Несмотря на свои недостатки, официальная статистика остается важной для расширения прав и возможностей людей:

Качественная статистика укрепляет демократию, предоставляя гражданам доступ к ключевой информации, которая повышает ответственность. Доступ к достоверной статистике является фундаментальным «правом», которое позволяет выбирать и принимать решения на основе информации. Без статистики не может быть прочной и широкой демократии.¹⁶¹

3.2 Большие данные

Распространенное определение больших данных: «Наборы данных, размер которых превышает возможности обычных программных средств по администрированию баз данных для сбора, хранения, управления и анализа».¹⁶² Некоторые отмечают, что это

¹⁶¹ Walter J. Radermacher, "The Future Role of Official Statistics", Power from Statistics Outlook Report, p. 3. Доступно по https://www.researchgate.net/publication/320616460_The_Future_Role_of_Official_Statistics.

¹⁶² James Manyika and others, "Big Data: The next frontier for innovation, competition, and productivity", McKinsey Global Institute, May 2011. Доступно по

неустойчивое определение (например, то, что обычно сегодня, может быть нетипичным завтра).

Другое важное определение: «Информационный актив, характеризуемый тремя V (volume-объем, variety-разнообразие и velocity-скорость)».¹⁶³ В этом определении «объем» относится к количеству данных, «разнообразию» относится к числу типов данных, а «скорость» относится к скорости обработки данных. С этой точки зрения, большие данные – это не только невероятный объем данных, но также широкое разнообразие и быстрая скорость генерирования данных.

Другие предлагают, что существует не только три, а пять V больших данных,¹⁶⁴ добавляя «veracity-правдивость» (надежность или достоверность данных) и «value-значение» (выгода или ценность) к объему, разнообразию и скорости.

Шумиха, связанная с большими данными, привела к критическому определению: «Культурный, технологический и научный феномен, основанный на взаимодействии: - (1) Технологии: максимизация вычислительных мощностей и алгоритмической точности для сбора, анализа, связывания и сравнения больших наборов данных; (2) Анализа: использование больших наборов данных для выявления закономерностей с целью предъявления экономических, социальных, технических и юридических требований; и (3) Мифологии: широко распространенное убеждение, что большие наборы данных предлагают более высокую форму интеллекта и знаний, которые могут генерировать идеи, которые ранее были невозможны, под аурой правды, объективности и точности».¹⁶⁵

Еще один способ рассмотрения феномена больших данных - через их возможности. 10 общих характеристик больших данных:

1. **Большие** – Большие наборы данных являются средством для достижения цели, они не являются самоцелью;
2. **Всегда работают** – Непрерывность больших данных позволяет изучать неожиданные события и проводить измерения в режиме реального времени;
3. **Инертные** – Измерения в источниках больших данных с меньшей вероятностью изменят поведение;
4. **Незавершенные** – Независимо от того, насколько велики ваши данные, в них, вероятно, нет нужной вам информации;
5. **Недоступные** – Исследователям сложно получить доступ к данным, хранящимся в компаниях и правительствах;
6. **Нерепрезентативные** – Непредставительные данные плохи для обобщений вне

https://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_exec_summary.ashx.

¹⁶³ Margaret Rouse, "Definition: 3Vs (volume, variety and velocity)", *WhatIs.com*, February 2013. Доступно по <http://whatis.techtarget.com/definition/3Vs>.

¹⁶⁴ Bernard Marr, "Big Data: The 5 Vs Everyone Must Know", *LinkedIn*, 6 March 2014. Доступно по <https://www.linkedin.com/pulse/20140306073407-64875646-big-data-the-5-vs-everyone-must-know>.

¹⁶⁵ danah boyd and Kate Crawford "Critical Questions For Big Data: Provocations for a cultural, technological, and scholarly phenomenon", *Information, Communication & Society*, vol. 15, no. 5 (June 2012), p. 663. Доступно по https://people.cs.kuleuven.be/~bettina.berendt/teaching/ViennaDH15/boyd_crawford_2012.pdf.

выборки, но могут быть весьма полезны для сравнений внутри выборки;

7. **Дрейфующие** – Миграция населения, изменение применения и тенденции развития систем затрудняют использование источников больших данных для изучения долгосрочных тенденций;
8. **Алгоритмически запутанные** – Поведение в системах больших данных не является естественным, оно обусловлено инженерными целями систем;
9. **Загрязненные** – Источники больших данных могут быть перегружены хламом и спамом;
10. **Конфиденциальные** – Часть информации, которой располагают компании и правительства, является конфиденциальной.¹⁶⁶

Источниками больших данных являются результаты интернет-поиска, транзакции по кредитным картам, электронные письма (Gmail использует алгоритмы для сканирования и последующего анализа содержимого электронных писем), публикации в социальных сетях (Facebook собирает и использует всю информацию, которая создается при использовании его платформы), смарт-часы и трекеры активности (которые измеряют ваши биометрические показатели, такие как частота сердечных сокращений, переменность сердечного ритма и температура тела). Правительства, а также университеты и исследовательские институты также производят наборы данных, которые можно использовать для анализа.

Также некоторые сторонники определяют большие данные как:

- **Исчерпывающие данные** – Пассивно собираемые данные об использовании людьми цифровых услуг, таких как мобильные телефоны, финансовые транзакции или результаты интернет-поисков.
- **Данные зондирования** – Активно собираемые данные с датчиков, например, в умных городах или от носимых устройств, а также с помощью дистанционного зондирования и спутниковых снимков.
- **Цифровой контент** – Открытый интернет-контент, активно создаваемый людьми: социальные сети, новостные статьи, блоги или объявления о работе. В отличие от исчерпывающих и данных зондирования цифровой контент специально кем-то редактируется, поэтому он субъективен или даже обманчив, в зависимости от намерений автора.¹⁶⁷

Большие данные важны, потому что они позволяют нам «использовать информацию новыми способами для получения полезной информации или производства товаров и услуг, представляющих значительную ценность».¹⁶⁸

166 Matthew J. Salganik, *Bit By Bit: Social Research in the Digital Age* (Princeton and Oxford, Princeton University Press, 2018), pp. 17-41.

167 Soenke Ziesche, *Innovative Big Data Approaches for Capturing and Analyzing Data to Monitor and Achieve the SDGs* (Bangkok, ESCAP, 2017), p. 18. Доступно по: <https://reliefweb.int/sites/reliefweb.int/files/resources/Innovative%20Big%20Data%20Approaches%20for%20Capturing%20and%20Analyzing%20Data%20to%20Monitor%20and%20Achieve%20the%20SDGs.pdf>.

168 Viktor Mayer-Schonberger and Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work and Think* (London, John Murray, 2013).

Вставка 12. Большие данные в бизнесе

Ниже приведены некоторые примеры того, как большие данные используются в бизнесе:

Профилирование моделей и таргетирование рекламы – Каждый раз, когда вы входите в Google или Facebook и видите рекламу, основанную на ваших предпочтениях, истории просмотров, лайках в Facebook, группах, к которым присоединились, на том, что понравилось вашим друзьям и т. д.

Управление доходами – Каждый раз, когда вы пытаетесь купить авиабилет онлайн, цены меняются в зависимости от маршрута, спроса, ожидаемого в последнюю минуту спроса, раннего бронирования и так далее.

Рекомендательные движки – Каждый раз, когда вы заходите на сайты электронной торговли и просматриваете продукт, вы начинаете получать рекомендации по другим продуктам, которые также были просмотрены другими посетителями. Если в итоге вы что-то купите, вы получите рекомендации по другим продуктам, которые связаны с приобретенным. Например, купите телефон, и он порекомендует чехол или защитную пленку для экрана. Если вы выполняете международные звонки, то вы можете получить рекомендацию по выгодному пакету для звонков. Идея состоит в том, чтобы преобразовать негарантированный будущий доход (без гарантии при совершении следующего международного звонка) в гарантированный ежемесячный доход (с помощью покупки пакета звонков, который позволяет определенное количество международных звонков по сниженной ставке).

Взято (с изменениями) из: Adityavijay Rathore's response to, "What are real-life examples of the application of big data analytics?" Quora, 9 January 2015. Доступно по <https://www.quora.com/What-are-real-life-examples-of-the-application-of-big-data-analytics>.

Большие данные могут быть использованы для развития и управления.

Инициатива Организации Объединенных Наций «Глобальный пульс» предлагает три категории потенциального применения больших данных для целей развития:

- **Раннее оповещение** – Раннее обнаружение аномалий в том, как население использует цифровые устройства и услуги, может обеспечить более быстрое реагирование во время кризиса;
- **Осведомленность в режиме реального времени** – Большие данные могут нарисовать детальное и текущее представление реальности, которое может помочь в разработке и целенаправлении программ и политик;
- **Обратная связь в режиме реального времени** – Возможность осуществления мониторинга населения в режиме реального времени позволяет понять, где не удастся проводить политику и реализовывать программы, и внести необходимые корректировки.¹⁶⁹

169 United Nations Global Pulse, "Big Data for Development: Challenges and Opportunities", May 2012. Доступно по <http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseMay2012.pdf>.

Эммануэль Летузе предлагает альтернативную классификацию использования больших данных для развития.¹⁷⁰ Большие данные могут быть:

- **Описательными** – Раннее обнаружение аномалий в том, как население использует цифровые устройства и услуги, может обеспечить более быстрое реагирование во время кризиса;
- **Предсказывающими** – Большие данные могут нарисовать детальное и текущее представление реальности, которое может помочь в разработке и целенаправлении программ и политик;
- **Предписывающими или диагностическими** – Возможность осуществления мониторинга населения в режиме реального времени позволяет понять, где не удастся проводить политику и реализовывать программы, и внести необходимые корректировки.

Вставка 13. Как Таиланд использует большие данные для власти правительства

После обеспечения Интернет-доступа для всех правительство Таиланда переключило свое внимание на еще одну сторону стратегии Таиланда 4.0 – данные.

Первые большие данные

В качестве первого шага правительство Таиланда объединит данные всех 20 министерств в централизованную систему управления большими данными. У всех министерств будут три первоначальные задачи: проверка списков наборов данных, идентификация наборов данных и определение основных целей применения для общественного блага.

Поскольку большинство наборов данных в Таиланде являются «традиционными и неструктурированными», простое преобразование их в электронные данные будет сложной задачей. Но когда система будет полностью интегрирована, все государственные ведомства получат доступ к данным, что позволит им лучше реализовывать политику и содействовать цифровой трансформации страны.

Кроме того, наборы данных будут доступны общественности, поэтому стартапы и инвесторы могут использовать данные правительства для разработки решений.

Большие данные в правительстве

Конечно, правительство также планирует использовать большие данные, чтобы сделать себя более эффективным.

Сак Сегхоунтод, президент Агентства по развитию цифрового правительства Таиланда, считает, что данные будут влиять на важные функции правительства, такие как бюджетирование, планирование и решение проблем граждан.

Он объясняет: «Прежде всего, правительство сможет обеспечить поступление бюджета по

170 Emmanuel Letouzé, "Big data for development: Facts and figures", *SciDevNet*, 15 April 2014. Доступно по <https://scidev.net/global/data/feature/big-data-for-development-facts-and-figures.html>.

правильным маршрутам. Во-вторых, будет повышена прозрачность в правительстве, поскольку данные станут более доступными. В-третьих, граждане будут иметь возможность участвовать в определенных взаимодействиях с правительством, поскольку теперь они могут быстрее получать полезную информацию».

Нуттапон Нимманфатчарин, президент и исполнительный директор Агентства по продвижению цифровой экономики Таиланда, согласился с тем, что открытая система больших данных может помочь в борьбе с коррупцией – это то, что «готовится в ближайшее время».

Взято (с изменениями) из: Wen Chuan Tan, "How Thailand is using big data to power the government", Tech in Asia, 18 October 2018. Доступно по <https://www.techinasia.com/thailand-big-data-government>.

Существуют определенные преимущества в анализе огромных объемов данных.

Небольшие выборки не идеальны для изучения редких событий (инциденты, которые происходят с низкой частотой, но могут иметь широкое влияние), неоднородностей (несхожих или разнообразных элементов) и небольших (но существенных) различий.¹⁷¹

Большие данные могут привести к более целенаправленной политике и совершенствованию имеющихся мер, а также разработке новых показателей. Минимизируя потребность в опросах, большие данные также могут снизить затраты на подготовку статистических данных.¹⁷²

Но «большие» не обязательно всегда лучшие.

Так, большие данные не могут быть репрезентативными для изучаемой группы или категории. В развивающихся странах обработка больших данных позволит получить полезную информацию только о тех слоях общества, которые подключены. Полученная информация, вероятно, будет верна только для небольшой и более привилегированной части общества.

Существуют и другие проблемы в использовании больших данных: (1) они не являются результатом процесса, соответствующего стандартной практике; (2) они не соответствуют методологиям, классификациям и определениям официальной статистики, и поэтому их трудно согласовать и передать в статистические структуры; и (3) они затрагивают вопросы безопасности, конфиденциальности, владения данными и стабильности доступа.¹⁷³

Также важно помнить, что большие данные – это не волшебное средство решения проблем.

Возьмем показатели ЦУР. Анализ показал, что 70 показателей ЦУР (или почти одна

171 Matthew J. Salganik, *Bit By Bit: Social Research in the Digital Age* (Princeton and Oxford, Princeton University Press, 2018), pp. 17-21.

172 Walter J. Radermacher, "The Future Role of Official Statistics", *Power from Statistics Outlook Report*, p. 3. Доступно по https://www.researchgate.net/publication/320616460_The_Future_Role_of_Official_Statistics.

173 Там же.

треть всех показателей) не подходят для расчетов больших данных.¹⁷⁴ Если говорить конкретнее, то:

- Свыше 45 процентов показателей уровня 3 (т.е. международно-признанной методологии или стандартов для этого показателя пока нет, но методология / стандарты разрабатываются или проверяются) и почти 11 процентов показателей уровня 2 (показатели концептуально ясны, методология или стандарты, признанные на международном уровне, доступны, но страны не передают данные на регулярной основе) не подходят для анализа больших данных.
- ЦУР 12 (Ответственное потребление и производство), 13 (Климатические меры) и 17 (Партнерство для достижения целей) имеют наибольшее количество показателей, которые не подходят для анализа больших данных.
- ЦУР 12 и 13 также имеют наибольшее количество показателей уровня 2 и уровня 3.¹⁷⁵

3.3 Данные в режиме реального времени

Данные в режиме реального времени относятся к данным, которые передаются и используются сразу после сбора. Однако в области развития это также относится к: «информации, которая создается и становится доступной за относительно короткий и соответствующий период времени, и информации, которая становится доступной в течение срока, позволяющего предпринять ответные действия».¹⁷⁶

Данные в режиме реального времени включают в себя сообщения в социальных сетях, спутниковые снимки, данные об уровне осадков и потоков, отслеживаемые датчиками, а также данные о местонахождении смартфонов. Чтобы принимать своевременные решения, данные в режиме реального времени нуждаются в цифровых технологиях для их сбора, обмена, управления, анализа и составления отчетов.¹⁷⁷

Данные в режиме реального времени могут повысить возможности действующих лиц по реагированию на изменения в рабочих условиях, обучению при постоянных оценках эффективности действий, выявлению аномалий, реагированию на возникающие проблемы, улучшению внутренней координации, оптимизации распределения ресурсов, принятию необходимых мер при обратной связи с гражданами и предвидению тенденций и будущих событий.¹⁷⁸

В настоящее время очень немногие правительства используют данные в режиме

174 Soenke Ziesche, *Innovative Big Data Approaches for Capturing and Analyzing Data to Monitor and Achieve the SDGs* (Bangkok, ESCAP, 2017), p. 98. Доступно по <https://reliefweb.int/sites/reliefweb.int/files/resources/Innovative%20Big%20Data%20Approaches%20for%20Capturing%20and%20Analyzing%20Data%20to%20Monitor%20and%20Achieve%20the%20SDGs.pdf>.

175 Там же.

176 United Nations Global Pulse, "Big Data for Development: Challenges and Opportunities", May 2012, p. 15. Доступно по <http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseMay2012.pdf>.

177 Andreas Pawelke and others, *Data for development: What's next? – Concepts, trends and recommendations for German development cooperation* (Bonn and Eschborn, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, 2017), p. 36. Доступно по http://webfoundation.org/docs/2017/12/Final_Data-for-development_Whats-next_Studie_EN-1.pdf.

178 Там же, стр. 38-39.

реального времени для отслеживания хода реализации программ. Правительства отслеживают деятельность и ход развития на ежеквартальной основе (если не реже), используя показатели в качестве ориентира для оценки прогресса. С данными в режиме реального времени ход развития может контролироваться на ежедневной или еженедельной основе.

В Соединенных Штатах Америки данные в режиме реального времени были успешно использованы для предотвращения мошенничества в сфере здравоохранения (см. Вставку 14).

Вставка 14. Аналитика данных в режиме реального времени в правительстве

В 2016 году Министерство юстиции Соединенных Штатов и Управление Генерального инспектора Министерства здравоохранения и социальных служб США (HHS-OIG: Health and Human Services) опубликовали ежегодный программный отчет о борьбе с мошенничеством и злоупотреблениями в здравоохранении (HCFAC: Health Care Fraud and Abuse Control), показывающий, что на каждый потраченный доллар на расследования мошенничества и злоупотреблений в сфере здравоохранения в рамках этой и других программ за последние три года правительство вернуло 6,1 долл. США. «Это впечатляющее возмещение», - сказала бывший секретарь HHS Кэтлин Себелиус, - «во многом связано с новой системой компьютерной аналитики, которая обнаруживает и останавливает мошенническое выставление счетов до того, как выплачиваются деньги».

В отчете по программе HCFAC за 2015 финансовый год объясняется, что комплексные инструменты анализа данных HHS-OIG включают в себя анализ данных, прогнозную аналитику, подходы к оценке тенденций и моделированию, которые позволяют лучше анализировать и планировать контроль над программами HHS. В отчете более подробно рассказывается, как инициативные группы по предотвращению и обеспечению защиты от мошенничества в здравоохранении используют данные в режиме реального времени для определения схем мошенничества, выявления предполагаемых тенденций осуществления мошенничества и для расчета соотношения разрешенных услуг по сравнению со средними показателями по стране, а также с другими оценками. Эти передовые технологии и методы, взятые из частного сектора, привели к беспрецедентным усилиям по внедрению инноваций в борьбу с мошенничеством в сфере здравоохранения.

Взято (с изменениями) из: John M. LeBlanc, "Real-Time Data Analytics in Government Investigations", Manatt, 23 June 2016. Доступно по <https://www.manatt.com/Insights/Articles/2016/Real-Time-Data-Analytics-in-Government-Investigations>.

В Африке проект Pulse Lab Kampala тесно сотрудничает с правительством Уганды в различных проектах по мониторингу качества предоставления государственных услуг в режиме реального времени, поскольку Уганда в своем Национальном плане развития приоритизировала этапы мониторинга и оценки.¹⁷⁹

¹⁷⁹ Pulse Lab Kampala, "Analytics in Real Time Can Help Monitor the Quality of Public Service Delivery", *United Nations Global Pulse*, 15 June 2016. Доступно по <https://www.unglobalpulse.org/real-time-analytics-for-public-service-delivery>.

Тем не менее, более широкое использование данных в режиме реального времени в правительстве было затруднено из-за соперничества с другими источниками данных, недоверием к качеству данных, недостаточной осведомленностью о доступных данных, данными, которые не были надлежащим образом превращены в информацию, отсутствием визуализации, и информацией, которая не предназначена для потребностей конечного пользователя.¹⁸⁰

3.4 Малые данные

Распространение понятия малых данных происходит в качестве «человеко-ориентированной альтернативы большим данным».¹⁸¹ Большие данные, как и малые собираются и обрабатываются для получения информации, но последние опираются на сочетании внимательных наблюдений за небольшими выборками и применением интуиции.

По Мартину Линдстрому малые данные:

Могут содержаться в держателе зубной щетки в ванной комнате в Тель-Авиве; или каким образом рулон туалетной бумаги прижимается к стене ванной комнаты в северной Бразилии. Это может проявляться в том, как обувь семьи выстроена в коридоре, или в зашифрованных буквах и цифрах, составляющих компьютерный пароль человека.¹⁸²

Малые данные – это подсказки, которые приводят к пониманию: «Одинокий участок небольших данных почти никогда не бывает достаточно значимой, чтобы открыть дело или выдвинуть гипотезу, но смешанные с другими идеями и наблюдениями ... эти данные в конечном итоге объединяются для создания решения, которое формирует основу будущего бренда или бизнеса».¹⁸³

Сбор малых данных является частью того, что Линдстром называет «исследованием подтекста» - процессом, который включает:

- **Сбор** – Получение как можно большего количества различных точек зрения из максимально возможного количества надежных источников, а также создание точек привязки, которые помогут структурировать начальные наблюдения и выработать гипотезу;
- **Подсказки** – Артефакты (небольшие данные), которые помогают создать повествование – связную историю, которая поддерживает или опровергает гипотезу, может быть, физическую или эмоциональную;
- **Связывание** – Придание смысла небольшим данным (подсказки начинают

180 Andreas Pawelke and others, *Data for development: What's next? – Concepts, trends and recommendations for German development cooperation* (Bonn and Eschborn, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, 2017), p. 39. Доступно по http://webfoundation.org/docs/2017/12/Final_Data-for-development_Whats-next_Studie_EN-1.pdf.

181 Roger Dooley, "Small Data: The Next Big Thing", *Forbes*, 16 February 2016. Доступно по <https://www.forbes.com/sites/rogerdooley/2016/02/16/small-data-lindstrom/#1811e5fb7870>.

182 Martin Lindstrom, *Small Data: The Tiny Clues that Uncover Huge Trends* (New York, Picador, 2017), p. 9.

183 Там же.

- срабатывать в одном направлении?);
- **Причинная связь** – «Мелкая добыча» - в поисках причины и следствия;
 - **Корреляция** – Нахождение взаимоотношений или связей;
 - **Возмещение** – Выявление неудовлетворенного желания или беспокойства;
 - **Концепция** – Определение «решения» для проблемы или желания, которые были выявлены.¹⁸⁴

Линдстром сравнивает свой процесс с этнографией – «запись и анализ культуры или общества, обычно основанные на наблюдениях участников и приводящие к письменному описанию людей, места или учреждения».¹⁸⁵

Другое определение малых данных – «небольшие наборы данных, которые могут влиять на решения в настоящее время».¹⁸⁶ Родственное определение – «данные в объеме и формате, которые делают их доступными, информативными и действенными».¹⁸⁷

Малые данные «связывают людей со своевременной, полезной информацией (полученной из больших данных и/или «локальных» источников), организованной и оформленной – часто визуальной – чтобы быть доступной, понятной и применимой для повседневных задач».¹⁸⁸ Она включает в себя результаты бейсбола, отчеты об инвентаризации, записи вождения, данные о продажах, историю поиска, прогнозы погоды и инструкции об использовании.¹⁸⁹

До сих пор непонятно? Вот мнение Джона Спейси: «Малые данные ранее просто назывались данными. Современный термин используется для различия между традиционными конфигурациями данных и большими данными».¹⁹⁰

Вставка 15. Почему малые данные?

- **Большие данные – это сложно:** Выполнение в масштабе и ожидание стимулирующих результатов может занять время. Не говоря уже о том, что большинству маркетологов и онлайн-стратегов не нужны полномасштабные большие данные для таргетирования своих кампаний или предоставления персонализированного сервиса.
- **Малые данные вокруг нас:** Социальные сети богаты малыми данными, которые готовы собирать для обслуживания маркетинговых и покупательских решений. На индивидуальном уровне мы постоянно создаем малые данные при каждой регистрации, поиске, просмотре,

184 Там же, стр. 219-225.

185 Discover Anthropology, "Ethnography", *Royal Anthropological Institute*. Доступно по <https://www.discoveranthropology.org.uk/about-anthropology/fieldwork/ethnography.html> (accessed on 8 January 2019).

186 Sahil Miglani, "Big Data and Small Data: What's the Difference?" *Dataversity*. Доступно по <http://www.dataversity.net/big-data-small-data/> (accessed on 8 January 2019).

187 Ahmed Banafa, "Small Data vs. Big Data: Back to the Basics", *OpenMind*, 25 July 2016. Доступно по <https://www.bbvaopenmind.com/en/small-data-vs-big-data-back-to-the-basics/>.

188 Small Data Group, "Defining Small Data", 18 October 2013. Доступно по <https://smalldatagroup.com/2013/10/18/defining-small-data/>.

189 Margaret Rouse, "Small Data", *WhatIs.com*, June 2014. Доступно по <https://whatIs.techtarget.com/definition/small-data>.

190 John Spacey, "8 Examples of Small Data", *Simplifiable*, 30 January 2018. Доступно по <https://simplifiable.com/new/small-data>.

публикации и т. д., создавая уникальную запись, которая дает представление о нашем цифровом и физическом здоровье.

- **Малые данные находятся в центре нового управления взаимоотношениями с клиентами:** Малые данные – это ключ к созданию обширных профилей, которые станут центром новых решений по управлению взаимоотношениями с клиентами.
- **Отдача от инвестиций:** На этапе так называемой «последней мили» больших данных предлагается использовать инвестиции в малые данные (10 миллиардов долларов США согласно расчетам IDC), потраченные на системы, инструменты и услуги более высокого уровня.
- **Управляемый данными маркетинг – это новый этап:** Маркетинг, основанный на больших (и малых) данных, может кардинально изменить способы взаимодействия бизнес-предприятий с клиентами, изменить способы доступа и использования (и даже ношения) полезных данных клиентами и, в конечном итоге, переопределить отношения между покупателями и продавцами.
- **Потребительских примеров предостаточно:** Потребители увидели потенциал малых данных по упрощению процесса покупок, улучшению физической подготовки или предоставлению рекомендаций по лучшей цене для своего следующего полета. С появлением более интеллектуальных носимых устройств, управляемых данными, появится еще больший рыночный спрос на обобщенные данные и устройства доставки данных, которые «соответствуют» потребностям повседневных потребителей.
- **Поставщики платформ и инструментов начинают обращать внимание:** Перспектива практического использования больших данных и «превращения информации в действие» является основным стимулом для многих громких имен в технологиях, включая SAP, Oracle и EMC.
- **Все для конечного пользователя:** Они включают в себя то, что им нужно и как они могут действовать. Сначала необходимо сфокусироваться на пользователе, и многие предлагаемые технологические решения станут яснее.
- **Простой:** Малые данные – правильные данные. Некоторые малые данные начинают жизнь как большие данные, но вам не нужно быть специалистом по данным, чтобы понимать или применять их для повседневных задач.

Взято (с изменениями) из: Ahmed Banafa, "Small Data vs. Big Data: Back to the Basics", OpenMind, 25 July 2016. Доступно по <https://www.bbvaopenmind.com/en/small-data-vs-big-data-back-to-the-basics/>.

3.5 Данные, генерируемые гражданами

Данные, сгенерированные гражданами – это «данные, которые люди или их организации производят для непосредственного мониторинга, запроса или стимулирования изменений по затрагивающим их вопросам».¹⁹¹ Примерами могут служить данные, созданные гражданами о качестве воздуха в Пекине, карта сексуальных домогательств в Египте (см. Вставку 16) и обновленные статусы по

191 DataShift, "What is Citizen-Generated Data and What is the Datashift Doing to Promote it?" n.d. Доступно по http://civicus.org/images/ER%20cgd_brief.pdf.

пунктам водоснабжения в Танзании.

Данные, сгенерированные гражданами, «предоставляют людям возможность принимать решения о том, какие данные собирать, для чего их использовать и как они будут использоваться».¹⁹² Они могут «привлечь внимание к вопросам, важным для людей, и представить свое мнение во время политических дискуссий на более высоком уровне».¹⁹³ Они также могут использоваться «для проверки официальных отчетов и наборов данных».¹⁹⁴ Это может означать «подотчетность процессов развития через независимый мониторинг».¹⁹⁵ Прежде всего, «данные, активно производимые гражданами с конкретной целью, могут обеспечить прямое представительство граждан и позволить им напрямую отслеживать, требовать или инициировать изменения по вопросам, которые их касаются».¹⁹⁶

Два тематических исследования инициатив в области развития, проведенных в Кении и Уганде, показывают, что инициативы по сбору данных, полученные от граждан, способствовали «улучшению обеспечения доступности и/или качества государственных услуг в рамках проектов в области развития» и «расширению прав и возможностей и участию местных субъектов в достижении подотчетности».¹⁹⁷ Они также раскрыли потенциал данных, производимых гражданами, для содействия мониторингу ЦУР на местном уровне.¹⁹⁸

Данные, производимые гражданами, дополняют институциональные данные и не должны рассматриваться в качестве замены или альтернативы. Данные, генерируемые гражданами, особенно полезны для понимания сообществ, в которых отсутствуют данные, связанные с положением женщин, уязвимых и маргинальных слоев населения.

Вставка 16. Картографирование случаев сексуальных домогательств в Египте

Приложение HarassMap было запущено в декабре 2010 года группой основателей из четырех

192 Andreas Pawelke and others, *Data for development: What's next? – Concepts, trends and recommendations for German development cooperation* (Bonn and Eschborn, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, 2017), p. 32. Доступно по http://webfoundation.org/docs/2017/12/Final_Data-for-development_Whats-next_Studie_EN-1.pdf.

193 Там же.

194 Там же.

195 Development Initiatives, "Citizen-generated data and sustainable development: Evidence from case studies in Kenya and Uganda report", March 2017, p. 10. Доступно по <https://hivos.org/sites/default/files/publications/15-citizen-generated-data-and-sustainable-development-evidence-from-case-studies-in-kenya-and-uganda.pdf>.

196 Andreas Pawelke and others, *Data for development: What's next? – Concepts, trends and recommendations for German development cooperation* (Bonn and Eschborn, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, 2017), p. 32. Доступно по http://webfoundation.org/docs/2017/12/Final_Data-for-development_Whats-next_Studie_EN-1.pdf.

197 Development Initiatives, "Citizen-generated data and sustainable development: Evidence from case studies in Kenya and Uganda report", March 2017, p. 5. Доступно по <https://hivos.org/sites/default/files/publications/15-citizen-generated-data-and-sustainable-development-evidence-from-case-studies-in-kenya-and-uganda.pdf>. The Kenya case examined the School Report Card initiative, an effort to increase parent participation in schooling of their children. The Uganda case examined a process of providing unsolicited citizen feedback to duty bearers and service providers in local communities in Uganda.

198 Andreas Pawelke and others, *Data for development: What's next? – Concepts, trends and recommendations for German development cooperation* (Bonn and Eschborn, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, 2017), p. 30. Доступно по http://webfoundation.org/docs/2017/12/Final_Data-for-development_Whats-next_Studie_EN-1.pdf.

женщин вместе с техническими партнерами, консультантами и волонтерами. Ниже четыре женщины размышляют над созданием и влиянием HarassMap:

«Мы все сталкивались со случаями сексуальных домогательств, которым мы, и почти все, кого мы знали, подвергались ежедневно.

В 2009 году один волонтер познакомил нас с Frontline SMS и Ushahidi, которые представляют собой бесплатное программное обеспечение, которое можно связать вместе, чтобы создать анонимную систему отчетности и картографирования для использования в онлайн и через службу коротких сообщений или СМС.

Поскольку около 97 процентов египтян, половина из которых составляли женщины, владели мобильным телефоном, данная технология выглядела в качестве возможности вновь привлечь внимание общественности к этой проблеме. Потребовался год, чтобы разработать нашу модель, мы проанализировали ситуацию и выработали подход к целевым областям, которые, по нашему мнению, не были затронуты программами по защите общественных интересов неправительственных организаций, которые уже действовали в то время.

Мы хотели быть уверенными, что HarassMap никогда не будет «просто картой», и для нас было важно, чтобы в нем был сильный компонент на уровне сообщества, который мог бы оказать сильное влияние на местах.

Отправной точкой было использование технологии отчетности и картирования для поддержки усилий по мобилизации сообщества в автономном режиме для разрушения стереотипов, прекращения оправданий преступников и убеждения людей высказываться и бороться против сексуальных домогательств.

Нам очень приятно видеть, что за последние годы сексуальные домогательства превратились из запретной темы в широко обсуждаемую. Наша команда добровольцев продолжает расти, и мы продолжаем расширять нашу работу на местах и в Интернете».

Взято (с изменениями) из: HarassMap, "Our Story". Доступно по <https://harassmap.org/who-we-are/our-story> (accessed on 8 January 2019).

4. Управление и администрирование данными

Задачи данного раздела:

- обсудить управление и администрирование данными, уделяя особое внимание Международной ассоциации по управлению данными, объединяющей знания.
-

Данные могут улучшить управление, когда они легко доступны, легко обмениваются и легко используются повторно.

Но обмен данными чрезвычайно затруднен, когда государственные органы используют разные системы ИКТ и хранят данные в разных форматах, а также при наличии перекрывающихся наборов данных и информационных хранилищ. Проблема усугубляется, когда стандарты информационной безопасности и качества данных слабы. Как отмечает Джелани Харпер:

Распределенная природа активов данных в государственном секторе еще сложнее, чем в частном секторе. Государственные органы сталкиваются с несколькими базами данных с различными точками доступа, местоположениями и архитектурой, а также проблемами безопасности регионального, национального и международного значения.¹⁹⁹

Управление и администрирование данными имеют ключевое значение для обеспечения доступности, надежности, качества и своевременности данных для пользователей.

4.1 Управление данными

Управление данными – это «комплексный процесс контроля целостности, использования, доступности, удобства использования и безопасности всех данных, принадлежащих или контролируемых предприятием».²⁰⁰ Это «система прав на принятие решений и ответственности за процессы, связанные с информацией, которые выполняются в соответствии с согласованными моделями, описывающими, кто может, какие действия предпринимать, с какой информацией и когда, при каких обстоятельствах, какими методами».²⁰¹

Как в государственном, так и в частном секторах ключевые тенденции,

199 Jelani Harper, “Lessons learned from big interoperability in government”, *KMWorld*, 30 October 2017. Доступно по <http://www.kmworld.com/Articles/Editorial/Features/Lessons-learned-from-big-interoperability-in-government-121253.aspx>.

200 Nate Lord, “What is Data Governance? Data Protection 101”, *Digital Guardian*, 10 September 2018. Доступно по <https://digitalguardian.com/blog/what-data-governance-data-protection-101>.

201 Data Governance Institute, “Definitions of Data Governance”. Доступно по http://www.datagovernance.com/adg_data_governance_definition/ (accessed on 8 January 2019).

обуславливающие необходимость управления данными, включают:

- Увеличение объемов данных из все большего числа источников, что приводит к несоответствиям данных, которые необходимо выявить и устранить до принятия решений с использованием неверной информации;
- Больше самостоятельного составления отчетности и аналитики (демократизация данных), что создает потребность в общем понимании данных во всей организации;
- Постоянное влияние нормативных требований, что делает еще более важным иметь четкое представление о том, где данные и как они используются;
- Растущая потребность в общем языке делового общения для межведомственного анализа и принятия решений.²⁰²

Цели управления данными состоят в том, чтобы:

- Предоставить организации возможность управлять своими данными как активом;
- Определить, утвердить, сообщить и реализовать принципы, политики, процедуры, показатели, инструменты и обязанности по управлению данными;
- Проводить мониторинг и осуществлять руководство политикой на предмет соответствия, предоставлять услуги по использованию и управлению данными.²⁰³

Управление данными:

- Способствует принятию более эффективных решений;
- Сокращает операционные разногласия;
- Заботится об удовлетворении потребностей заинтересованных сторон данных;
- Обучает менеджмент и персонал по выработке общих подходов к решению проблем, связанных с данными;
- Определяет стандарты и создает повторяемые процессы;
- Снижает затраты;
- Повышает эффективность за счет координации усилий;
- Обеспечивает прозрачность процессов.²⁰⁴

Орган (например, совет по управлению данными, комитет или группа) обычно является официально зарегистрированным, и на него возлагается обязанность определять стратегическое направление для ЧЕГО должна выполняться программа управления данными и КОГДА она должна выполнять ее.²⁰⁵ Этот орган также должен делать все

202 Nancy Couture, "Why data governance?" *CIO*, 3 January 2018. Доступно по <https://www.cio.com/article/3245588/governance/why-data-governance.html>.

203 DAMA International, *Data Management Body of Knowledge*, second edition (DAMA-DMBOK 2) (New Jersey, Technics Publications, 2017), Chapter 3, pp. 67-95.

204 Data Governance Institute, "Goals and Principles for Data Governance". Доступно по http://www.datagovernance.com/adg_data_governance_goals/ (accessed on 8 January 2019).

205 George F., "Data governance council – what is it and why do you need one?" *Lights on Data*, 18 July 2018. Доступно по <http://www.lightsondata.com/data-governance-council/>.

возможное, чтобы все о нем знали и получали поддержку при необходимости.

Деятельность по управлению данными включает следующее:

1. Определение управления данными для организации
 - Разработать стратегию управления данными (план для поддержания и улучшения качества данных, целостности, безопасности и доступа);
 - Провести оценку готовности;
 - Выполнить исследование и бизнес-согласование;
 - Разработать организационные аспекты.
2. Создание стратегии управления данными –
 - Определить операционную структуру управления данными;
 - Разработать цели, принципы и политики;
 - Утвердить проекты управления данными;
 - Участвовать в управлении изменениями;
 - Участвовать в управлении работой по проблемам;
 - Выполнить оценку соответствия нормативным требованиям.
3. Реализация управления данными –
 - Профинансировать разработку стандартов данных и процедур;
 - Разработать глоссарий бизнес-операций;
 - Проводить координацию между группами по созданию архитектуры;
 - Профинансировать оценку активов данных.
4. Внедрение управления данными.²⁰⁶

Основными результатами являются:

- Стратегия управления данными;
- Стратегия накопления данных;
- План действий стратегии данных;
- Принципы работы с данными;
- Политики и процессы управления данными;
- Операционная структура;
- План действий и стратегия реализации;

²⁰⁶ DAMA International, *Data Management Body of Knowledge*, second edition (DAMA-DMBOK 2) (New Jersey, Technics Publications, 2017), Chapter 3, pp. 67-95.

- Оперативный план;
- Глоссарий бизнес-операций;
- Система учета результатов управления данными;
- Вебсайт управления данными;
- План коммуникаций;
- Значение признанных данных;
- Усовершенствованные методы управления данными.²⁰⁷

NASCIO – ассоциация Соединенных Штатов Америки, представляющая государственных руководителей информационных служб, руководителей и менеджеров по информационным технологиям – выступает за то, чтобы правительства:

- Включали управление данными в качестве приоритетного пункта повестки дня, который необходимо изучить и рассмотреть в рамках общей инициативы по управлению ИКТ;
- Признавали управление данными как часть более широкой области управления накопленными знаниями;
- Понимали, что управление данными должно быть правильно спланировано и согласовано (с политикой, организацией, процессом, обязательством, коммуникациями, структурами, методами и процедурами, оценкой и инструментами);
- Использовали модели зрелости для планирования контролируемых этапов управления данными;
- Установили процесс информирования о преимуществах и ожидаемых результатах управления данными высокоприоритетным для привлечения и поддержки участия заинтересованных сторон.²⁰⁸

В Юго-Восточной Азии Таиланд и Индонезия признали важность управления данными.

Вставка 17. Satu Data Индонезия

Распорядители данных в государственном секторе могут извлечь выгоду из простых в понимании технических руководящих принципов для эффективной реализации новых процедур управления. Проект президентского постановления о Satu Data Индонезия содержит такие рекомендации.

Три области, которые предоставляют возможности для усиления управления данными в государственном секторе:

1. Организационная структура – Плохая координация между государственными структурами – одна из трудностей, с которыми распорядители данных продолжают сталкиваться. Для ее

²⁰⁷ Там же

²⁰⁸ NASCIO, “Data Governance – Managing Information as an Enterprise Asset: Part I – An Introduction”, April 2008, p. 11. Доступно по <https://www.nascio.org/Portals/0/Publications/Documents/NASCIO-DataGovernance-Part1.pdf>.

улучшения необходимы конкретные руководящие указания по организационной структуре, которые могут прояснить процедуры координации, связанные с управлением данными. Эти руководящие указания могут включать в себя подробное объяснение шагов по координации данных, обеспечению соответствия между компетентными распорядителями данных и конкретными задачами, а также по организации форума по управлению данными среди государственных заинтересованных сторон.

2. **Поток данных** – Несмотря на то, что распорядители данных в государственном секторе понимают важность поддержки качества данных, не все из них полностью осведомлены о необходимых процессах – от этапа сбора данных до момента использования данных в разработке политики. Руководящие указания, которые могут перевести этот общий процесс в четкие шаги, могут помочь распорядителям данных лучше выполнять свои обязанности. В частности, были определены пять сегментов движения данных, которые могут помочь в реализации политики Satu Data Индонезия. Эти сегменты включают: потоки сбора и распространения данных, поток стандартизации, поток процесса проверки, поток взаимодействия и поток распространения данных.
3. **Формат данных** – Как правило, распорядители данных в ведомствах и министерствах государственного сектора не имеют единого общего формата для каталогизации цифровых данных. В результате люди используют любые предпочтительные форматы, и это приводит к снижению эффективности, когда дело доходит до обмена данными между государственными ведомствами. Преимущество разработки директивы по стандартным форматам данных заключается в том, что она помогает всем игрокам соблюдать принципы взаимодействия, изложенные в Satu Data Индонезия, одновременно улучшая навыки управления данными, связанные со сбором, обработкой и распространением данных. Эта директива также может служить глоссарием терминологии данных, формата метаданных и формата данных.

Взято (с изменениями) из: Pulse Lab Jakarta, "Identifying Opportunities to Strengthen Data Governance in the Public Sector", Medium, 8 October 2017. Доступно по <https://medium.com/pulse-lab-jakarta/identifying-opportunities-to-strengthen-data-governance-in-the-public-sector-92d9bc7db5ad>.

Управление данными отличается от администрирования данными.

4.2 Администрирование данными

Администрирование данными представляет собой: «Разработку и выполнение процессов, архитектур, политик, практик и процедур для управления информацией, генерируемой организацией».²⁰⁹ Оно включает в себя получение, проверку, хранение, защиту и обработку данных, чтобы обеспечить их доступность, надежность и своевременность для своих пользователей.²¹⁰

Администрирование данными помогает обеспечить доступность данных, когда и где они необходимы. Это особенно необходимо, когда государственные данные хранятся в

209 Blue-Pencil, "What is Data Management and Why it is Important", 23 November 2015. Доступно по <http://www.blue-pencil.ca/what-is-data-management-and-why-it-is-important/>.

210 Molly Galetto, "What is Data Management?" NGDATA, 31 March 2016. Доступно по <https://www.ngdata.com/what-is-data-management/>.

различных форматах и системах, которые поддерживают деятельность конкретных программ, отделов и организаций.²¹¹

Преимущества хорошего администрирования данными включают в себя:

- Повышение уровня информированности и понимания того, какие данные доступны для текущего и будущего использования, благодаря лучшей каталогизации и архивированию данных;
- Улучшенный доступ к данным, свободный от ненужных препятствий, защищенный от разглашения личной информации или нарушения юридических и договорных обязательств;
- Повышение качества и более своевременная информация, то есть доступ к нужной информации в нужное время, что объясняется более быстрой идентификацией потребностей клиентов и недопущением неверной или противоречивой информации благодаря использованию эффективных метаданных;
- Лучшее соотношение цены и качества за счет четких, справедливых и согласованных размеров сборов оплаты и условий использования данных, которые признают необходимость бесплатного доступа для соответствующих клиентов;
- Повышение эффективности использования данных в целом, благодаря более легкому обмену данными и интеграции с другими гармонизированными данными;
- Повышение эффективности деятельности правительства и его ведомств в результате использования более качественных данных.²¹²

Администрирование данными осуществляется «посредством единой инфраструктуры технологических ресурсов и управляющей структуры, которая определяет административные процессы, используемые на протяжении всего жизненного цикла данных».²¹³

Администрирование данными обеспечивает организации получение выгоды из своих данных. Руководство осуществляется процессами управления данными.

Правительствам, заинтересованным в управлении на основе данных, следует рассмотреть возможность использования *Международного руководства Международной ассоциации по управлению данными (DAMA) по Своду знаний управления данными (DAMA-DMBOK 2)*. Руководство представляет собой набор процессов и областей знаний, которые обычно считаются передовой практикой в рамках дисциплины управления данными.

211 Francesca El-Attrash, "What Government Needs to Know About Data Management", *GovLoop*, 9 August 2017. Доступно по <https://www.govloop.com/government-needs-know-data-management/>.

212 Intra-Governmental Group on Geographic Information, *The Principles of Good Data Management*, second edition (London, Office of the Deputy Prime Minister, 2005). Доступно по https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/14867/Good_dataMan.pdf.

213 Technopedia, "Data Management". Доступно по <https://www.techopedia.com/definition/5422/data-management> (accessed on 8 January 2019).

Основным недостатком DAMA-DMBOK 2 заключается в рассмотрении управления данными в качестве элемента администрирования данными. Однако это не умаляет его ценности как полезного всеобъемлющего руководства для понимания и реализации управления данными.

В оставшейся части этого раздела основное внимание будет уделено следующим областям знаний DAMA-DMBOK 2: архитектура данных; моделирование и дизайн данных; хранение данных и операции; безопасность данных; интеграция данных и совместимость; управление документами и контентом; справочные и основные данные; хранение данных и бизнес-аналитика; управление метаданными; и управление качеством данных.

4.2.1 Архитектура данных

Под архитектурой данных понимается: «Организованное расположение компонентов, предназначенное для оптимизации функции, производительности, осуществимости, стоимости и эстетики всей структуры или системы. ... Она определяет потребности предприятия в данных (независимо от структуры), а также разрабатывает и поддерживает основной план для удовлетворения этих потребностей».²¹⁴

Цели управления архитектурой данных:

- Определить требования к хранению и обработке данных;
- Разработать структуры и планы для удовлетворения текущих и долгосрочных потребностей к данным предприятия;
- Стратегически подготовить организации к быстрому развитию своих продуктов, услуг и данных, чтобы использовать возможности, которые открывают появляющиеся технологии.

Деятельность включает:

1. Создание архитектуры данных организации –
 - Оценить существующие спецификации архитектуры данных;
 - Разработать план действий;
 - Удовлетворить потребности организации в рамках проектов;
2. Интеграция с корпоративной архитектурой.

Основными результатами являются:

²¹⁴ This subsection is drawn from: DAMA International, *Data Management Body of Knowledge*, second edition (DAMA-DMBOK 2) (New Jersey, Technics Publications, 2017), Chapter 4, pp. 97-120.

- Дизайн архитектуры данных;
- Потоки данных;
- Цепочки производства данных;
- Модель данных организации;
- План практических мероприятий.

Вставка 18. Шесть принципов современной архитектуры данных

1. **Рассматривать данные в качестве общий ресурс:** Организации, которые концептуально рассматривают данные в качестве общего актива, в конечном итоге превосходят своих конкурентов.
2. **Предоставить подходящие интерфейсы для пользователей, чтобы использовать данные:** Размещение данных в одном месте недостаточно для достижения перспективы организации, управляемой на основе данных. Чтобы люди (и системы) могли использовать общий ресурс данных, необходимо предоставить интерфейсы, облегчающие пользователям использование этих данных.
3. **Обеспечить контроль безопасности и доступа:** Ориентируйтесь на технологии, которые позволяют разрабатывать систему безопасности и предоставлять широкий доступ к самообслуживанию без ущерба для контроля.
4. **Согласовать общую терминологию:** Без этого общей терминологии вы будете тратить больше времени на обсуждение или согласование результатов, чем на повышение производительности.
5. **Обрабатывать данные:** Без надлежащей обработки данных (которая включает в себя моделирование важных взаимосвязей, очистку необработанных данных и настройку ключевых параметров и показателей) конечные пользователи могут испытать разочарование – что значительно снизит потенциальную и реализуемую ценность исходных данных.
6. **Исключить копирование и перемещение данных:** Каждый раз, когда данные перемещаются, это влияет на стоимость, точность и время. Устраняя необходимость в дополнительном перемещении данных, современные корпоративные архитектуры данных могут снизить затраты (время, усилия, точность), повысить «свежесть» данных и оптимизировать общую гибкость корпоративных данных.

Взято (с изменениями) из: Joshua Klahr, "The 6 Principles Of Modern Data Architecture", AtScale, 19 January 2018. Доступно по <https://www.atscale.com/blog/the-six-principles-of-modern-data-architecture>.

4.2.2 Моделирование и дизайн данных

Под моделированием и дизайном данных понимается: «процесс выявления, анализа и определения требований к данным, а затем представления и передачи этих требований к данным в точной форме, называемой моделью данных».²¹⁵

Цель состоит в том, чтобы подтвердить и задокументировать понимание различных точек зрения, что приводит к приложениям, которые более точно соответствуют текущим и будущим бизнес-требованиям. Это создает основу для успешного завершения широкомасштабных инициатив, таких как программы администрирования основными данными и управления данными.

Мероприятия состоят из:

1. Планирование моделирования данных.
2. Построение моделей данных –
 - Создать концептуальную модель данных;
 - Создать логическую модель данных;
 - Создать физическую модель данных.

Основными результатами являются:

- Концептуальная модель данных;
- Логическая модель данных;
- Физическая модель данных.

Вставка 19. Что такое моделирование данных?

Моделирование данных – это процесс создания модели данных для тех данных, которые будут храниться в базе данных.

Моделирование данных помогает визуальному представлению данных и обеспечивает соблюдение рабочих правил, нормативных требований и государственной политики в отношении данных. Модели данных обеспечивают согласованность по правилам наименования, значений по умолчанию, семантики и безопасности при обеспечении качества данных.

Модели данных уделяют внимание необходимости каких данных, и как они должны быть организованы, а не какие операции должны выполняться над данными. Модель данных – это как план здания архитектора, который помогает построить концептуальную модель и установить отношения между элементами данных.

²¹⁵ Данный подраздел взят из: DAMA International, *Data Management Body of Knowledge*, second edition (DAMA-DMBOK 2) (New Jersey, Technics Publications, 2017), Chapter 5, pp. 123-166.

Взято (с изменениями) из: Guru99, "What is Data Modelling? Conceptual, Logical, & Physical Data Models".
Доступно по <https://www.guru99.com/data-modelling-conceptual-logical.html> (accessed on 8 January 2019).

4.2.3 Хранение и операции с данными

Хранение и операции с данными включают в себя проектирование, реализацию и поддержку хранимых данных, чтобы обеспечить их максимальную пользу в течение всего жизненного цикла: от создания/сбора до утилизации.²¹⁶ Хранение и операции с данными включают два элемента деятельности – обслуживание баз данных и техническая поддержка данных.

Цели состоят в том, чтобы обеспечить:

- Управление доступностью данных на протяжении всего жизненного цикла данных;
- Целостность массивов данных;
- Управление производительностью транзакций данных.

Деятельность включает:

1. Управление технологиями баз данных –
 - Понять и оценить технологии баз данных;
 - Обеспечить управление и мониторинг технологий баз данных.
2. Управление операциями с базами данных –
 - Изучить требования;
 - Составить план обеспечения бесперебойного функционирования систем;
 - Создать копии баз данных;
 - Обеспечить управление производительностью баз данных;
 - Осуществить управление тестовыми базами данных;
 - Выполнить миграцию данных.

Основными результатами являются:

- Критерии оценки технологий баз данных;

²¹⁶ Данный подраздел взят из: DAMA International, *Data Management Body of Knowledge*, second edition (DAMA-DMBOK 2) (New Jersey, Technics Publications, 2017), Chapter 6, pp. 169-214.

- Конфигурации баз данных;
- Перенесенные/реплицированные/версионные данные;
- Планы обеспечения бесперебойного функционирования систем;
- Соглашения об оказании услуг по производительности баз данных.

Вставка 20. Типы данных для хранения

Первым шагом в создании эффективной политики хранения данных является ответ на вопрос: какие типы данных я должен хранить в течение длительного периода времени, а какие можно мгновенно удалить?

Ниже приведена общая иерархия, в которой указано, какие типы данных следует держать под рукой. Данные в верхней части списка являются наиболее важными для хранения как можно дольше, в то время как данные в нижней части наименее важны:

- Данные, которые должны храниться в соответствии с политиками по вопросам соблюдения или регулирования. Если по закону от вас требуется хранить определенный тип данных, вы обязательно должны хранить эти данные.
- Данные, которые относятся к вашим клиентам и, изучив «клиентов со всех углов», помогают вам взаимодействовать с ними. Понять ваших клиентов сложно, и вы не хотите отказываться от данных, которые помогут вам в решении этой проблемы.
- Деловые документы, контракты и т.д. Это важно хранить как можно дольше.
- Данные, которые генерируются повседневными бизнес-операциями, но не упорядочиваются. Эти данные могут быть полезны для анализа или планирования за прошлые периоды, но в них нет строгой необходимости.
- Машинные данные, генерируемые сетевым оборудованием, серверами, датчиками или другими типами автоматических источников. Машинные данные, как правило, являются наименее полезным типом данных для долгосрочного хранения. Иногда полезно иметь возможность просматривать машинные данные при исследовании технического инцидента или планировании расширения инфраструктуры, но по большей части машинные данные полезны только в режиме реального времени, потому что состояние вашей инфраструктуры меняется очень быстро.

Конечно, точные типы данных, которые нужно расставить по приоритетам для долгосрочного хранения, будут различаться в зависимости от учреждения. Данная иерархия просто представлена в качестве общего ориентира.

Взято (с изменениями) из: Christopher Tozzi, "Best Practices in Data Storage (Part 1): What Types of Data Should be Retained?" Syncsort, 17 July 2017. Доступно по <http://blog.syncsort.com/2017/07/big-data/data-storage-best-practices-data-types/>.

4.2.4 Безопасность данных

Безопасность данных – это «планирование, разработка и выполнение политик и процедур безопасности для обеспечения надлежащей аутентификации, авторизации,

доступа и аудита массивов данных и информации».²¹⁷

Цели состоят в том, чтобы:

- Разрешить соответствующий и предотвратить несанкционированный доступ к данным и их изменению;
- Изучить и соблюдать все соответствующие положения и политики в отношении вопросов неразглашения, защиты и конфиденциальности;
- Обеспечить соблюдение и аудит относительно неразглашения и конфиденциальности для всех заинтересованных сторон.

Деятельность включает в себя:

1. Определение соответствующих требований безопасности данных;
2. Разработка политики безопасности данных;
3. Определение стандартов безопасности данных;
4. Оценка текущих рисков безопасности;
5. Реализация системы контроля и осуществление процедур.

Основными результатами являются:

- Архитектура обеспечения безопасности данных;
- Политики по обеспечению безопасности данных;
- Стандарты по обеспечению конфиденциальности и неразглашения данных;
- Меры контроля за безопасным доступом к данным;
- Нормативные представления доступа к данным;
- Документированные классификации обеспечения безопасности;
- Система аутентификация и история доступа пользователей;
- Отчеты аудита безопасности данных.

Вставка 21. Предотвращение угроз

Организации часто тратят большую часть своих ресурсов по обеспечению информационной

²¹⁷ Данный подраздел взят из: DAMA International, *Data Management Body of Knowledge*, second edition (DAMA-DMBOK 2) (New Jersey, Technics Publications, 2017), Chapter 7, pp. 217-266.

безопасности, внедряя решения по управлению рисками, предотвращению инцидентов и антивирусные системы. Эти традиционные мероприятия все еще необходимы, но они формируют только один уровень в рамках более сложной стратегии. Чтобы уменьшить возможность атак через конечные точки и сети, и сделать успешную атаку как можно более затруднительной, организации должны:

Создайте механизм жесткого контроля и процедур, присущих только для вашей организации – Только вы знаете особенности вашей организации и ваших пользователей. Создайте уникальные для вашей организации процессы и убедитесь, что роли пользователей четко определены.

Обучите сотрудников быть вашей «первой линией защиты» – Обучая сотрудников, вы можете значительно снизить риск уязвимостей, вызванных человеческими ошибками. Убедитесь, что все сотрудники имеют доступ к соответствующему обучению, даже по темам, которые кажутся вам базовыми, например, не забывать выходить из системы, когда они покидают свои рабочие места, или хранить свои пароли в секрете.

Изучайте отраслевые сценарии взлома данных – Изучите информацию об угрозах, чтобы понять характерные типы атак и приемов.

Применять многоуровневый подход к технологиям обеспечения безопасности – Разнообразные решения по безопасности уменьшат характер угроз и предотвратят сложные атаки в вашей сети. Шифрование данных является обязательным. Дополните безопасность конечных точек сетевыми элементами управления.

Взято (с изменениями) из: Ryan St. Hilaire, "Data Security Best Practices Not Good Enough", eSecurity Planet, 8 April 2015. Доступно по <https://www.esecurityplanet.com/network-security/data-security-best-practices-not-good-enough.html>.

4.2.5 Интеграция и совместимость данных

Интеграция и совместимость данных «описывают процессы, связанные с перемещением и консолидацией данных внутри и между хранилищами данных, приложениями и организациями».²¹⁸

Цели состоят в том, чтобы:

- Обеспечить безопасное предоставление данных с соблюдением нормативных требований в необходимом формате и сроках;
- Уменьшить стоимость и сложность управления решениями за счет разработки общих моделей и интерфейсов;
- Определять значимые события и автоматически подавать сигналы оповещения и

218 This subsection is drawn from: DAMA International, *Data Management Body of Knowledge*, second edition (DAMA-DMBOK 2) (New Jersey, Technics Publications, 2017), Chapter 8, pp. 269-299.

начинать действия;

- Оказывать поддержку процессам анализа рабочих процессов, аналитики, управления основными данными и повышения эффективности работы.

Деятельность включает:

1. Планирование и анализ –
 - Определить требования к интеграции данных и их жизненного цикла;
 - Выполнить поиск данных;
 - Задokumentировать происхождение данных;
 - Кратко описать данные;
 - Изучить на соответствие бизнес-правилам.
2. Разработка решений для интеграции и взаимодействия данных –
 - Разработать компоненты решения;
 - Сопоставить источники с целями;
 - Разработать организационную структуру данных.
3. Разработка решений для интеграции и совместимости данных –
 - Разработать услуги передачи данных;
 - Разработать организацию потоков данных;
 - Разработать способы миграции данных;
 - Разработать комплексную обработку событий;
 - Поддерживать интеграцию данных и совместимость метаданных.
4. Внедрение и мониторинг.

Основными результатами являются:

- Архитектура интеграции и взаимодействия данных;
- Спецификации обмена данными;
- Соглашения о доступе к данным;
- Услуги передачи данных;
- Комплексная обработка событий;
- Пороговые значения и предупреждения.

Вставка 22. Совместимость данных и стихийные бедствия

Если бы передача данных была совместимой, некоторых трудностей, связанных с реагированием на ураган «Катрина» и восстановлением после него, можно было избежать. Например, если бы совместимость данных позволила улучшить взаимодействие между правительством, проводящим эвакуацию, и частными добровольными организациями, оказывающими помощь в эвакуации, и приютами, это позволило бы им легче сводить родителей с пропавшими детьми.

Функциональная совместимость данных влияет не только на правительства, но и на предприятия частного сектора. Несмотря на то, что большинство первых ответных мер находится в руках правительства, большая часть критически важной инфраструктуры в Соединенных Штатах Америки принадлежит частному сектору.

Реальная совместимость обеспечила бы немедленную ситуационную осведомленность с разных сторон ликвидации бедственного положения: коммунальное предприятие, которое должно устранить проблему, и государственная организация, которая должна принять меры по оказанию медицинской помощи, поставке генераторов для жизнеобеспечения людей, насосных станций водоснабжения и канализации и множество других функций, которые находятся в ведении местного самоуправления.

Устранение последствий стихийных бедствий в разных частях мира может быть значительно улучшено за счет использования международных стандартов. Воздействие от такого процесса стандартизации касается не только Соединенных Штатов, но и международного сообщества. Независимая организация должна иметь международные связи и членство, а международное сообщество должно участвовать в процессе разработки стандартов. Хорошим примером преимуществ, которые мы могли бы извлечь из этого процесса, являются более качественные результаты, которые мы получили бы благодаря свободному потоку информации на международном уровне во время азиатского цунами в 2004 году. Хотя меры реакции на это бедствие были очень хорошими, более эффективная передача данных могла бы оказаться полезной в процессе реагирования и восстановления.

Взято (с изменениями) из: Bill Lent, "Facing the Challenge Of Data Interoperability", Disaster Resource Guide. Доступно по http://disaster-resource.com/index.php?option=com_content&view=article&id=335%3Afaciing-the-challenge-of-data-interoperability&catid=9%3Acrisis-response&Itemid=15 (accessed on 8 January 2019).

4.2.6 Управление документацией и контентом

Управление документацией и контентом означает «контроль сбора, хранения, доступа и использования данных и информации, хранящихся вне реляционных баз данных».²¹⁹

Цели состоят в том, чтобы:

- Соблюдать юридические обязательства и соответствовать ожиданиям клиентов в отношении управления записями;
- Обеспечить эффективное и качественное хранение, поиск и использование

²¹⁹ Данный подраздел взят из: DAMA International, *Data Management Body of Knowledge*, second edition (DAMA-DMBOK 2) (New Jersey, Technics Publications, 2017), Chapter 9, pp. 303-344.

документов и контента;

- Обеспечить возможности интеграции между структурированным и неструктурированным контентом.

Деятельность включает:

1. Планирование управления жизненным циклом –
 - Создать план управления записями;
 - Разработать стратегию по контенту.
2. Создание политики обработки контента, включая поиск с электронным обнаружением.
3. Определение информационной архитектуры и управление жизненным циклом контента –
 - Производить сбор и управление записями и контентом;
 - Сохранять и размещать записи и контент;
 - Архивировать записи и контент.
4. Опубликование и подача контента.

Основными результатами являются:

- Стратегия управления контентом и записями;
- Политика и процедуры;
- Репозиторий контента;
- Управляемые записи во многих медиа-форматах;
- Проведение сквозной ревизионной проверки и ведение журнала.

Вставка 23. Управление документацией среди правительств американских штатов

Ниже приводятся результаты опроса, проведенного Центром цифрового правительства в 2016 году среди 203 должностных лиц Соединенных Штатов Америки в отношении использования ими решений для управления документацией:

1. **Управление документацией доминирует** – Более 80 процентов респондентов согласны с тем, что улучшение документооборота является приоритетом для них лично и для их организации.
2. **Сейчас самое время для модернизации** – Шестьдесят восемь процентов организаций активно модернизируют свой подход к управлению документами, а 15 процентов планируют сделать это в будущем.

3. **Автоматизация процессов является естественным дополнением** – Четверть респондентов планируют автоматизировать процесс в течение 12 месяцев.
4. **Вы можете начать с малого или взяться за все сразу** – Тридцать семь процентов организаций осуществляют управление документооборотом в масштабах всего агентства, 23 процента в одном отделе и 22 процента в масштабе штата.
5. **Быстро и настраиваемые – ключевые факторы** – При решении покупать более трети организаций предпочитают широко настраиваемое коммерческое решение. Только 10 процентов предпочитают разрабатывать решения собственными силами.
6. **Многофункциональные – это новая ценная тенденция** – Наиболее важными функциями автоматизированного управления документами являются: извлекаемые записи (77 процентов); оцифрованные электронные формы (59 процентов); электронная маршрутизация (50 процентов); и распознавание ошибок (30 процентов).
7. **Не все решения являются равноценными** – Респонденты говорят, что некоторые потребности в современных решениях для управления документами остаются невыполненными: простота использования (48 процентов); хранение (35 процентов); и безопасность (2 процента).
8. **Существуют препятствия для внедрения** – Тридцать восемь процентов респондентов считают, что интеграция с другими системами является самой большой проблемой при обосновании необходимости внедрения решения.
9. **Главы департаментов утверждают сделку** – Шестьдесят четыре процента респондентов считают, что руководители департаментов, такие как топ-менеджеры и вице-президенты, чаще всего принимают участие в процессе отбора.
10. **Это все о 3S (security, simplicity and savings) – безопасность, простота и экономия** – Ключевые факторы при выборе решения: 60-процентов – это безопасность; 49-процентов – это простота и удобство использования; и 45-процентов – это экономия.

Взято (с изменениями) из: Laserfiche, "10 Things to Know About Document Management in the Public Sector". Доступно по <https://www.laserfiche.com/ecmblog/infographic-10-things-to-know-about-document-management-in-the-public-sector/> (accessed on 8 January 2019).

4.2.7 Справочные и базовые данные

Справочные и базовые данные означают «управление общими данными для достижения целей организации, снижения рисков, связанных с избыточностью данных, обеспечения более высокого качества и снижения стоимости интеграции данных».²²⁰

Цели состоят в том, чтобы:

²²⁰ Данный подраздел взят из: DAMA International, *Data Management Body of Knowledge*, second edition (DAMA-DMBOK 2) (New Jersey, Technics Publications, 2017), Chapter 10, pp. 347-379.

- Обеспечить возможность совместного использования информационных активов между бизнес-направлениями и приложениями в организации;
- Предоставить достоверный источник зарегистрированных и оцененных базовых и справочных данных;
- Снизить стоимость и сложность благодаря использованию стандартов, общих моделей данных и схем интеграции.

Деятельность включает:

1. Определение движущих факторов и требований;
2. Анализ и определение источников данных;
3. Определение архитектурного подхода;
4. Моделирование данных;
5. Определение процессов управления и обслуживания;
6. Создание политики управления;
7. Внедрение услуг обмена данными/интеграции.

Основными результатами являются:

- Требования к базовым и справочным данным;
- Модели данных и документация;
- Достоверные справочные и базовые данные;
- Многократно используемые услуги передачи данных.

Вставка 24. Что такое базовые данные и справочные данные?

Базовые данные

Базовые данные - это ключевая бизнес-информация, которая поддерживает транзакции.

Основные данные описывают клиентов, продукты, детали, сотрудников, материалы, поставщиков, местоположения и т. д., участвующих в транзакциях. Они обычно относятся к местам (местоположения, география, места), заинтересованным сторонам (люди, клиенты, поставщики, служащие) и вещам (продукты, предметы, материалы, транспортные средства).

Основные данные обычно создаются и используются в ходе обычной деятельности существующими бизнес-процессами. К сожалению, эти операционные бизнес-процессы адаптированы для варианта использования этих базовых данных в зависимости от конкретных видов применения, и, следовательно, не позволяют выполнить общее требование предприятия, которое предписывает часто используемые базовые данные для приложений с высокими стандартами качества и общим управлением.

Справочные данные

Справочные данные – это данные, на которые ссылаются и которые используются несколькими системами. Большая часть справочных данных относится к понятиям, которые влияют либо на бизнес-процессы, например, на статус заказа (ВЫПОЛНЕНО | УТВЕРЖЛЕНО | ОТКЛОНЕНО), либо используется в качестве дополнительной стандартизированной семантики, которая дополнительно разъясняет интерпретацию записи данных, например, должность сотрудника (МЛАДШИЙ | СТАРШИЙ | ВП).

Некоторые из справочных данных могут быть универсальными и / или стандартизированными (например, страны - ISO 3166-1). Другие справочные данные могут быть «согласованы» внутри предприятия (статус клиента) или в рамках определенного бизнес-направления (классификации продуктов).

Справочные данные часто рассматриваются в качестве подмножества базовых данных. Полное название для этой категории данных - «базовые справочные данные».

Взято (с изменениями) из: FX Nicolas, "Back to Basics: Transactional, Master, Golden and Reference Data explained", Semarchy, 10 October 2018. Доступно по https://blog.semarchy.com/backtobasics_data_classification.

4.2.8 Хранение данных и бизнес-аналитика

Хранение данных и бизнес-аналитика означают «процессы планирования, внедрения и контроля для предоставления данных в целях принятия решений и поддержки работников умственного труда, занимающихся отчетностью, запросами и анализом».²²¹

Цели состоят в том, чтобы:

- Создавать и поддерживать технологическую среду, технические и бизнес-процессы, необходимые для предоставления интегрированных данных в целях поддержки действующих функций, требований соответствия и деятельности бизнес-аналитики;
- Поддерживать и обеспечивать эффективный бизнес-анализ и принятие решений работниками умственного труда.

Деятельность включает:

1. Изучение требования –
 - Определить и поддерживать архитектуру хранения данных и бизнес-аналитики;
2. Разработка хранилища данных и системы класса «data marts» (информационных витрин);
3. Заполнение хранилища данных;

²²¹ Данный подраздел взят из: DAMA International, *Data Management Body of Knowledge*, second edition (DAMA-DMBOK 2) (New Jersey, Technics Publications, 2017), Chapter 11, pp. 381-414.

4. Внедрение пакета бизнес-аналитики;
5. Поддержка работоспособности продуктов, полученных при обработке или передачи данных.

Основными результатами являются:

- Архитектура хранения данных и бизнес-аналитики;
- Продукты от обработки или передачи данных;
- Процесс заполнения;
- Деятельность по управлению;
- Словарь происхождения;
- План обучения и усвоения;
- План выпуска;
- Процесс поддержки производства;
- Деятельность по настройке нагрузки;
- Мониторинг деятельности бизнес-аналитики.

Вставка 25. Что такое хранилище данных?

Хранилище данных – это метод сбора и управления данными из различных источников для обеспечения ценной информации для организации. Это смесь технологий и компонентов, которая позволяет стратегически использовать данные.

Это электронное хранилище большого объема информации организации и предназначено для обработки запросов и анализа вместо обработки транзакций. Это процесс преобразования данных в информацию и своевременного предоставления их пользователям с целью изменения положения к лучшему.

Хранилище данных необходимо для всех типов пользователей, таких как:

- Лица, принимающие решения, которые полагаются на большое количество данных;
- Пользователи, которые используют специальные сложные процессы для получения информации из нескольких источников данных;
- Люди, которым нужна простая технология для доступа к данным;
- Люди, предпочитающие системный подход к принятию решений;
- Пользователи, которым нужна высокая производительность для обработки огромного количества данных, что необходимо для отчетов, таблиц или диаграмм;
- Пользователи, которые хотят обнаружить «скрытые закономерности» в потоках и группировках данных.

Взято (с изменениями) из: Guru99, "What Is Data Warehousing? Types, Definition & Example". Доступно по <https://www.guru99.com/data-warehousing.html> (accessed on 8 January 2019).

4.2.9 Управление метаданными

Метаданные или данные о данных «содержат информацию о технических и бизнес-процессах, правилах и ограничениях данных, а также логических и физических структурах данных».²²²

Цели состоят в том, чтобы:

- Обеспечить организационные договоренности по условиям и использованию;
- Собрать и интегрировать метаданные из разных источников;
- Предоставить стандартный способ доступа к метаданным;
- Обеспечить качество и безопасность метаданных.

Деятельность включает:

1. Определение стратегии по метаданным.
2. Изучение требований метаданных –
 - Требования корпоративных пользователей;
 - Технические требования.
3. Определение архитектуры метаданных –
 - Создать метамоделю;
 - Применить стандарты метаданных;
 - Обеспечить управление хранилищами метаданных.
4. Создание и поддержка метаданных –
 - Интегрировать метаданные;
 - Распространить и доставить метаданные.
5. Запросы, отчетность и анализ метаданных.

Основными результатами являются:

- Стратегия по метаданным;
- Стандарты метаданных;
- Архитектура метаданных;

²²² Данный подраздел взят из: DAMA International, *Data Management Body of Knowledge*, second edition (DAMA-DMBOK 2) (New Jersey, Technics Publications, 2017), Chapter 12, pp. 417-448.

- Метамодель;
- Унифицированные метаданные;
- Хранилища метаданных;
- Происхождение данных;
- Анализ воздействия;
- Анализ зависимостей;
- Процесс контроля метаданных.

Вставка 26. Основы метаданных

В целом, метаданные – это «данные о данных», и при правильном управлении они генерируются всякий раз, когда данные создаются, собираются, добавляются, удаляются или обновляются в любом хранилище данных и системе данных в рамках архитектуры данных организации.

Метаданные предоставляют ряд очень важных преимуществ для организации, в том числе:

- **Согласованность определений** – Метаданные содержат информацию о данных, которые помогают согласовать разницу в терминологии, как например: «клиенты» и «заказчики», «доходы» и «продажи» и т. д.
- **Четкость взаимоотношений** – Метаданные помогают устранить неоднозначность и несоответствия при определении связей между сущностями, хранящимися в среде данных. Например, если клиент объявляет «получателя» в одной заявке, а этот получатель называется «участником» в другой заявке, определения метаданных помогут прояснить ситуацию.
- **Ясность происхождения данных** – Метаданные содержат информацию о происхождении конкретного набора данных и могут быть достаточно детальными для определения информации на уровне атрибута. Метаданные могут поддерживать допустимые значения для атрибута данных, его соответствующего формата, местоположения, владельца и распорядителя. В оперативном плане метаданные могут содержать: проверяемую информацию о пользователях, приложениях и процессах, которые создают, удаляют или изменяют данные; точная временная метка изменения; и разрешение, которое было использовано для выполнения этих действий.

Взято (с изменениями) из: Alex Berson and Larry Dubov, “The benefits of metadata and implementing a metadata management strategy”, TechTarget. Доступно по <https://searchchannel.techtarget.com/feature/The-benefits-of-metadata-and-implementing-a-metadata-management-strategy> (accessed on 8 January 2019).

4.2.10 Управление качеством данных

Управление качеством данных относится к: «Планированию, внедрению и контролю

действий, в которых применяются методы управления качеством к данным, чтобы обеспечить их пригодность для использования и удовлетворения нужд потребителей данных». ²²³

Цели состоят в том, чтобы:

- Разработать принцип управления, при котором данные соответствовали целям, исходя из требований потребителей данных;
- Разработать стандарты, требования и спецификации для контроля качества данных как часть жизненного цикла данных;
- Определить и внедрить процессы для измерения, мониторинга и отчетности по уровням качества данных;
- Выявить и поддерживать возможности улучшения качества данных за счет улучшения процессов и систем.

Деятельность включает в себя:

1. Определение данных высокого качества.
2. Разработка стратегии обеспечения качества данных.
3. Определение объема для выполнения первоначальной оценки –
 - Определить критические данные;
 - Определить существующие правила и закономерности.
4. Выполнение начальной оценки качества данных –
 - Определить и расставить приоритеты;
 - Выполнить анализ основных причин проблем.
5. Определение приоритетов для улучшений –
 - Приоритизировать действия на основе бизнес-воздействия;
 - Разработать профилактические и корректирующие действия;
 - Подтвердить запланированные действия.
6. Разработка и развертывание операций по обеспечению качества данных –
 - Разработать операционные процедуры по обеспечению качества данных;
 - Скорректировать качество данных.

Основными результатами являются:

223 Данный подраздел взят из: DAMA International, Data Management Body of Knowledge, second edition (DAMA-DMBOK 2) (New Jersey, Technics Publications, 2017), Chapter 13, pp. 451-494.

- Улучшение качества данных;
- Оперативный анализ управления данными;
- Профили данных;
- Отчеты о сертификации качества данных;
- Соглашения об уровне обслуживания по обеспечению качества данных.

Вставка 27. Шесть составляющих качества данных

На первый взгляд, очевидно, что качество данных – это очистка плохих данных – данных, которые отсутствуют, неверны или недействительны в некотором роде. Но для того, чтобы обеспечить достоверность данных, важно понять следующие шесть ключевых аспектов качества данных, чтобы оценить, насколько эти данные «плохие»:

1. **Полнота** определяется как ожидаемая комплексность. Данные могут быть полными, даже если дополнительные данные отсутствуют. До тех пор, пока данные соответствуют ожиданиям, они считаются полными.
2. **Согласованность** означает, что данные во всех системах отражают одну и ту же информацию, и они синхронизируются друг с другом по всей организации.
3. **Соответствие** означает, что данные следуют набору стандартных определений данных, таких как тип данных, размер и формат. Например, дата рождения клиента представляется в формате «дд/мм/ггг».
4. **Точность** – это степень, с которой данные правильно отражают реальный объект или описываемое событие.
5. **Целостность** означает достоверность данных во всех взаимосвязях и обеспечивает условие того, что все данные в базе данных могут быть отслежены и связаны с другими данными.
6. **Своевременность** указывает на доступность информации, когда она ожидается и необходима. Своевременность данных очень важна.

Взято (с изменениями) из: Somasekhar Thatipamula, "Data Done Right: 6 Dimensions of Data Quality", Smart Bridge, 9 August 2013. Доступно по <https://smartbridge.com/data-done-right-6-dimensions-of-data-quality/>.

4.2.11 Оценка прогресса

Одним из способов определения состояния управления данными в организации является оценка зрелости управления данными.

Оценка зрелости управления данными – это «метод ранжирования процессов

обработки данных в организации для характеристики текущего состояния управления данными и его влияния на организацию».²²⁴

Правительства, стремящиеся определить свой уровень зрелости управления данными, могут использовать (общие) шесть уровней зрелости управления данными следующим образом:

- **Уровень 0 – Нет способности:** Нет организованных методов управления данными или формальных корпоративных процессов для управления данными.
- **Уровень 1 – Начальный/Специальный:** Практика управления данными общего назначения с использованием ограниченного набора инструментов, практически без управления.
- **Уровень 2 – Повторяемый:** Появление согласованных инструментов и определение роли для поддержки выполнения процесса. Организация начинает использовать централизованные инструменты и обеспечивает больше контроля над управлением данными.
- **Уровень 3 – Определенный:** Новые возможности управления данными. Внедрение и институционализация масштабируемых процессов управления данными и представление управления данными в качестве организационного инструмента.
- **Уровень 4 – Управляемый:** Стандартизированные инструменты для управления данными от настольных компьютеров до инфраструктуры в сочетании с хорошо сформированными функциями централизованного планирования и управления. Достижение этого уровня представляют собой ощутимое повышение качества данных и возможностей всей организации, таких как сквозной аудит данных.
- **Уровень 5 – Оптимальный:** Распространение данных контролируется для предотвращения ненужного дублирования. Хорошо понятные метрики используются для управления и измерения качества данных и процессов.²²⁵

Повышенная зрелость управления данными обеспечивает не только доступность и качество данных, но и успешное будущее организации. Как утверждает Мэтью Чейз:

Те ведомства, которые оптимизировали стратегии управления данными, теперь будут подготовлены к успешному будущему, чтобы идти в ногу со многими продвинутыми приложениями для данных, которые появляются, включая визуализацию данных, машинное обучение, искусственный интеллект и так далее. Те, кто оказался ближе к неактивным аутсайдерам, вероятно, продолжат отставать, поскольку технологические инновации становятся все более сложными и развиваются еще более быстрыми темпами.²²⁶

224 DAMA International, *Data Management Body of Knowledge*, second edition (DAMA-DMBOK 2) (New Jersey, Technics Publications, 2017), p. 533.

225 Там же, pp. 534-536.

226 Matthew Chase, "The state of data management in the public sector in 2018", *Experian*, 8 February 2018. Доступно по <https://www.edq.com/blog/the-state-of-data-management-in-the-public-sector-in-2018/>.

5. Аналитика данных, алгоритмическая необъективность и интуиция

Задачи данного раздела:

- Объясните четыре типа анализа данных – описательный, диагностический, прогнозирующий и предписывающий;
 - Изучить понятие алгоритмической необъективности;
 - Исследовать роль интуиции при принятии решений.
-

5.1 Аналитика данных

Аналитика данных создает новые значения в данных.

Это «стремление извлекать смысл из необработанных данных с использованием специализированных компьютерных систем ... которые преобразуют, систематизируют и моделируют данные, чтобы сделать выводы и определить закономерности».²²⁷

Существует четыре типа аналитики: описательная, диагностическая, прогнозирующая и предписывающая.²²⁸

5.1.1 Описательная аналитика

Описательная аналитика дает ответ на вопрос: что произошло?²²⁹

Агрегация и извлечение данных – два основных метода, используемых в описательной аналитике.²³⁰ Результаты включают в себя сводку расходов, табулирование социальных показателей на государственных сайтах в Facebook и Twitter, а также отчеты об общих тенденциях, таких как уровень инфляции или занятости.

Вставка 28. Описательная аналитика и бизнес-аналитика

Описательная аналитика – один из самых основных элементов бизнес-аналитики, который будет использовать организация. Несмотря на то, что описательная аналитика может быть

227 Informatica, “What is Data Analytics”. Доступно по <https://www.informatica.com/services-and-training/glossary-of-terms/data-analytics-definition.html#fbid=goi4ogkVXL0> (accessed on 8 January 2019).

228 Alex Bekker, “4 types of data analytics to improve decision-making”, *ScienceSoft*, 11 July 2017. Доступно по <https://www.scnsoft.com/blog/4-types-of-data-analytics>.

229 Halo, “Descriptive, Predictive, and Prescriptive Analytics Explained”. Доступно по <https://halobi.com/blog/descriptive-predictive-and-prescriptive-analytics-explained/> (accessed on 8 January 2019).

230 Anushka Mehta, “Four Types of Business Analytics to Know”, *Analytics Insight*, 13 October 2017. Доступно по <https://www.analyticsinsight.net/four-types-of-business-analytics-to-know/>.

очень узкоотраслевой – например, связанной с сезонными колебаниями времени завершения доставки – она также может включать общепринятые меры, которые используются в финансовой сфере. Рентабельность инвестированного капитала представляет собой описательную аналитику, созданную путем взятия трех показателей данных – чистого дохода, дивидендов и общего капитала – и превращения их в простой для понимания процент, который можно использовать для сравнения эффективности одной компании с другими. Как правило, чем крупнее и сложнее компания, тем более описательной аналитикой она будет пользоваться для измерения своей эффективности.

Взято (с изменениями) из: Investopedia, “Descriptive Analytics”, 10 March 2018. Доступно по <https://www.investopedia.com/terms/d/descriptive-analytics.asp#ixzz5NYWlugiV>.

Описательная аналитика редко пытается исследовать или устанавливать причинно-следственные связи. Она не может дать ответ на такие важные вопросы, как: *Как мы можем избежать данную проблему? Как мы можем распространить это успешное решение?* Они рассматриваются диагностической аналитикой.

5.1.2 Диагностическая аналитика

Диагностическая аналитика сравнивает исторические данные с другими данными, чтобы ответить на вопрос: *Почему что-то случилось?*

Диагностическая аналитика фокусируется на процессах и причинах. Используемые методы включают детализацию, обнаружение данных, анализ данных и корреляции.²³¹ Алгоритмы обучения для классификации и регрессии также подпадают в данный тип аналитики.²³²

Диагностическая аналитика используется для:

- **Выявления аномалии** – Основываясь на результатах описательного анализа, аналитики должны определить области, которые требуют дальнейшего изучения, поскольку там находятся вопросы, на которые невозможно ответить, просто взглянув на данные;
- **Детализации в аналитике (открытие)** – Аналитики должны определить источники данных, которые помогут им объяснить эти аномалии. Зачастую этот шаг требует от аналитиков поиска закономерностей за пределами существующих наборов данных, а также может потребоваться извлечение данных из внешних источников для выявления корреляций и определения того, является ли какой-либо из них причинным по своей природе;
- **Определения причинно-следственных связей** – Скрытые отношения выявляются при рассмотрении событий, которые могли привести к выявленным

231 CornerStone, “Diagnostic Analytics”. Доступно по <https://www.cornerstoneondemand.com/glossary/diagnostic-analytics> (accessed on 8 January 2019).

232 Anushka Mehta, “Four Types of Business Analytics to Know”, *Analytics Insight*, 13 October 2017. Доступно по <https://www.analyticsinsight.net/four-types-of-business-analytics-to-know/>.

Вставка 29. Как выполнять диагностическую аналитику

1. Определите, что стоит исследовать

Первый шаг в диагностической аналитике – найти объект для исследования. Как правило, это что-то плохое, например, снижение дохода или числа клики, но это также может быть неожиданное повышение производительности.

2. Выполните анализ

Диагностическая аналитика может быть такой же простой, как и поиск единственной основной причины – например, доход упал в прошлом месяце из-за того, что количество новых клиентов сократилось. Однако для более сложного анализа может потребоваться несколько наборов данных и поиск корреляции с использованием регрессионного анализа.

На этом шаге необходимо найти статистически достоверную взаимосвязь между двумя наборами данных, где рост (или падение) одного вызывает рост (или падение) другого.

Более продвинутые методы в этой области включают анализ данных и анализ основных компонентов, но простой регрессионный анализ – отличное начало для работы.

3. Выполните выборочный отбор ваших оценок

Несмотря на то, что может быть интересно, какие различные факторы способствовали изменению производительности, не стоит перечислять все возможные причины в отчете. Вместо этого аналитик должен стремиться обнаружить единственный или не более двух наиболее влиятельных факторов в диагностируемой проблеме.

4. Четко сформулируйте свое заключение

В заключение, аналитический отчет должен содержать вывод с очень четким обоснованием. Не обязательно включать всю подготовительную работу, но вы должны:

- Определите проблему, которую вы анализируете;
- Укажите, почему вы думаете, что это произошло;
- Предоставьте подтверждающие доказательства.

Взято (с изменениями) из: Jeff Rajeck, "Analytics approaches every marketer should know #2: Diagnostic analytics", Econsultancy, 31 July 2017. Доступно по <https://www.econsultancy.com/blog/69300-analytics-approaches-every-marketer-should-know-2-diagnostic-analytics>.

5.1.3 Прогностическая аналитика

Прогностическая аналитика использует статистические модели и методы прогнозов, чтобы ответить на вопрос: *Что может случиться?*²³⁴

Она предоставляет организациям практические выводы, объединяя исторические данные в целях определения закономерностей в данных, и применяет статистические модели и алгоритмы для сбора взаимосвязей между различными наборами данных.

Частный сектор использует прогностическую аналитику для предвидения поведения покупателей и закономерностей покупок, выявления тенденций в деятельности по продажам и прогнозирования спроса на ресурсы для цепочки поставок, операций и запасов.²³⁵ Это наглядно показано в практике Otto – второго по величине интернет-магазина Германии.²³⁶ Otto позволяет своей ИКТ-системе приобретать около 200 000 единиц товара в месяц у сторонних партнеров без участия человека. Чем объясняется доверие к системе? ИКТ-система Otto анализирует около трех миллиардов прошлых транзакций и 200 переменных (включая прошлые продажи, поиски на веб-сайте Otto и информацию о погоде) для предсказания (с точностью до 90%), что клиенты купят за неделю до заказа.

Вставка 30. Розничная торговля использует прогностическую аналитику для обслуживания беременных женщин

Эндрю Поул был нанят компанией Target (крупным американским ритейлером) для определения тех уникальных моментов в жизни потребителей, когда их покупательские привычки становятся особенно уступчивыми, и правильная реклама или предоставление купонов заставят их начать тратить по-новому.

Среди жизненных событий нет ничего более важного, чем рождение ребенка. В этот момент привычки новых родителей более податливы, чем в любое другое время их взрослой жизни. Если компании могут идентифицировать беременных покупателей, они могут заработать миллионы.

При просмотре данных с помощью компьютера Поул смог идентифицировать около 25 продуктов, которые при совместном анализе позволили ему присвоить каждому покупателю индекс «возможности беременности». Что еще более важно, он мог оценить срок беременности женщины с точностью до небольшого промежутка времени, поэтому Target мог отправлять купоны, приуроченные к очень конкретным этапам ее беременности.

Примерно через год после того, как Поул создал свою модель прогнозирования беременности, в Target за пределами Миннеаполиса вошел мужчина и потребовал встречи с менеджером. Он сжимал купоны, которые были отправлены его дочери, и он был зол, по словам сотрудника, который участвовал в разговоре.

234 Halo, “Descriptive, Predictive, and Prescriptive Analytics Explained”. Доступно по <https://halobi.com/blog/descriptive-predictive-and-prescriptive-analytics-explained/> (accessed on 8 January 2019).

235 Там же

236 Economist, “How Germany’s Otto uses artificial intelligence”, 12 April 12, 2017. Доступно по <https://www.economist.com/news/business/21720675-firm-using-algorithm-designed-cern-laboratory-how-germanys-otto-uses?etear=sasexpectexceptional>.

«Моя дочь получила это по почте!» - сказал он. «Она все еще учится в старшей школе, и вы отправляете ей купоны на детскую одежду и детские кроватки? Вы пытаетесь убедить ее забеременеть?»

Менеджер понятия не имел, о чем говорил мужчина. Он посмотрел на адрес. Конечно же, купоны были адресованы дочери мужчины и содержали рекламу одежды для беременных, детской мебели и фотографии улыбающихся младенцев. Менеджер извинился, а затем позвонил через несколько дней, чтобы снова извиниться.

По телефону, однако, голос отца был несколько смущенным. «Я поговорил с моей дочерью», - сказал он. «Оказывается, в моем доме произошли какие-то события, о которых я не был полностью осведомлен. Она должна родить в августе. Я должен извиниться перед вами».

Взято (с изменениями) из: Charles Duhigg, "How Companies Learn Your Secrets", New York Times Magazine, 16 February 2012. Доступно по <https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=1&r=1&hp>.

Интересное использование прогностической аналитики для государственного управления заключается в анализе настроений. Например, правительства могут предсказать настроения граждан по конкретным вопросам в качестве позитивных, негативных или нейтральных путем анализа сообщений в социальных сетях.²³⁷

Прогностическая аналитика использует такие алгоритмы машинного обучения, как случайные леса (random forests) и метод опорных векторов, и статистику для изучения и тестирования данных. Самые популярные инструменты для прогностической аналитики включают Python, R и RapidMiner.²³⁸

5.1.4 Предписывающая аналитика

Предписывающая аналитика подсказывает, какие действия предпринять.

Она использует комбинацию таких методов и инструментов, как бизнес-правила, процедуры машинного обучения и компьютерного моделирования, и применяет их к различным наборам данных, включая исторические и транзакционные данные, а также данные в реальном времени.²³⁹

Компании используют предписывающую аналитику для «оптимизации производства, планирования и инвентаризации в цепочке поставок, чтобы обеспечить доставку необходимых продуктов в нужное время и улучшение качества обслуживания

237 Anushka Mehta, "Four Types of Business Analytics to Know", *Analytics Insight*, 13 October 2017. Доступно по <https://www.analyticsinsight.net/four-types-of-business-analytics-to-know/>.

238 Там же

239 Halo, "Descriptive, Predictive, and Prescriptive Analytics Explained". Доступно по <https://halobi.com/blog/descriptive-predictive-and-prescriptive-analytics-explained/> (accessed on 8 January 2019).

Вставка 31. Совершенствование сектора здравоохранения с помощью предписывающей аналитики

Когда поставщики медицинских услуг объединят такие данные, как записи пациентов, медицинские данные, экономические данные, демографические и социографические данные, информацию о тенденциях в области здравоохранения и данные больниц, они смогут предложить более качественное медицинское обслуживание за меньшие деньги, улучшить будущие капитальные вложения в новые объекты или больничное оборудование, а также повысить эффективность больниц.

Комбинация различных наборов данных также может быть использована для предложения врачам рекомендаций в отношении наилучшего возможного лечения пациента. Благодаря объединению и анализу нескольких наборов данных, Центру Aurora Health Care удалось улучшить медицинское обслуживание и снизить показатели повторного приема на 10 процентов, что позволило сэкономить 6 миллионов долларов США в год.

Кроме того, фармацевтические организации могут извлечь выгоду из предписывающей аналитики, улучшая разработку лекарств и сокращая время выхода на рынок новых медицинских препаратов. Процесс моделирования лекарств может снизить время разработки медикаментов, и становится легче найти подходящего пациента для клинических испытаний, основанных на нескольких переменных.

Взято (с изменениями) из: Mark van Rijmenam, "The Future of Big Data? Three Use Cases of Prescriptive Analytics", Dataflog. Доступно по <https://dataflog.com/read/future-big-data-use-cases-prescriptive-analytics/668> (accessed on 8 January 2019).

В то время как предписывающая аналитика считается «будущим больших данных, ... но до того момента, как она станет общераспространенным, еще далеко». ²⁴¹

5.1.5 Аналитика данных в правительстве

Правительства могут использовать аналитику данных для улучшения предоставления государственных услуг, стимулирования экономического развития за счет повышения эффективности использования и распределения ресурсов, а также сокращения мошенничества, растрат и злоупотреблений. ²⁴²

240 Там же

241 Mark van Rijmenam, "The Future of Big Data? Three Use Cases of Prescriptive Analytics", *Dataflog*. Доступно по <https://dataflog.com/read/future-big-data-use-cases-prescriptive-analytics/668> (accessed on 8 January 2019).

242 Sid Frank and Traci Gusher, "Better data, better government: Effective use of data and analytics at all levels deliver improved citizen services and outcomes", *KPMG Government Institute*, June 2016, p. 5. Доступно по <https://assets.kpmg/content/dam/kpmg/pdf/2016/06/co-gv-6-better-data,-better-government.pdf>.

Аналитика данных в правительстве включает:

- **В администрировании** – Определение наименее затратного поставщика с наивысшим рейтингом, производительность людей на основе взятых больничных листов, неэффективных отделов, типов необходимых навыков, а также, превысят ли расходы выделенный бюджет и почему;
- **В социальных службах** – В частности, в сфере охраны детства при определении наиболее подверженных риску детей, тех, кто с наименьшей вероятностью будет воссоединен, и выявлении наиболее подходящих услуг для каждого ребенка;
- **В органах правопорядка** – Мониторинг, отслеживание и отображение статистики преступности на ежедневной основе. В сочетании с визуализацией геопространственных данных аналитика позволяет полицейским департаментам принимать более обоснованные решения о том, куда направлять полицейские ресурсы, что повышает безопасность и эффективность затрат;
- **В здравоохранении** – Обеспечение научных достижений, повышение эффективности в медицинских учреждениях, развитие профилактики, диагностики и лечения заболеваний, а также помощь в наблюдении за заболеваниями (см. Вставку 32).

Вставка 32. Аналитика и наблюдение за заболеваниями

Точная и своевременная информация о глобальных проблемах здравоохранения является ключом к возможности быстрой оценки и реагирования на возникающие риски для здоровья во всем мире. Агентство по вопросам общественного здравоохранения Канады разработало Глобальную информационную сеть общественного здравоохранения (GPHIN: Global Public Health Intelligence Network). Информация от GPHIN предоставляется Всемирной организации здравоохранения, правительствам других стран и неправительственным организациям, которые затем могут быстро реагировать на инциденты в области общественного здравоохранения.

GPHIN - это безопасная интернет-система раннего предупреждения, которая собирает предварительные отчеты, имеющие большое значение для здоровья населения, в режиме реального времени, 24 часа в сутки, 7 дней в неделю. Эта уникальная многоязычная система собирает и распространяет соответствующую информацию о вспышках заболеваний и других событиях в области общественного здравоохранения путем мониторинга глобальных источников информации, таких как новостные ленты и веб-сайты. Данный мониторинг осуществляется на восьми языках с машинным переводом, который используется для перевода неанглоязычных статей на английские и английские статьи на другие языки. Информация фильтруется на предмет актуальности с помощью автоматизированного процесса, который затем дополняется анализом, осуществляемым людьми. Результаты классифицированы и сделаны доступными для пользователей. Уведомления о событиях общественного здравоохранения, которые могут иметь серьезные последствия для здоровья населения, немедленно передаются пользователям.

GPHIN имеет широкую сферу применения. Она отслеживает такие события, как вспышки заболеваний, инфекционные болезни, зараженные продукты питания и вода, биотерроризм и воздействие химических веществ, стихийные бедствия, а также вопросы, связанные с

безопасностью продуктов, лекарств и медицинских изделий.

Взято (с изменениями) из: *Abla Mawudeku and Michael Blench, "Global Public Health Intelligence Network (GPHIN)", n.d. Доступно по <http://mt-archive.info/MTS-2005-Mawudeku.pdf> (accessed on 8 January 2019).*

Существуют политические, организационные, культурные, ресурсные и технические барьеры, препятствующие более широкому использованию аналитики в государственном управлении.²⁴³

Правительства сталкиваются с техническими проблемами, которые включают следующее: данные в различных форматах; разные степени качества данных; многочисленные распределенные хранилища данных; устаревшие инфраструктуры управления данными; отсутствие доступа к частному сектору, содержащему большие данные; и недоступность передовых аналитических инструментов.²⁴⁴

Организационные проблемы включают в себя: отсутствие организационных структур, процессов или инструментов управления данными; отсутствие спонсоров/лидеров по внедрению управления данными; нежелание персонала обмениваться данными либо на основании обоснованных правовых/политических соображений, либо на основе жестких исторических методов управления; ограниченное понимание «искусства возможного» в использовании передовой аналитики и больших данных для получения ценной информации; и изменение приоритетов.²⁴⁵

5.2 Алгоритмическая необъективность

Использование аналитики включает в себя «применение алгоритмического или механического процесса для получения информации».²⁴⁶

Алгоритм – это «Набор математических инструкций или правил, которые, особенно с помощью компьютера, помогут вычислить ответ на проблему».²⁴⁷ Его эффективность зависит от данных – качество данных «влияет на смысл выполнения аналитики и придает уверенность в принятии решений на основе результатов».²⁴⁸

Алгоритмы используются для обработки (нахождения смысла) больших данных. Организации – частные и государственные – полагаются на алгоритмы для решения проблем и принятия решений, потому что они «способны обрабатывать гораздо

243 Там же

244 Там же

245 Там же

246 Avantika Monnappa, "Data Science vs. Big Data vs. Data Analytics", *simplilearn*, 2 March 2018. Доступно по <https://www.simplilearn.com/data-science-vs-big-data-vs-data-analytics-article>.

247 Cambridge Dictionary, "Algorithm". Доступно по <https://dictionary.cambridge.org/dictionary/english/algorithm> (accessed on 8 January 2019).

248 Royal Academy of Engineering, "Algorithms in decision-making: A response to the House of Commons Science and Technology Committee inquiry into the use of algorithms in decision-making", April 2017, p. 3. Доступно по <https://www.raeng.org.uk/publications/responses/algorithms-in-decision-making>.

больший диапазон входных данных и переменных для принятия решений и могут делать это со скоростью и надежностью, которые намного превышают возможности человека».²⁴⁹

Королевская инженерная академия Великобритании приводит следующую рекомендацию:

Правительству, компаниям и государственным органам необходимо будет рассмотреть вопрос об использовании ими алгоритмов при принятии решений, при этом провести широкие консультации и обеспечить наличие механизмов для выявления и устранения любых ошибок или непреднамеренных последствий принятых решений.²⁵⁰

Это потому, что алгоритмы не являются объективными или нейтральными. Алгоритмическая необъективность имеет место, когда человеческие предрассудки и предвзятость присутствуют на этапе проектирования.²⁵¹ Следовательно, дискриминация (неодинаковое отношение) встраивается в модель.

Чем больше организаций используют алгоритмы для поддержки (или замены) людей при принятии решений, тем больше алгоритмическая необъективность становится проблемой. Как отметил Уилл Найт: «Если предвзятость, скрывающаяся в алгоритмах, принимающих все более важные решения, останется непризнанной и непроверенной, это может иметь серьезные негативные последствия, особенно для более обездоленных сообществ и меньшинств».²⁵²

Вставка 32. Алгоритмическая необъективность

В 2015 году приложение Google для работы с фотографиями ошибочно пометило фотографию двух чернокожих как горилл, поскольку его алгоритм не был обучен на достаточном количестве изображений темнокожих людей. В другом случае «судья» конкурса красоты, работающего на основе искусственного интеллекта (ИИ), в основном выбирал белых участников в качестве победителей, потому что его обучение проводилось на изображениях белых людей.

Это тривиальные случаи, которые можно легко исправить, предоставив ИИ больше выборок в областях, где у него недостаточно данных. В других случаях, когда ИИ работает с огромными объемами существующих данных в бескрайнем море онлайн-информации, поиск и противодействие необъективности становится намного более сложным.

Алгоритмическая необъективность может иметь еще более разрушительный эффект в других областях, например, как в правоохранительных органах. В 2016 году расследование ProPublica

249 Keith Kirkpatrick, "Battling Algorithmic Bias", *Communications of the ACM*, vol. 59, no. 10 (October 2016), pp. 16-17. Доступно по <https://cacm.acm.org/magazines/2016/10/207759-battling-algorithmic-bias/abstract>.

250 Royal Academy of Engineering, "Algorithms in decision-making: A response to the House of Commons Science and Technology Committee inquiry into the use of algorithms in decision-making", April 2017, p. 4. Доступно по <https://www.raeng.org.uk/publications/responses/algorithms-in-decision-making>.

251 Keith Kirkpatrick, "Battling Algorithmic Bias" *Communications of the ACM*, vol. 59, no. 10 (October 2016), pp. 16-17. Доступно по <https://cacm.acm.org/magazines/2016/10/207759-battling-algorithmic-bias/abstract>.

252 Will Knight, "Biased Algorithms Are Everywhere, and No One Seems to Care", *MIT Technology Review*, 12 July 2017. Доступно по <https://www.technologyreview.com/s/608248/biased-algorithms-are-everywhere-and-no-one-seems-to-care/>.

показало, что инструмент на основе искусственного интеллекта, используемый правоохранительными органами, с большей вероятностью объявляет чернокожих людей подверженными высокому риску рецидивизма по сравнению с белыми людьми. В некоторых штатах судьи полагаются на такие инструменты, чтобы решить, кто будет находиться в тюрьме, а кто выйдет на свободу, не проводя самостоятельных дополнительных расследований.

Подобные случаи могут случиться в других областях, таких как одобрение выдачи ссуды, где люди, которые недостаточно представлены, будут далее маргинализированы и лишены обслуживания. В сфере здравоохранения, где ИИ активно внедряется в диагностику и лечение заболеваний, алгоритмы могут нанести вред населению, данные которого не были включены в обучающую выборку.

На самом деле, если не принимать во внимание алгоритмическую необъективность, это может привести к усилению предвзятого отношения со стороны человека. Под иллюзией, что программное обеспечение не является предвзятым, люди склонны доверять решениям алгоритмов ИИ, забывая, что эти решения уже отражают их собственные предубеждения. В результате мы, не сомневаясь в них, будем принимать решения, основанные на искусственном интеллекте, и создадим более необъективные данные для этих алгоритмов, что приведет к дальнейшему росту необъективности.

Взято (с изменениями) из: Ben Dickson, "What is algorithmic bias?" TechTalks, 26 March 2018. Доступно по <https://bdtechtalks.com/2018/03/26/racist-sexist-ai-deep-learning-algorithms/>.

Алгоритмическая необъективность может возникнуть в результате проблем на разных этапах.²⁵³ Они включают:

- **Предвзятые или, по-другому, исходные данные низкого качества** – Данные могут быть необъективными, неполными или низкого качества, что может привести к тому, что алгоритм приведет к плохим и, возможно, дискриминационным результатам.
- **Плохо определенные правила** – Данные, используемые в качестве исходных данных для алгоритмических решений, могут быть плохо взвешены.
- **Недостаток контекстной осведомленности** – Определение качества данных обучения и надежности правил и весов часто зависят от конкретного контекста. Алгоритмы, которые хорошо работают в контексте, для которого они были разработаны, могут отличаться, если используются в другом контексте.
- **Петли обратной связи** – Алгоритмы не работают в вакууме. Их деятельность влияет на среду, из которой они извлекают данные, которые они, в свою очередь, используют в качестве исходных данных. Предвзятый алгоритм может усилить свои необъективности, что можно считать самоисполняющимся циклом ожидаемых результатов.²⁵⁴

Алгоритмическая необъективность, согласно Кейт Кроуфорд, может привести к двум типам ущерба: распределительный вред и репрезентативный вред.²⁵⁵

²⁵³ World Wide Web Foundation, *Algorithmic Accountability: Applying the concept to different country contexts* (Washington D.C., 2017), p. 9. Доступно по http://webfoundation.org/docs/2017/07/Algorithms_Report_WF.pdf.

²⁵⁴ Там же

²⁵⁵ Sidney Fussell, "AI Professor Details Real-World Dangers of Algorithm Bias [Corrected]", *Gizmodo*, 8 December

Распределительный вред возникает, когда алгоритм использует или удерживает определенные возможности или ресурсы на основе предвзятых предположений. Примером является алгоритм оценки банковского кредитного риска, который систематически отклоняет заявки на кредит для женщин.

Репрезентативный вред «возникает, когда системы усиливают зависимость некоторых групп по линии идентичности».²⁵⁶ Здесь технологии усиливают стереотипы. Например, алгоритм при идентификации восточных азиатов считает, что они мигают, когда они фактически улыбаются.

Для устранения алгоритмической необъективности и предотвращения создаваемого ущерба была предложена подотчетность алгоритмов. Подотчетность – это не только прозрачность, но и «обязанность сообщать и обосновывать принятие алгоритмических решений и смягчать любые негативные социальные последствия или потенциальный вред».²⁵⁷

В 2016 году сообщество «Справедливость, ответственность и прозрачность в машинном обучении» – группа ученых, разработчиков и исследователей в области компьютерных технологий – выпустило пять руководящих принципов для подотчетности алгоритмов. Принципы (см. Вставку 33) направлены на то, чтобы помочь разработчикам в разработке и реализации алгоритмических систем способами, подотчетными перед общественностью.

Вставка 33. Принципы подотчетности алгоритмов: справедливость, подотчетность и прозрачность в машинном обучении

1. **Справедливость** – Убедитесь, что алгоритмические решения не создают дискриминационных или несправедливых воздействий при сравнении разных демографических показателей.
2. **Объяснимость** – Убедитесь, что алгоритмические решения, а также любые данные, определяющие эти решения, могут быть объяснены конечным пользователям и другим заинтересованным сторонам в нетехнических терминах.
3. **Проверяемость** – Предоставьте заинтересованным третьим сторонам возможность исследовать, понимать и анализировать поведение алгоритма путем раскрытия информации, позволяющей осуществлять мониторинг, проверку или критику, в том числе посредством предоставления подробной документации, технически приемлемых интерфейсов программирования приложений (API) и допустимых условий использования.
4. **Ответственность** – Обеспечьте доступность видимых извне способов возмещения неблагоприятных индивидуальных или общественных последствий алгоритмической системы принятия решений и обозначьте внутри организации роль человека,

2017. Доступно по <https://gizmodo.com/microsoft-researcher-details-real-world-dangers-of-algo-1821129334>.

256 Там же

257 Nicholas Diakopoulos and Sorelle Friedler, "How to Hold Algorithms Accountable", *MIT Technology Review*, 17 November 2016. Доступно по <https://www.technologyreview.com/s/602933/how-to-hold-algorithms-accountable/>.

ответственного за своевременное решение таких проблем.

5. **Точность** – Идентифицируйте, зарегистрируйте и сформулируйте источники ошибок и неопределенностей по всему алгоритму и его источникам данных, чтобы можно было понять ожидаемые и наихудшие последствия, и проинформируйте о мерах по смягчению ущерба.

Взято (с изменениями) из: World Wide Web Foundation, *Algorithmic Accountability: Applying the concept to different country contexts* (Washington D.C., 2017), p. 11. Доступно по http://webfoundation.org/docs/2017/07/Algorithms_Report_WF.pdf.

Государственные органы начали предпринимать действия по продвижению алгоритмической подотчетности.

На местном уровне городской совет Нью-Йорка принял законопроект об алгоритмической подотчетности в декабре 2017 года.²⁵⁸ Данный законопроект требует создания целевой группы, которая будет изучать: (1) как городские ведомства используют алгоритмы для принятия решений, которые влияют на жизнедеятельность; (2) является ли какая-либо из систем дискриминационной по отношению к людям по признаку возраста, расы, религии, пола, сексуальной ориентации или гражданства; и (3) как сделать эти процессы принятия решений понятными для общественности.²⁵⁹

В отчете за 2015 год, озаглавленном «*Большие данные: использование возможностей, сохранение ценностей*», администрация Обамы признала «потенциал использования технологии больших данных для дискриминации отдельных лиц, преднамеренно или нечаянно, потенциально приводящих к дискриминационным результатам, сокращению возможностей и вариантов, доступных для них».²⁶⁰ В отчете было рекомендовано несколько мер по снижению данного риска, включая расширение технических знаний для предотвращения дискриминации и углубление понимания различий в определении расходов. В последующем отчете «*Большие данные: отчет об алгоритмических системах, возможностях и гражданских правах*» указывалось на необходимость принятия мер «обеспечения равных возможностей с момента разработки» для устранения потенциального вреда больших данных.²⁶¹

Великобритания приняла Концепцию этики данных со следующими принципами:

- Начните с определения понятных потребностей пользователей и общественных интересов;

258 Elizabeth Zima, “Could New York City’s AI Transparency Bill Be a Model for the Country?” *Government Technology*, 4 January 2018. Доступно по <http://www.govtech.com/policy/Could-New-York-Citys-AI-Transparency-Bill-Be-a-Model-for-the-Country.html>.

259 Lauren Kirchner, “New York City moves to create accountability for algorithms”, *Ars Technica*, 19 December 2017. Available at <https://arstechnica.com/tech-policy/2017/12/new-york-city-moves-to-create-accountability-for-algorithms/>.

260 Executive Office of the President of the United States, “Big Data: Seizing Opportunities, Preserving Values –Interim Progress Report”, February 2015. Доступно по https://obamawhitehouse.archives.gov/sites/default/files/docs/20150204_Big_Data_Seizing_Opportunities_Preserving_Values_Memo.pdf.

261 Executive Office of the President of the United States, “Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights”, May 2016. Доступно по https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf.

- Будьте в курсе соответствующего законодательства и практических руководств;
- Используйте данные, которые соответствуют потребностям пользователя;
- Изучите ограничения данных;
- Обеспечьте эффективными практическими методами и работайте в рамках ваших компетенций;
- Сделайте свою работу прозрачной и подотчетной;
- Внедряйте ответственное использование данных.²⁶²

Общее положение о защите данных (GDPR) Европейского союза (ЕС), вступившее в силу 25 мая 2018 года, также способствует обеспечению алгоритмической подотчетности за счет прозрачности.²⁶³

GDPR обязывает организации предоставлять субъектам данных необходимую информацию об автоматизированных системах, которые они используют для обработки персональных данных, чтобы они могли принимать взвешенные решения относительно того, следует ли отказаться от такой обработки данных. Учитывая, что автоматизированные системы, которые обрабатывают личные данные граждан, часто основываются на принципах машинного обучения, обозреватель отметил, что «процессы машинного обучения должны стать прозрачными – если не по-настоящему прозрачными, то, по крайней мере, намного менее похожими на черный ящик – для компаний, которые подпадают под GDPR для соответствия требованиям».²⁶⁴

5.3 Интуиция и принятие решений

Означает ли соответствующее использование алгоритмов и аналитики конец интуиции в процессе принятия решений?

Малкольм Гладуэлл в своей книге «*Озарение. Сила мгновенных решений*» 2005 года отстаивает непреходящее значение интуиции. В своей книге Гладуэлл описывает теорию «тонких срезов» – «способность нашего бессознательного находить закономерности в ситуациях и поведении на основе очень узких срезов опыта».²⁶⁵ Итак, насколько тонкий является тонким? Гладуэлл и другие ученые считают, что мы можем точно сделать выводы об эмоциях и отношениях людей, с которыми взаимодействуем, наблюдая их поведение только несколько секунд.²⁶⁶

Основываясь на исследованиях ученых, специализирующихся в вопросах поведения, Гладуэлл утверждает, что: (1) у всех нас есть возможность извлекать огромное

262 United Kingdom Department for Digital, Culture, Media and Sport, “Guidance: Data Ethics Framework (updated 30 August 2018)”. Доступно по <https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework>.

263 Andre Burt, “Is there a ‘right to explanation’ for machine learning in the GDPR?” *International Association of Privacy Professionals*, 1 June 2017. Доступно по <https://iapp.org/news/a/is-there-a-right-to-explanation-for-machine-learning-in-the-gdpr/>.

264 Juraj Jánošík, “Transparency of machine-learning algorithms is a double-edged sword”, *welivesecurity*. 13 November 2017. Доступно по <https://www.welivesecurity.com/2017/11/13/transparency-machine-learning-algorithms/>.

265 Malcom Gladwell, *Blink: The Power of Thinking Without Thinking* (New York and Boston, Blackbay Books, 2005), p. 24.

266 Jeff Thompson, “Thin Slices & First Impressions”, *Psychology Today*, 24 March 2012. Доступно по <https://www.psychologytoday.com/us/blog/beyond-words/201203/thin-slices-first-impressions>.

количество полезной информации из самого тонкого среза опыта; и (2) «Если мы хотим научиться повышать качество принимаемых нами решений ... мы должны уважать тот факт, что можно узнать, не зная, почему мы это знаем и принимаем – иногда – нам от этого не лучше».²⁶⁷

Исследования показали точность тонких срезов в: первых впечатлениях от незнакомцев с самооценками; способности определять сексуальную ориентацию, работоспособность оператора кол-центра, рейтинги учителей и благонадежность продавцов; и обсуждениях между студентами-медиками и пациентами, интервьюерами и соискателями, а также студентами и руководителями.²⁶⁸

Научные исследования теории тонкого среза также меняют наше понимание процесса принятия решений. Исследователи обнаружили, что «когда мы просим людей обдумать, прежде чем они примут решение, оно, как правило, не так хорошо, как если бы они сделали его неосознанно».²⁶⁹ Кроме того, при определенных условиях больше информации не помогает, а может привести к неправильным решениям.²⁷⁰

Но могут ли политики полагаться только на интуицию?

Исследования выявили следующие условия, когда можно положиться на интуицию: наличие опыта, характер проблемы и доступное время.²⁷¹

Чем больше опыта в определенной области, тем больше шансов использовать интуицию для принятия решений по конкретным направлениям. Например, опытный судья обычно может определить, когда свидетель лжет. Сколько нужно для этого опыта? Согласно исследованиям, для выработки точных интуитивных суждений требуется 10-летний опыт работы в конкретной области (и в течение этих 10 лет необходимы повторение и обратная связь).²⁷²

Интуиция может быть использована для неструктурированных проблем – тех, в которых нет четких правил принятия решений или мало объективных критериев для принятия решения.²⁷³ С другой стороны, было бы неразумно использовать интуицию для решения проблем, которые имеют четкие правила принятия решений, объективные критерии и много данных, с помощью которых можно выполнить анализ.

Интуитивное суждение также хорошо работает в ситуациях, которые требуют быстрого решения, и для подробного анализа остается мало времени. Поведенческая наука

267 Malcom Gladwell, *Blink: The Power of Thinking Without Thinking* (New York and Boston, Blackbay Books, 2005), p. 53.

268 Jeff Thompson, "Thin Slices & First Impressions", *Psychology Today*, 24 March 2012. Доступно по <https://www.psychologytoday.com/us/blog/beyond-words/201203/thin-slices-first-impressions>.

269 Rosie Ifould, "Acting on impulse", *The Guardian*, 7 March 2009. Доступно по <https://www.theguardian.com/lifeandstyle/2009/mar/07/first-impressions-snap-decisions-impulse>.

270 Malcom Gladwell, *Blink: The Power of Thinking Without Thinking* (New York and Boston, Blackbay Books, 2005), p. 140.

271 Connson Chou Locke, "When It's Safe to Rely on Intuition (and When It's Not)", *Harvard Business Review*, 30 April 2015. Доступно по <https://hbr.org/2015/04/when-its-safe-to-rely-on-intuition-and-when-its-not>.

272 Там же

273 Там же

говорит нам: «Когда информации и времени мало, использование эвристики, такой как интуиция, часто может быть столь же эффективной, как и рациональный подход».²⁷⁴

Ключевой момент данного обсуждения заключается в том, что интуиция по-прежнему играет роль в принятии решений на основе данных.

Тем не менее, Гари Кляйн в «*Источниках силы*» отвергает идею о том, что для принятия решений можно полагаться только на шестое чувство и интуицию.²⁷⁵ Он утверждает, что принятие решений в реальном мире – это двухэтапный процесс, начинающийся с интуиции, поскольку лица, принимающие решения, осознают, как они должны ответить, после чего идет неторопливая оценка, поскольку они мысленно моделируют возможный ответ, чтобы увидеть, будет ли он работать.²⁷⁶

Эндрю Макафи и Эрик Бриньольфссон идут дальше:

Имеются убедительные доказательства того, что всякий раз, когда доступны варианты, использование одних только данных и алгоритмов обычно приводит к лучшим решениям и прогнозам, чем полагаться на мнение даже опытных людей и «экспертов».²⁷⁷

274 Там же

275 Gary Klein, *Sources of Power: How People Make Decisions – 20th Anniversary Edition* (Cambridge and London, MIT Press, 2017).

276 Там же, p. xvii.

277 Andrew McAfee and Erik Brynjolfsson, *Machine, Platform, Crowd: Harnessing Our Digital Future* (New York and London, W. W. Norton & Company, 2017), p. 64.

6. Благоприятная политика

В этом разделе рассматриваются стимулирующие политики и программы для реализации управления на основе данных. Они включают в себя политику и программы, которые обеспечивают:

- Укрепление процесса сбора данных, в том числе совершенствование национальной системы статистики, совершенствование гендерной статистики и стимулирование генерирования данных населением;
 - Институционализацию управления данными;
 - Повышение общественного доступа и доверия к данным путем принятия политики открытых государственных данных и укрепления конфиденциальности данных;
 - Готовность к большим данным.
-

6.1 Укрепление процесса сбора данных

Мероприятия по укреплению сбора данных включают в себя улучшение национальных статистических систем, улучшение гендерной статистики и стимулирование генерирования гражданами данных.

6.1.1 Совершенствование национальной системы статистики

Необходимость совершенствования национальных статистических систем хорошо известна.

В связи с этим была инициирована инициатива в рамках системы Организации Объединенных Наций, призванная обеспечить «основу для обсуждения, а также планирования и осуществления наращивания статистического потенциала, необходимого для достижения масштабов и целей» ЦУР.²⁷⁸

Кейптаунский глобальный план действий в области данных по устойчивому развитию направлен на: «планирование необходимых действий для регулярного получения качественных и своевременных данных для информирования об устойчивом развитии на требуемом уровне разбивки и охвата населения, включая наиболее уязвимых и труднодоступных групп».²⁷⁹

Основными принципами Кейптаунского глобального плана действий являются:

²⁷⁸ High-level Group for Partnership, Coordination and Capacity-Building for Statistics for the 2030 Agenda for Sustainable Development, "Cape Town Global Action Plan for Sustainable Development Data", adopted by the United Nations Statistical Commission at its 48th Session, March 2017, p. 2. Доступно по https://unstats.un.org/sdgs/hlg/Cape_Town_Global_Action_Plan_for_Sustainable_Development_Data.pdf

²⁷⁹ Там же, стр. 3.

- **Полнота охвата** – План должен охватывать все аспекты координации, создания и использования данных для устойчивого развития;
- **Подотчетность** – Современная подготовка статистических данных требует всестороннего взаимодействия между поставщиками данных, производителями и пользователями;
- **Сотрудничество** – В Плате признается решающая роль сотрудничества между странами, региональными организациями и другими международными организациями и заинтересованными сторонами в поддержке планов и усилий стран в укреплении потенциала.²⁸⁰

Кейптаунский глобальный план действий определяет шесть стратегических направлений для действий.

Первое направление – это «координация и стратегическое лидерство в области данных для устойчивого развития».²⁸¹ Цели состоят в том, чтобы обеспечить:

- Укрепление национальных статистических систем и координационной роли национальных статистических управлений;
- Повышение координации между национальными статистическими системами и региональными и международными организациями, занимающимися подготовкой данных и статистики для устойчивого развития.

Рекомендуемые ключевые действия включают следующее:

- Провести оценку потребностей для укрепления потенциала в области статистики на национальном уровне и оценку имеющихся ресурсов для удовлетворения этих потребностей;
- Создать и/или улучшить координационный механизм для сбора, обмена и распространения статистики в целях устойчивого развития между национальными статистическими системами, а также между национальными, региональными и международными статистическими системами.

Второе стратегическое направление для улучшения – это «инновации и модернизация национальных статистических систем».²⁸² Если конкретно, то это:

- Модернизация системы управления и институциональной базы, чтобы национальные статистические системы могли соответствовать потребностям и возможностям постоянно развивающихся экосистем данных;
- Модернизация статистических стандартов, особенно тех, которые направлены на облегчение интеграции данных и автоматизации обмена данными на разных этапах процесса подготовки статистических данных;
- Содействие применению новых технологий и новых источников данных в основной статистической деятельности.

280 Там же

281 Там же, стр. 4.

282 Там же

Предлагаемые действия для второго направления:

- Содействовать пересмотру статистических законов и нормативно-правовой базы, где это необходимо, в соответствии с Основными принципами официальной статистики, чтобы:
 - (а) повысить статус, независимость и координирующую роль национальных статистических ведомств;
 - (б) расширить их доступ к данным, в том числе улучшить обмен данными в рамках всей национальной статистической системы, и тем самым повысить их способность более эффективно реагировать на поступающие данные и статистические потребности;
 - (в) разработать механизм использования данных из альтернативных и инновационных источников в официальной статистике;
 - (г) повысить прозрачность и доступ общественности к официальной статистике;
 - (д) усилить доступность устойчивого финансирования для национальных статистических систем;
- Определить и внедрить стандартизированные структуры для обмена и интеграции данных и метаданных на основе социальных, экономических и экологических принципов устойчивого развития и на всех уровнях (глобальном, региональном, национальном и субнациональном) в соответствии с обменом статистическими данными и метаданными²⁸³ и соответствующими стандартами;
- Определить спецификации для совместимых технологий с открытым исходным кодом для обеспечения гибкости информационных систем, необходимой для стратегического использования новых и появляющихся технологий для сбора, обработки, распространения и анализа официальных данных.

Третье стратегическое направление – «укрепление основных статистических мероприятий и программ с особым акцентом на удовлетворение потребностей в мониторинге Повестки дня на период до 2030 года».²⁸⁴ В частности, обеспечить:

- Укрепление и расширение программ обследований домашних хозяйств, комплексных систем обследований, программ обследований предприятий и других экономических систем, программ переписи населения и учета жилищного фонда, программ учета гражданского состояния и естественного движения населения и Программы международных сопоставлений с учетом потребностей, изложенных в Повестке дня на период до 2030 года;
- Повышение качества национальных статистических реестров и расширение использования административных данных, интегрируя их с данными обследований и других новых источников данных, для составления комплексной социальной, экономической и экологической статистики с целью использования для осуществления Повестки дня на период до 2030 года;

283 See the official site for the SDMX community: A global initiative to improve Statistical Data and Metadata eXchange. Доступно по <https://sdmx.org/>.

284 High-level Group for Partnership, Coordination and Capacity-Building for Statistics for the 2030 Agenda for Sustainable Development, "Cape Town Global Action Plan for Sustainable Development Data", adopted by the United Nations Statistical Commission at its 48th Session, March 2017, pp. 5-6. Доступно по https://unstats.un.org/sdgs/hlg/Cape_Town_Global_Action_Plan_for_Sustainable_Development_Data.pdf.

- Укрепление и расширение Системы национальных счетов и Системы эколого-экономического учета;
- Интеграцию геопространственных данных в программы подготовки статистической информации на всех уровнях;
- Укрепление и расширение сбора данных по всем группам населения, чтобы никто не остался вне поля зрения;
- Укрепление и расширение сбора данных в тех направлениях, которые в настоящее время недостаточно охвачены официальной статистикой.

Некоторые из ключевых действий для третьего стратегического направления:

- Усилить гармонизацию и обеспечить участие стран в финансируемых на международном уровне программах обследований домашних хозяйств (таких как DHS, MICS, LSMS, Обследование детского труда, WHS, CWIQ и т. д.) путем укрепления существующей сети обследований домашних хозяйств и Межсекретариатской рабочей группы по обследованиям домашних хозяйств;
- Обеспечить разработку, стандартизацию и улучшение охвата реестров населения, имущества и предприятий для статистических целей;
- Поддерживать внедрение Системы национальных счетов и Системы эколого-экономического учета, принимая во внимание опыт стран на сегодняшний день и текущие потребности для улучшения практической реализации;
- Содействовать интеграции современных систем управления геопространственной информацией в основные программы подготовки статистических данных с фокусом на взаимодействие между двумя системами;
- Улучшить производство высококачественных, доступных, своевременных, надежных и дезагрегированных данных по всем характеристикам, соответствующим национальным условиям, для обеспечения того, чтобы никто не остался без внимания;
- Обеспечить разработку, стандартизацию и улучшение охвата и качества данных, которые сегодня не рассматриваются официальной статистикой.

Четвертое направление для улучшения – это «распространение и использование данных для обеспечения устойчивого развития».²⁸⁵ В частности, необходимо разрабатывать и продвигать инновационные стратегии для обеспечения распространения на должном уровне и использования данных для устойчивого развития.

Среди лучших рекомендуемых действий отметим следующие:

- Содействовать развитию технологической инфраструктуры для эффективного распространения данных;
- Разработать эффективные стратегии коммуникации и распространения данных и руководящих принципов для государственно-частного диалога, ориентированных на

политиков, законодателей, средства массовой информации, широкую общественность, экономическую деятельность и т. д.

Пятое направление – это «многосторонние партнерства для получения данных по устойчивому развитию».²⁸⁶ В частности, развивать и укреплять партнерские отношения национальных и международных статистических систем с правительствами, научными кругами, гражданским обществом, частным сектором и другими заинтересованными сторонами, участвующими в производстве и использовании данных для устойчивого развития.

Ключевые действия:

- Повысить прозрачность и доступность официальной статистики для общественности;
- Предоставить частые и периодические возможности для консультаций со всеми заинтересованными сторонами по вопросам подготовки и использования статистических данных для устойчивого развития.

Шестое направление – это «мобилизация ресурсов и координация усилий по наращиванию статистического потенциала».²⁸⁷ В частности, обеспечить наличие ресурсов для реализации необходимых программ и действий, как указано в глобальном плане действий (как внутри страны, так и в рамках международного сотрудничества).

Самые распространенные действия:

- Предоставить обзор потребностей в потенциале на основе выполненных или существующих оценок потребностей и рассмотреть подходящие соответствия между типами помощи и потребностей;
- Определить и скоординировать существующие ресурсы, включая трехсторонние механизмы сотрудничества и по линии Юг-Юг, для стратегического удовлетворения потребностей и выявления пробелов в ресурсах;
- Разработать программу для наращивания статистического потенциала на основе потребностей в ресурсах.

Совершенствование национальных статистических систем будет способствовать развитию управления на основе данных и повышению статистических данных по гендерным вопросам.

6.1.2 Улучшение гендерной статистики

Главный статистик организации «ООН-женщины» отмечает, что, хотя мир производит огромные объемы данных, «существуют вопиющие белые пятна и зияющие пробелы»,

286 Там же

287 Там же

когда речь идет о данных, имеющих отношение к женщинам и девочкам.²⁸⁸

Гендерная статистика – это «научное представление и интерпретация статистических данных, которые соответствующим и полным образом отражают условия жизни и положение женщин и мужчин во всех сферах и областях политики».²⁸⁹ В частности, они представляют собой:

- Данные, которые собираются и представляются в разбивке по полу в качестве основной и общей классификации;
- Данные, которые отражают гендерные проблемы;
- Данные, основанные на концепциях и определениях, которые должным образом отражают разнообразие женщин и мужчин и охватывают все аспекты их жизни;
- Использование методов сбора данных, учитывающих стереотипы и социальные и культурные факторы, которые могут вызывать предвзятость по признаку пола (в том числе занижение экономической активности женщин, не полный учет девочек, их рождения или смерти или занижение данных о насилии в отношении женщин).²⁹⁰

В 2013 году Статистическая комиссия ООН выпустила минимальный набор гендерных показателей.²⁹¹ Этот набор включает 52 количественных показателя и 11 качественных показателей, сгруппированных в пять категорий: (1) экономические структуры и доступ к ресурсам; (2) образование; (3) здравоохранение и сопутствующие услуги; (4) общественная деятельность и принятие решений; и (5) защита прав женщин и детей.

Подобно показателям ЦУР, Минимальный набор гендерных показателей также разбит на три уровня, которые отражают проблемы, связанные с данными и методологией:

- Уровень 1 – Показатель является концептуально ясным с методологией и стандартами, принятыми на международном уровне, и данными, регулярно предоставляемыми странами с достаточным охватом, позволяющим отслеживать прогресс во времени;
- Уровень 2 – Показатель концептуально ясен, с международно-установленными методологией и стандартами, но данные не регулярно предоставляются странами;
- Уровень 3 – Показатель не имеет международно-установленных методологий или стандартов, и данные не регулярно предоставляются странами.

Если более подробно, то Экономическая и социальная комиссия ООН для Азии и Тихого океана (ЭСКАТО) в 2015 году обнародовала «базовый набор гендерных показателей для Азиатско-Тихоокеанского региона» в дополнение к глобальному Минимальному набору гендерных показателей.²⁹²

288 UN Women, "Take five with Papa Seck: Getting better at gender data—why does it matter?" 21 September 2016.

Доступно по <http://www.unwomen.org/en/news/stories/2016/9/feature-story-take-five-with-papa-seck-on-gender-data>.

289 Gender Stats, "Gender Statistics in 5 Charts". Доступен по <http://genderstats.org/> (accessed on 8 January 2019).

290 European Institute for Gender Equality, "Gender Statistics". Доступно по <http://eige.europa.eu/gender-mainstreaming/methods-tools/gender-statistics> (accessed on 8 January 2019).

291 United Nations Statistics Division, "Minimum Set of Gender Indicators". Доступно по <https://genderstats.un.org/#/home> (accessed on 8 January 2019).

292 ESCAP, "Core set of gender indicators for Asia and the Pacific: Note by the secretariat", Committee on Statistics,

Азиатско-тихоокеанский базовый набор гендерных показателей состоит из следующего:

- **Шесть основных категорий** – Они предназначены для предоставления информации по общерегиональным вопросам гендерного равенства и расширения прав и возможностей женщин;
- **Пять дополнительных категорий** – Они связаны с вопросами гендерного равенства и расширения прав и возможностей женщин, которые рассматриваются в качестве приоритетов некоторыми подгруппами стран региона;
- **Приоритетная целевая группа (сельские женщины)** – Это подмножество показателей в основных категориях, которые отражают проблемы, которые вызывают особую обеспокоенность у сельских женщин – группы населения, которая, в целом, находится в неблагоприятном положении во многих странах Азиатско-Тихоокеанского региона;
- **Качественные показатели, связанные с национальными критериями** – Они отслеживают, как работает национальное законодательство в обеспечении гендерного равенства путем ратификации соответствующих международных конвенций и институционализации конкретных мер и политики по ликвидации дискриминации в отношении женщин и поощрению гендерного равенства.²⁹³

Для обеспечения гендерно-справедливого управления, основанного на данных, правительствам необходимо оценить свои данные и механизмы сбора данных в свете приведенных выше рекомендаций.

6.1.3 Стимулирование генерации данных гражданами

Данные, создаваемые гражданами, дополняют официальную статистику. Они также напрямую касаются вопроса справедливости на основе данных, поскольку данные, созданные гражданами, могут сделать видимыми отдельных лиц и группы, которые невидимы для правительства из-за отсутствия данных.

Важным элементом стимулирования данных, генерируемых гражданами, и решений, управляемых данными, является поддержка сообщества.

Существует, как минимум, шесть факторов, которые необходимы для создания поддержки сообществом решений, управляемых данными. Это:

1. **Создание права собственности** – Во многих случаях усилия по вовлечению сообществ предпринимаются после того, как лидеры разработали и дали старт инициативам, основанным на данных. Но взаимодействие должно начаться раньше, чтобы члены сообщества были заинтересованы в поддержке

Fourth Session, 23 January 2015 (E/ESCAP/CST(4)/10). Доступно по http://www.asiapacificgender.org/sites/default/files/pdf/statistics_documents/Core_Set_Gender_Indicators_Asia_Pacific.pdf.

инициативы;

2. **Учитывание сложности** – Лидеры должны адаптироваться к сложной системе воздействий, которые влияют на успех любого решения, управляемого данными;
3. **Работа с местными учреждениями** – Местные организации нарабатывают социальный капитал, который создает благоприятные условия для успешных действий на основе данных;
4. **Обеспечение равенства** – Слишком часто усилия по социальным изменениям не вовлекают правильное сочетание людей;
5. **Придание импульса усилиям** – Работа по привлечению сообществ требует принятия безотлагательных мер. Успехи на ранних этапах помогут сообществу описать деятельность в положительном ключе, которая может заменить существующие истории, основанные на очевидной неразрешимости социальных проблем. Кроме того, быстрые результаты позволят членам сообщества увидеть, что их участие имеет значение;
6. **Управление участниками вследствие изменений** – Руководители, которые переходят на новую структуру управления на основе данных, должны проследить за тем, как различные группы будут реагировать на это изменение. Хороший способ – это начать с разграничения между техническими и адаптационными проблемами (связанными с поведением и убеждениями людей).²⁹⁴

Те, кто стремится к тому, чтобы стимулировать генерацию данных гражданами, должны:

- **Совместить интересы ключевых заинтересованных сторон для возможности налаживания партнерских отношений** – Успешные проекты по данным, сгенерированным гражданами, объединяют участников с разными интересами в одних и тех же данных. Данные служат точкой соприкосновения для участников и являются объектом для сотрудничества. Часто существует разница между производством, использованием и распространением, и выгоды, связанные с каждой стадией, могут быть разными. Участники могут оценивать разные аспекты данных, и понимание того, как участники осознают эту пользу, является ключом к созданию многосторонних партнерских отношений.
- **Обеспечить, чтобы данные, сгенерированные гражданами, использовались несколькими способами для максимального повышения распространения и воздействия** – Чем больше способов использовать массивы данных, тем больше разных типов участников будут заинтересованы в данных. Для содействия различным вариантам использования разными участниками, данные должны быть доступными и представленными в совместимом формате.
- **Использовать существующие ресурсы и процессы, чтобы упростить процесс**

294 Melody Barnes and Paul Schmitz, "Community Engagement Matters (Now More Than Ever)", *Stanford Social Innovation Review* (Spring 2016). Доступно по https://ssir.org/articles/entry/community_engagement_matters_now_more_than_ever#.

производства и использования данных, сгенерированных гражданами – Это включает в себя использование технологий, с которыми граждане знакомы и которые они используют, а также использование установленных процедур и динамики поведения в группах, таких как существующие бюрократические процессы или общественные форумы.

- **Предусмотреть конкретные стимулы, которые зависят от контекста и цели** – Ключевые аспекты, которые следует учитывать, включают в себя нацеленность проекта на прямую связь с правительством или нет, а также социально-политическую и управленческую среду. Это включает, среди прочего, следующие вопросы: принимает ли правительство ответные меры, существует ли сильная правовая база и высокий уровень доверия, или есть ли достаточно информации по конкретной проблеме.²⁹⁵

Для правительства действия по стимулированию генерации данных гражданами включают следующее:

- Признать, что данные, генерируемые гражданами, и данные гражданского общества могут играть различную и дополняющую роль для данных государственного сектора. Правительству следует искать способы взаимодействия, признания и поддержки таких инициатив.
- Изучить модели финансирования и обеспечения устойчивости для данных, генерируемых гражданами, и данных гражданского общества. Во многих случаях сгенерированные гражданами данные следует считать дополнительными для сбора государственных данных, а не в качестве предлога для прекращения или сокращения финансирования последних.
- Поддерживать процедуры консультаций и участия, позволяющих гражданам и гражданскому обществу вносить вклад в сбор государственных данных, в том числе посредством проведения мероприятий, оперативных каналов коммуникации и процессов совместной разработки.
- Поддерживать дальнейшие исследования и разработки ресурсов в этой области, которые могут быть использованы для того, чтобы позволить инфраструктурам государственных данных лучше учитывать интересы и проблемы гражданского общества.²⁹⁶

6.2 Институционализация управления данными

Напомним, что управление данными – это: «Организация контроля качества для добавления четкости и дисциплины в процесс управления, использования, улучшения и

²⁹⁵ Danny Lammerhirt, Shazade Jaeson and Eko Presetyo, "Making Citizen Generated Data Work: Towards a Framework Strengthening Collaborations Between Citizens, Civil Society Organisations, and Others", *Data Shift*, March 2017. Доступно по http://civicus.org/thedatashift/wp-content/uploads/2017/03/Making-Citizen-Generated-Data-Work-short-report_.pdf.

²⁹⁶ Jonathan Gray, Danny Lammerhirt and Liliana Bounegru, "Changing What Counts: How Can Citizen-Generated and Civil Society Data be Used as an Advocacy Tool to Change Official Data Collection?" *Data Shift*, March 2016. Доступно по <http://civicus.org/thedatashift/wp-content/uploads/2016/03/changing-what-counts-2.pdf>.

защиты организационной информации».²⁹⁷ При этом устанавливаются «параметры для управления данными и их использования, создания процессов для решения проблем с данными и предоставления возможности пользователям принимать решения на основе высококачественных данных и хорошо управляемых информационных активов».²⁹⁸

Без хорошего управления данными организации не могут быть более эффективными, результативными и не могут стать более прозрачными.²⁹⁹

Для институционализации управления данными, правительства должны сначала понять свои факторы, обеспечивающие исполнение:

- **Стратегическое намерение** – ПОЧЕМУ управление данными представляет ценность; конечное состояние, к которому стремится правительство; и основополагающая политика, которая описывает мотивацию исполнительного руководства;
- **Модель зрелости управления данными** – Путь от того, что ЕСТЬ к тому, что ДОЛЖНО БЫТЬ относительно управления данными, информацией и знаниями. Модели зрелости управления данными предоставляют средства для оценки прогресса. Представляя промежуточные этапы, а также желаемое конечное состояние, модели зрелости помогают планировать, КАК правительство достигнет следующего уровня эффективности, а также КОГДА и ГДЕ в структуре правительства;
- **Организационные модели, матрицы ролей и ответственности** – КТО должен быть вовлечен в принятие решений, реализацию, мониторинг и поддержание работоспособности;
- **Руководящие положения или фреймворк** – ЧТО управляется, включая связанные понятия, компоненты и взаимосвязи между ними;
- **Методология для навигации по фреймворку** – Методы и процедуры, КАК ориентироваться в фреймворке, создавать артефакты, описывающие организацию, и поддерживать деятельность с течением времени;
- **Показатели эффективности работы** – Измерение и оценка прогресса и эффективности осуществления инициативы. По ним можно проследить обратно до стратегических намерений и связанных с ними моделей зрелости. Данные показатели должны постоянно оцениваться на предмет соответствия;
- **Оценка и безопасность государственных информационных активов** – Достоверная оценка данных и информации будет определять уровень инвестиций для обеспечения качества и соответствующей безопасности на протяжении жизненного цикла информационных активов.³⁰⁰

297 IBM, "The IBM Data Governance Council Maturity Model: Building a roadmap for effective data governance", October 2007. Доступно по https://www-935.ibm.com/services/uk/cio/pdf/leverage_wp_data_gov_council_maturity_model.pdf.

298 TechTarget, "Building an effective data governance framework", October 2013. Доступно по <https://searchdatamanagement.techtarget.com/essentialguide/Building-an-effective-data-governance-framework>.

299 Colin Wood, "Data Governance: The Public Sector's Next Big Frontier", *Government Technology*, 29 April 2014. Доступно по <http://www.govtech.com/data/Data-Governance.html>.

300 NASCIO, "Data Governance Part II: Maturity Models – A Path to Progress", March 2009, pp. 2-3. Доступно по

При разработке политики и процедур управления данными для достижения успеха необходимо преодолеть следующие препятствия:

- Определение правил и требований, включая интерпретацию и понимание правил, касающихся источников данных;
- При разработке политики получение согласия всех заинтересованных сторон;
- Разработка новых инструментов и программного обеспечения для управления данными;
- Стоимость внедрения политики;
- Несовместимые системы;
- Коллизии приоритетов внутри организации;
- Руководства должно понять, что необходимо;
- Налаживание управления проектом.³⁰¹

Следующим шагом является разработка стратегии управления данными. Это включает:

- **Устав** – Определите драйверы, видение, миссию и принципы управления данными, включая оценку готовности, выявление внутренних процессов, а также текущих проблем или критериев успеха;
- **Оперативные руководящие положения и система подотчетности** – Определите структуры и обязанности для управления данными;
- **«Дорожная карта» реализации** – Определите временные рамки для осуществления политики и директив, бизнес-гlossария, архитектуры, оценки активов, стандартов и процедур, ожидаемых изменений в бизнес-процессах и технологических процессах, а также результатов для поддержки аудиторской деятельности и соответствия нормативным требованиям;
- **План успешного практического осуществления** – Опишите конечное состояние устойчивой деятельности по управлению данными.³⁰²

Наконец, правительствам следует рассмотреть возможность принятия принципов управления данными, например, таких, как разработанные Институтом управления данными, приведенные ниже:

- **Целостность** – Участники управления данными будут стараться придерживаться принципа целостности в отношениях друг с другом; они будут правдивы и откровенны при обсуждении факторов, ограничений, вариантов и воздействий при

<https://nascio.org/Portals/0/Publications/Documents/NASCIO-DataGovernancePTII.pdf>.

301 NASCIO, "Data Governance – Managing Information as an Enterprise Asset: Part I – An Introduction", April 2008, p. 5. Доступно по <https://www.nascio.org/Portals/0/Publications/Documents/NASCIO-DataGovernance-Part1.pdf>.

302 DAMA International, *Data Management Body of Knowledge*, second edition (DAMA-DMBOK 2) (New Jersey, Technics Publications, 2017), p. 82.

принятии решений, связанных с данными;

- **Прозрачность** – Управление данными и управленческие процессы будут демонстрировать прозрачность; всем участникам и проверяющим должно быть ясно, как и когда решения и средства управления, связанные с данными, применялись в процессах;
- **Возможность проверки** – Решения, связанные с данными, процессы и средства контроля, имеющие отношение к управлению данными, будут подлежать проверке; они будут сопровождаться процедурами документирования для обеспечения требований на соответствие и проведения аудита деятельности;
- **Подотчетность** – Управление данными будет определять ответственность за решения, процессы и средства управления на основе данных, затрагивающих все подразделения организации;
- **Выполнение работы** – Управление данными будет определять ответственность за действия по выполнению работы, которые являются обязанностями отдельных участников, а также ответственностью группы распорядителей данными;
- **Сдержки и противовесы** – Управление данными будет определять ответственность таким образом, чтобы поддерживать систему сдержек-и-противовесов между бизнесом-командой и технологическими группами, а также между теми, кто создает/собирает данные, теми, кто ими управляет, теми, кто их использует, и теми, кто вводит стандарты и требования соответствия;
- **Стандартизация** – Управление данными будет внедрять и поддерживать стандартизацию данных организации;
- **Управление изменениями** – Управление данными будет поддерживать упреждающие и ответные действия по управлению изменениями для справочных значений данных и структуры/использования базовых данных и метаданных.³⁰³

6.3 Повышение общественного доступа и доверия к данным

Действия, которые необходимо рассмотреть для улучшения доступа общественности и доверия к данным, включают открытие государственных данных и обеспечение конфиденциальности данных.

6.3.1 Открытие государственных данных

Открытые государственные данные определяются как: «Данные, произведенные или введенные в действие правительством или организациями, контролируемые государством, которые могут свободно и повторно использоваться и распространяться любым лицом».³⁰⁴ Это также относится к: «Философии – все большее число политик,

303 The Data Governance Institute, “Goals and Principles for Data Governance”. Доступно по http://www.datagovernance.com/adg_data_governance_goals/ (accessed on 8 January 2019).

304 Open Knowledge Foundation, “What is Open Government Data”. Доступно по <https://opengovernmentdata.org/>

которые способствуют прозрачности, подотчетности и созиданию, делая государственные данные доступными для всех».³⁰⁵

Восемь принципов открытых государственных данных:

1. **Полные** – Все государственные данные становятся доступными. Государственные данные – это данные, на которые не распространяются действующие ограничения по вопросам конфиденциальности, безопасности или наличия привилегий;
2. **Первичные** – Данные собираются в источнике с максимально возможным уровнем детализации, а не в агрегированной или измененной форме;
3. **Своевременные** – Данные становятся доступными настолько быстро, насколько это необходимо для сохранения ценности данных;
4. **Доступные** – Данные доступны широкому кругу пользователей для самых разных целей. Данные должны быть доступны в Интернете, чтобы охватить самый широкий круг практического применения и пользователей;
5. **Обрабатываемые компьютером** – Данные достаточно структурированы для автоматической обработки;
6. **Недискриминационные** – Данные доступны для всех, без необходимости регистрации. Для государственных данных должен быть разрешен анонимный доступ к данным, включая доступ через анонимные прокси. Данные не должны быть заключены в «золотую клетку»;
7. **Непатентованные** – Данные доступны в формате, который не должен предоставлять кому-либо исключительные права контроля. Патентованные форматы добавляют ненужные ограничения на то, кто может использовать данные, как их можно использовать и передавать, и будут ли данные использоваться в будущем;
8. **Безлицензионные** – На данные не распространяются никакие положения об авторском праве, патенте, торговой марке или коммерческой тайне. Разумные ограничения по вопросам конфиденциальности, безопасности и наличия привилегий могут быть разрешены.³⁰⁶

Открытые данные способствуют: прозрачности и демократическому контролю; участию; самоутверждению; появлению усовершенствованных или новых продуктов и услуг; инновациям; внедрению более эффективных и действенных государственных услуг; оценке воздействия политики; и получению новых знаний из объединенных источников данных и моделей в больших объемах данных.³⁰⁷

Три составных элемента открытых данных для достижения перспектив развития: (1) публикация открытых данных правительствами; (2) преобразование данных в полезную информацию организациями-посредниками; и (3) использование данных гражданами,

(accessed on 8 January 2019).

305 OECD, “Open Government Data”. Доступно по <http://www.oecd.org/gov/digital-government/open-government-data.htm> (accessed on 8 January 2019).

306 Joshua Tauberer, “The Annotated 8 Principles of Open Government Data”. Доступно по <https://opengovdata.org/> (accessed on 8 January 2019).

307 Open Knowledge International, “Why Open Data?” *Open Data Handbook*. Доступно по <http://opendatahandbook.org/guide/en/why-open-data/> (accessed on 8 January 2019).

государственными должностными лицами и другими заинтересованными сторонами для достижения результатов развития.³⁰⁸

Фонд Sunlight предоставило правительствам полезное руководство по разработке и реализации политики открытых данных.³⁰⁹

Первая часть руководства посвящена тому, *какие данные должны быть общедоступными*. Рекомендации следующие:

- **Активно публиковать государственные данные в интернете** – Инициативное раскрытие информации – это публикация государственных данных до того, как лицо их запросит. В 21-м веке это означает активное размещение новых данных в Интернете.
- **Ссылаться и опираться на существующие государственные политики подотчетности и доступа** – Сильная политика открытых данных основывается на принципах, воплощенных в существующих законах и политиках, которые защищают и устанавливают уровень доступа населения, часто определяя стандарты качества данных, их раскрытия и публикации.
- **Опираться на ценности, цели и миссию общества и правительства** – Политика открытых данных может проводиться с целью достижения многих различных общественных благ, в том числе большей прозрачности правительства, честности, подотчетности, эффективности, гражданской активности и экономического роста. Четкое изложение целей, ценностей или намерений может помочь уточнить результаты, которые правительство надеется достичь с помощью политики открытых данных.
- **Создать общедоступный и исчерпывающий список всех имеющихся данных** – Правительствам следует провести инвентаризацию существующих данных на ранних этапах разработки политики открытых данных, чтобы правительство и другие заинтересованные стороны знали обо всех возможных последствиях обнародования данных.
- **Указать методику определения приоритетов обнародования данных** – В то время как политика открытых данных в идеале обеспечивает возможность онлайн-публикации всех государственных данных, обнародование данных может в конечном итоге стать поэтапным процессом по практическим причинам, таким как недостаточное финансирование или кадровое обеспечение. Правительствам следует четко определить комплекс возможных методик, которые можно использовать при определении очередности обнародования данных.
- **Оговорить условия обращения с открытыми данными подрядчиками и полугосударственными организациями** – Данные, которые собираются от общественности с использованием государственных средств, должны оставаться общедоступными, независимо от решения правительства о передаче полномочий

308 Andreas Pawelke and others, *Data for development: What's next? – Concepts, trends and recommendations for German development cooperation* (Bonn and Eschborn, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, 2017), p. 29. Доступно по http://webfoundation.org/docs/2017/12/Final_Data-for-development_Whats-next_Studie_EN-1.pdf.

309 Sunlight Foundation, "Open Data Policy Guidelines", *Open Data Policy Hub*. Доступно по <https://opendatapolicyhub.sunlightfoundation.com/guidelines/> (accessed on 8 January 2019).

на управление.

- **Соответствующим образом защитить конфиденциальные данные** – Хорошо продуманная политика открытых данных дополняет ранее существовавшее законодательство и директивы о доступе к государственным данным – это означает, что она может интегрировать ранее существовавшие исключения из закона о публичном доступе для данных, которые конфиденциальны с точки зрения безопасности или по другим причинам.³¹⁰

Вторая часть о том, *как сделать данные общедоступными*. Рекомендации следующие:

- **Использовать форматы данных для максимального технического доступа** – Данные должны быть обнародованы в форматах, которые можно легко и эффективно использовать с помощью технологий. Это означает обнародование данных в открытых форматах (или открытых стандартах) и в машиночитаемых форматах, которые соответствующим образом структурированы (или обрабатываются машиной).
- **Обеспечить всесторонние и подходящие форматы для различного использования** – Должны быть рассмотрены соответствующие методы распространения, чтобы максимизировать степень доступа, использования и качества публикуемых данных.
- **Снять ограничения на доступ к данным** – Для обеспечения действительно открытого доступа необходимо предоставить право на повторное использование государственных данных без каких-либо технических ограничений, таких как требования по регистрации, платы за доступ и ограничения использования и т.д.
- **Данные не должны быть обременены лицензиями** – Если данные должны быть действительно общедоступными и максимально пригодными для повторного использования, не должно быть никакого препятствия, связанного с лицензиями, для повторного использования государственных данных.
- **Обязать ведомства, создающие данные, создать соответствующую форму цитирования** – Практику цитирования государственных данных необходимо поощрять, если непосредственные распорядители данных разработают типовые ссылки для своих наборов данных. Эти типовые ссылки должны перечислять ключевые элементы идентичности источника, которые потребуются для эффективной идентификации отдельного источника данных и определения подразделения правительства, которая создала или хранит данные.
- **Требовать публикации метаданных** – Предоставление общей и полностью описанной базовой схемы метаданных (а также другой документации) может быть полезным как для общественности, так и для правительства. Качественная схема метаданных берет свое начало от общих международных мета-атрибутов (таких как DCAT) и позволяет издателям данных классифицировать контекстные поля или элементы в своих наборах данных.
- **Требовать публикации процессов создания данных** – Краткое изложение процессов, которые использовались для создания определенного набора данных, предоставляет ценную информацию, которая может быть не увидена только с помощью метаданных и должна сопровождать процесс опубликования массива данных.

- **Обязать использование уникальных идентификаторов** – Использование уникальных идентификаторов в наборах данных и между ними повышает качество и точность анализа данных. Без уникальных идентификаторов некоторые случаи анализа могут стать трудоемкими или невозможными, поскольку похожие имена могут относиться или не относиться к одним и тем же объектам. Важно отметить, что идентификаторы должны быть непатентованными и публичными.
- **Требовать совместного использования кода или публикации в качестве открытого источника** – В дополнение к данным код, используемый для создания государственных веб-сайтов, порталов, инструментария и других онлайн-ресурсов, может быть полезным в качестве открытых данных.
- **Требовать проведения оцифровки и распространения архивных материалов** – Вопросы о том, какие архивные материалы должны быть оцифрованы и какие сроки являются реалистичными для оцифровки архивных материалов, могут быть получены с помощью того же самого процесса расстановки приоритетов, который используется для общей процедуры обнародования данных.
- **Создать единое место, посвященное публикации данных и политики** – Порталы данных и аналогичные веб-сайты могут способствовать распространению открытых данных, предоставляя удобный для поиска хаб с возможностью поиска для нескольких массивов данных.
- **Публиковать массивы данных** – Широкомасштабный доступ обеспечивает простой, но эффективный способ полной публикации наборов данных, позволяя общественности сразу загрузить все данные, хранящиеся в базе данных.
- **Создать публичные API для доступа к данным** – Государственные органы могут разрабатывать API-интерфейсы, которые позволяют сторонним субъектам автоматически искать, извлекать или отправлять данные непосредственно из баз данных в режиме онлайн.
- **Оптимизировать методы сбора данных** – Для оптимизации качества и своевременности данных в правилах раскрытия информации должны использоваться методы сбора данных в режиме онлайн. Электронная подача заявок, также известная как «э-подача», является одним из методов оптимизации качества и своевременности сбора данных.
- **Обязать выполнение обнародования и обновления данных на регулярной основе** – Идеал онлайн-данных – это доступные в режиме реального времени – данные должны быть доступны как можно ближе к моменту их сбора. Недостаточно требовать одноразового обнародования массива данных, поскольку он становится неполным, как только дополнительные данные создаются, но не публикуются.
- **Создать постоянный, долговременный доступ к данным** – После обнародования оцифрованные государственные данные должны оставаться постоянно доступными, которые можно «найти» в постоянном онлайн-месте или в архивах с неограниченным сроком хранения.³¹¹

Третья часть – как реализовать политику открытых данных. Рекомендации следующие:

- **Создать или назначить надзорный орган** – Некоторые вопросы могут не решаться легко в процессе создания политики открытых данных, поэтому целесообразно определить единый орган, уполномоченный разрешать конфликты и обеспечивать соответствие новым мерам по открытым данным.
- **Создать руководство или другие обязательные правила для реализации** – Политика открытых данных должна быть на практике «желаемой» – это означает, что она должна определять видение того, почему политика реализуется, и также должна предоставить практические шаги для правительства и надзорных органов, которым необходимо следовать, чтобы довести политику до реализации. Создание правил или руководств может обеспечить сильную, надежную политику и обычно означает разницу между политикой, принятой для шоу, и политикой, принятой по существу.
- **Учесть точку зрения общественных интересов при реализации политики** – Общественность должна быть вовлечена в текущую оценку и анализ реализации политики. Правительствам следует создавать реальные возможности для организации обратной связи с общественностью для получения информации о качестве, количестве, выборе и формате данных, а также удобстве ресурсов доступа для пользователя.
- **Установить соответствующие реалистичные сроки реализации** – Установление четких сроков может продемонстрировать целеустремленность обязательства и поможет преобразовать обязательства в результаты. Сроки также могут помочь четко выявить невыполнения, открывая возможности для общественного контроля.
- **Создать процессы для обеспечения качества данных** – Качество данных не будет обеспечено только путем обнародования данных. Необходимо приложить усилия, чтобы поддерживать актуальность, точность и доступность данных.
- **Обеспечить достаточное финансирование для реализации** – Как и любая другая инициатива, реализация политики открытых данных должна осуществляться с учетом долгосрочной устойчивости. Одним из способов добиться этого является рассмотрение источников финансирования для реализации политики, а также ее будущая поддержка. Достаточное финансирование может означать разницу между успешной и неудачной политикой.
- **Создать или изучить возможности партнерства** – Партнерство может быть полезным в различных важных мероприятиях, связанных с обнародованием данных, таких как: повышение доступности открытых данных, определение неотъемлемых приоритетов для обнародования данных и подключение государственных данных к тем, которые хранятся в некоммерческих организациях, аналитических центрах, академических учреждениях и соседних правительствах.
- **Обязать проведение будущего пересмотра для потенциальных изменений в этой политике** – Для того, чтобы не отставать от современного передового опыта и осуществлять обратную реакцию на результаты контроля за существующей политикой, политика открытых данных должна предусматривать проведение в будущем пересмотра самой политики, а также любых рекомендаций, созданных политикой или другими процессами реализации.³¹²

Организация Open Data Barometer рекомендует следующие ключевые шаги «для

гарантирования того, что «революция данных» приведет к подлинной революции в прозрачности и эффективности правительств»:

- Политическая приверженность высокого уровня активному раскрытию данных государственного сектора, особенно данных, наиболее важных для подотчетности;
- Непрерывные инвестиции в поддержку и обучение широкого круга гражданского общества и предпринимателей для эффективного понимания и использования данных;
- Контекстуализация инструментов открытых данных и подходов к местным потребностям, например, делая данные визуально доступными в странах с более низким уровнем грамотности;
- Поддержка инициатив по открытым данным на уровне города в качестве дополнения к программам на национальном уровне;
- Правовая реформа, которая содержит гарантии права на информацию и права на неприкосновенность частной жизни, лежащие в основе инициатив в области открытых данных.³¹³

Несмотря на публичные обязательства правительств на высоком уровне открыть данные и принять ряд многосторонних глобальных инициатив, переход на открытые данные происходил медленно. В отчете *Open Data Inventory 2017* года – оценка охвата и открытости официальной статистики, опубликованной на веб-сайтах национальных статистических управлений – отмечено, что «прогресс до настоящего времени был медленным».³¹⁴

6.3.2 Укрепление конфиденциальности данных

Слежение – или «целенаправленное, рутинное, систематическое и сфокусированное внимание, уделяемое личным данным, для контроля, предоставления прав, управления, влияния или защиты» – стало нормой.³¹⁵

Исследование, проведенное в августе 2018 года, показало, что, не извещая владельцев смартфонов, Google отслеживают местоположение устройств Android даже тогда, когда телефон не работает. За период в 24 часа устройство Android отправляет в Google около 4,4 МБ, а iPhone - 0,76 МБ. Это же исследование показало, что «айфоны отправляют данные в 10 раз реже на серверы Apple, чем устройства Android на серверы Google. Apple также собирает данные о местоположении, но в среднем только один раз в день».³¹⁶

В государственном секторе «правоохранительные органы и другие правительственные

313 Web Foundation, “Key Findings”, *Open Data Barometer*. Доступно по <https://opendatabarometer.org/2ndEdition/summary/index.html> (accessed on 8 January 2019).

314 Open Data Watch, “Open Data Inventory 2017 Annual Report: A Progress Report on Open Data”, 2018. Доступно по <https://opendatawatch.com/publications/open-data-inventory-2017-annual-report/>.

315 David Lyon, “Surveillance Society”, presentation made at Piacenza, Italy, 28 September 2008. Доступно по http://www.festivaldeldiritto.it/2008/pdf/interventi/david_lyon.pdf.

316 Liam Tung, “Want Google to track you less? Get an iPhone, ditch the Android”, *ZDNet*, 23 August 2018. Доступно по <https://www.zdnet.com/article/want-google-to-track-you-less-get-an-iphone-ditch-the-android/>.

учреждения считают, что Facebook, Twitter, Instagram и другие сайты социальных сетей являются богатыми источниками данных, чтобы покопаться для различных целей».³¹⁷

В наше время проблема конфиденциальности данных заключается в следующем: «Определить, когда и как с этической точки зрения анализировать информацию, что искать в данных, какие вопросы задавать для данных и масштаб, в котором разумно делать прогнозы о будущих событиях и действиях на основе этих данных».³¹⁸

С 1970-х годов законодательство о конфиденциальности данных стало традиционным ответом на растущее слежение. Конфиденциальность данных (или конфиденциальность информации) – «это право иметь некоторый контроль над тем, как ваша личная информация собирается и используется».³¹⁹

Первый закон о конфиденциальности – Закон о защите данных – вступил в силу в немецком штате Гессен в 1971 году.³²⁰ Шведский закон о данных 1973 года считается первым национальным законом о защите данных. Конгресс США принял Закон о конфиденциальности в 1974 году, который регулирует сбор и использование записей федеральными агентствами и дает людям право на доступ и исправление их личной информации.³²¹ Франция, пионер в области законодательства о конфиденциальности, приняла Закон о защите данных в 1978 году (в который были внесены поправки в 2005 году).³²²

В Азии Япония сформулировала свой закон о защите данных, обрабатываемых компьютером, который был принят административными органами в 1988 году.³²³ В 1995 году Тайвань принял Закон о защите персональных данных, обрабатываемых компьютером. Первый закон Республики Корея о защите данных (Закон о защите данных государственных ведомств) был принят в 1995 году. В том же году колониальное правительство Гонконга приняло Постановление о конфиденциальности личных данных, которое считается первым в Азии всеобъемлющим законом о конфиденциальности данных. Закон об официальной информации Таиланда был принят в 1997 году. Филиппины приняли закон о конфиденциальности данных в 2012 году.

На региональном уровне Организация экономического сотрудничества и развития (ОЭСР) опубликовала свои принципы защиты данных в 1980 году, которые заключаются в следующем:

317 Rachel Levinson-Waldman, "Government Access to and Manipulation of Social Media: Legal and Policy Challenges", *Howard Law Journal*, vol. 61, no. 3 (August 2018), p. 560. Доступно по <https://www.brennancenter.org/analysis/government-monitoring-social-media-legal-and-policy-challenges>.

318 Jens-Erik Mai, "Big data privacy: The datafication of personal information", *The Information Society*, vol. 32, no. 3 (2016), p. 194. Доступно по <http://download.xuebalib.com/xuebalib.com.22693.pdf>.

319 International Association of Privacy Professionals, "What does privacy mean?" Доступно по <https://iapp.org/about/what-is-privacy/> (accessed on 8 January 2019).

320 Jan Holvast, "History of Privacy", in *The Future of Identity*, V. Matyáš and others, eds. (International Federation for Information Processing, 2009), p. 28. Доступно по https://link.springer.com/content/pdf/10.1007/978-3-642-03315-5_2.pdf.

321 Daniel J. Solove, "A Brief History of Information Privacy Law", *GW Law*, 2006. Доступно по https://scholarship.law.gwu.edu/cgi/viewcontent.cgi?article=2076&context=faculty_publications.

322 Deutsche Welle, "France maintains long tradition of data protection", 26 January 2011. Доступно по <http://www.dw.com/en/france-maintains-long-tradition-of-data-protection/a-14797711>.

323 Graham Greenleaf, "A Brief History of Data Privacy Laws in Asia", *Tiki-Toki*. Доступно по <http://www.tiki-toki.com/timeline/entry/381411/A-Brief-History-of-Data-Privacy-Laws-in-Asia/> (accessed on 8 January 2019).

- **Принцип ограничения сбора** – Должны быть ограничения на сбор персональных данных, и любые данные должны быть получены законными и справедливыми способами и, при необходимости, с ведома или согласия субъекта данных.
- **Принцип качества данных** – Персональные данные должны соответствовать целям, для которых они будут использоваться, и на том необходимом уровне, должны быть точными, полными и постоянно обновляться.
- **Принцип спецификации цели** – Цели, для которых собираются персональные данные, должны быть указаны не позднее, чем во время сбора данных, и последующее использование ограничено выполнением тех или иных целей, которые не являются несовместимыми с этими целями и которые указываются при каждом изменении цели.
- **Принцип ограничения использования** – Персональные данные не должны разглашаться, предоставляться или иным образом использоваться для целей, отличных от указанных, кроме как с согласия субъекта данных или в соответствии с законом.
- **Принцип обеспечения безопасности** – Персональные данные должны быть защищены обоснованными мерами безопасности от таких рисков, как потеря или несанкционированный доступ, уничтожение, использование, изменение или раскрытие данных.
- **Принцип открытости** – Должна быть общая политика открытости в отношении событий, практики и политики в отношении персональных данных. Средства должны быть легко доступны для установления существования и характера персональных данных, основных целей их использования, а также личности и места жительства распорядителя данных.
- **Принцип индивидуального участия** – Человек должен иметь право на:
 - a. Получение от распорядителя данных или иным образом подтверждение того, есть ли у распорядителя данные, относящиеся к нему/ней;
 - b. Сообщение ему/ей о данных, относящихся к нему/ней; (ii) за плату, если таковая имеется, которая не является чрезмерной; (iii) на приемлемых условиях; и (iv) в форме, которая легко понятна ему/ей;
 - c. Указание причин, если запрос отклонен, и возможность оспорить такой отказ;
 - d. Возможность оспорить данные, относящиеся к нему/ней, и, если апелляция прошла успешно, удалить, исправить, дополнить или изменить данные.
- **Принцип ответственности** – Распорядитель данных должен нести ответственность за соблюдение мер, которые реализуют принципы, изложенные выше.³²⁴

Данный документ и его принципы оказали влияние на разработку национальных законодательств и типовых кодексов не только ОЭСР, но и других стран.

В 2005 году Азиатско-Тихоокеанское экономическое сотрудничество (АТЭС) приняло Концепцию конфиденциальности АТЭС, которая «соответствует основным ценностям Руководства ОЭСР по защите конфиденциальности и трансграничных потоков»

персональных данных и подтверждает важность обеспечения конфиденциальности для отдельных лиц и информационного общества».³²⁵ Десять лет спустя АТЭС выпустил обновленную структуру, основанную на концепциях, представленных в Руководстве ОЭСР по конфиденциальности 2013 года, с должным вниманием различным правовым особенностям и контексту региона АТЭС.³²⁶

Наиболее важным недавним событием в области конфиденциальности данных является вступление в силу Общего регламента о защите персональных данных (GDPR) – новой нормативно-правовой базы, которая объединяет законы о защите данных в ЕС.

GDPR имеет семь ключевых принципов:

1. Законность, справедливость и прозрачность;
2. Конкретизация цели;
3. Минимизация данных;
4. Точность;
5. Ограничение хранения данных;
6. Целостность и конфиденциальность (безопасность);
7. Подотчетность.³²⁷

Среди его ключевых особенностей, которые характерны для организаций, можно перечислить следующее:

- Назначение сотрудника по защите данных;
- Обязательное проведение оценки последствий для неприкосновенности частной жизни;
- Расширенные права субъектов данных;
- Уведомление об утечке данных в течение 72 часов;
- Новые обязательства по соблюдению.³²⁸

Для государств, не входящих в ЕС, GDPR важен из-за его экстерриториальной применимости. GDPR применяется к:

- Всем компаниям, обрабатывающим персональные данные субъектов данных, проживающих в ЕС, независимо от местоположения компании;
- Обработке персональных данных субъектов данных в ЕС контролирующим или обрабатывающим субъектом, не установленным в ЕС, где эти виды деятельности связаны с предложением товаров или услуг гражданам ЕС (независимо от того, требуется ли оплата) и мониторингом поведения, которое происходит в рамках ЕС.³²⁹

325 АПЕС, *APES Privacy Framework* (Singapore, 2015). Доступно по [https://www.apec.org/Publications/2017/08/APES-Privacy-Framework-\(2015\)](https://www.apec.org/Publications/2017/08/APES-Privacy-Framework-(2015)).

326 АПЕС, *APES Privacy Framework* (Singapore, 2015). Доступно по [https://www.apec.org/Publications/2017/08/APES-Privacy-Framework-\(2015\)](https://www.apec.org/Publications/2017/08/APES-Privacy-Framework-(2015)).

327 Information Commissioner's Office, United Kingdom, "The Principles". Доступно по <https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/principles/> (accessed on 8 January 2019).

328 Aditya Vats, "10 Key Issues Of General Data Protection Regulation (GDPR)", *Medium*, 25 May 2017. Доступно по <https://medium.com/@adityavats/10-key-issues-of-general-data-protection-regulation-gdpr-d70e3875b59e>.

329 EUGDPR.org, "GDPR Key Changes". Доступно по <https://www.eugdpr.org/key-changes.html> (accessed on 8

Кроме того, компании за пределами ЕС, обрабатывающие данные граждан ЕС, должны будут назначить представителя в ЕС.³³⁰

Недавняя и интересная тенденция развития – использование этики в защите конфиденциальности данных. Этические нормы, впервые разработанные Супервизором по конфиденциальности данных ЕС, а затем Гонконгом и Филиппинами, является важным дополнением к юридическим обязательствам по достижению целей обеспечения конфиденциальности.³³¹ Кроме того, этика рассматривается как руководство к поведению в среде, где закон не в состоянии идти в ногу с технологическим развитием. При таком подходе соблюдение закона является просто основой (базой) для инициатив по защите данных.

Но адекватно ли защищают нас законы о конфиденциальности данных и этические нормы в эпоху больших данных?

Нобелевский лауреат Джозеф Стиглиц считает, что существует необходимость принятия правил для следующего: «Какие данные могут хранить технологические фирмы; какие данные они могут использовать; могут ли они объединять разные наборы данных; цели, для которых они могут использовать эти данные; и какую степень прозрачности они должны предоставить при вопросе о том, что они делают с данными».³³²

Дженс-Эрик Май утверждает, что существующие модели конфиденциальности (и в соответствии с законодательством о расширении сферы деятельности) «разделяют общую направленность на сбор данных и ... заботятся о том, как и для каких целей собранные данные впоследствии используются».³³³ Он считает, что в эпоху обработки данных (датафикации) и больших данных нам необходимо сместить акцент с сбора данных на обработку и анализ данных. Это потому, что «обеспечение неприкосновенности частной жизни в действии – это создание новых знаний, идей или реалий на основе доступных данных».³³⁴ Май предлагает нам объединить существующие подходы к конфиденциальности с «датафикационной» моделью обеспечения неприкосновенности частной жизни, которая «фокусируется на анонимном создании новой личной информации, реинтерпретации и статистическом анализе данных, а также на коммерциализированном характере личной информации».

Как это отражено в законодательстве или нормативных актах – это то, что все еще

January 2019).

330 Там же

331 European Data Privacy Supervisor, "Towards a new digital ethics: Data, dignity and technology", Opinion 4, 2015, p. 6. Доступно по https://edps.europa.eu/sites/edp/files/publication/15-09-11_data_ethics_en.pdf. For Hong Kong, see Information Accountability Foundation, "Ethical Accountability Framework for Hong Kong, China: A Report prepared for the Office of the Privacy Commissioner for Personal Data", n.d. Доступно по https://www.pcpd.org.hk/misc/files/Ethical_Accountability_Framework.pdf. For the Philippines, see National Privacy Commission, Philippines, "NPC launches DPO ACE Program, sets benchmark for data privacy training in PH", 12 December 2018. Доступно по <https://www.privacy.gov.ph/2018/12/npc-launches-dpo-ace-program-sets-benchmark-for-data-privacy-training-in-ph/>.

332 Ian Sample, "Joseph Stiglitz on artificial intelligence: 'We're going towards a more divided society'", *The Guardian*, 8 September 2018. Доступно по <https://www.theguardian.com/technology/2018/sep/08/joseph-stiglitz-on-artificial-intelligence-were-going-towards-a-more-divided-society>.

333 Jens-Erik Mai, "Big data privacy: The datafication of personal information", *The Information Society*, vol. 32, no. 3 (2016), p. 198. Доступно по <http://download.xuebalib.com/xuebalib.com.22693.pdf>.

334 Там же

нуждается в доработке.

6.4 Готовимся к большим данным

Потенциал больших данных в предоставлении своевременной и актуальной статистики для решения проблем развития хорошо известен.³³⁵ Тем не менее, ряд вопросов, касающихся методологии и технологии, законодательства, конфиденциальности, управления и финансов, еще предстоит решить. Существует также проблема создания «среды, в которой общественность доверяет использованию больших данных для официальной статистики и где может быть обеспечена неприкосновенности частной жизни и конфиденциальность личной информации».³³⁶

Создание благоприятной среды для больших данных требует лидерства.

Бен Росси считает, что «государственные лидеры должны взять на себя обязательство возглавить (инициативы в области больших данных) на всех уровнях и во всех областях».³³⁷ Росси добавляет, что для эффективного лидерства руководители должны обладать «фундаментальным знанием» обработки данных, аналитики и алгоритмов. Компетентность может повлиять на улучшении жизни граждан инициативами в области больших данных, или эти же граждане могут стать жертвами данных, аналитики и алгоритмов.

Помимо руководства, британское правительство выдвигает на первый план две основные проблемы, которые необходимо преодолеть, чтобы успешно реализовывать инициативы в области больших данных: (1) завоевание и сохранение общественного доверия; и (2) наращивание потенциала государственной службы по сбору, хранению, анализу, обмену и использованию данных.³³⁸

Для британского правительства: «Общественное доверие абсолютно необходимо для достижения наших амбиций в отношении правительства, управляемого данными».³³⁹ Заслужить общественное доверие значит: (1) персональные данные, собранные и хранящиеся в ведомстве, используются надлежащим образом и эффективно; и (2) персональные данные в безопасности, особенно когда они передаются между различными ведомствами. Конкретной мерой, применяемой для завоевания общественного доверия, является структура Управления национальной статистики, которая называется «Five safes», для создания и поддержания доверия и уверенности в использовании государственных данных.³⁴⁰

335 United Nations Global Working Group for Big Data, “Using Big Data for the Sustainable Development Goals”. Доступно по <https://unstats.un.org/bigdata/taskteams/sdgs/> (accessed on 8 January 2019).

336 Там же

337 Ben Rossi, “Why the Governments’ Data Science Ethical Framework is a Recipe for Disaster”, *Information Age*, 2 June 2016. Доступно по <https://www.information-age.com/why-governments-data-science-ethical-framework-recipe-disaster-123461541/>.

338 John Manzoni, “Big data in government: The challenges and opportunities”, *Gov.UK*, 21 February 2017. Доступно по <https://www.gov.uk/government/speeches/big-data-in-government-the-challenges-and-opportunities>.

339 Там же

340 Peter Stokes, “The ‘Five Safes’ – Data Privacy at ONS”, *Office for National Statistics*, United Kingdom, 27 January 2017. Доступно по <https://blog.ons.gov.uk/2017/01/27/the-five-safes-data-privacy-at-ons/>. Это инициатива по защите

Что касается второй проблемы – наращивания потенциала – британское правительство запустило программы, направленные на развитие сообщества специалистов в области данных, а также на повышение уровня грамотности по данным для специалистов, не связанных с данными, на государственной службе. В связи с этим была создана программа «Data Science Campus» в Управлении национальной статистики Великобритании. Целью Программы является: «Создание нового поколения инструментов и технологий для использования повышения и доступности инновационных источников данных, а также для обеспечения всестороннего обоснованного измерения и анализа экономики, глобальной окружающей среды и общества в целом».³⁴¹ «Data Science Campus», созданная в 2017 году, ставит своей целью подготовку «500 квалифицированных аналитиков данных для правительства в ближайшие годы».³⁴² Для достижения этой цели инициированы следующие мероприятия: обучение в области анализа данных; магистратура в области анализа данных для правительства; модули непрерывного профессионального развития; совместное финансирование и/или совместное обучение и программы MSc и PhD; поддержка и финансирование студентов магистратуры; серия семинаров по науке о данных; тренинг «Искусство возможного»; акселератор данных и наставничество Национальной академии данных Управления национальной статистики; и представители, популяризирующие STEM-образование.³⁴³

Правительства, стремящиеся использовать мощь больших данных, должны решить также операционные проблемы.

Исследователи из Университета Манчестера разработали «модель, показывающую разницу между реальностью и дизайном для больших данных, используемых в целях развития». Чтобы понять препятствия для использования больших данных в целях развития, модель измеряет «разрыв между требованиями или допущениями к проектированию больших данных по сравнению с текущей реальностью на местах».³⁴⁴ Она рассматривает семь аспектов:

1. **Информация** – Включает в себя как информацию, так и данные, предшествующие созданию информации;
2. **Технологии** – Основное внимание уделяется ИКТ, которые обрабатывают данные;
3. **Процессы** – Действия, предпринятые для производства, сбора, анализа, представления и использования данных;
4. **Цели и ценности** – Компонент «цели» охватывает вопросы достижения собственных интересов и реализации политики, а также неформальные и формальные стратегии; и компонент «ценности» охватывает культуру (например, что заинтересованные стороны считают правильным и неправильным способом

частной информации, собираемой правительством, может противоречить политике открытых данных.

341 Office for National Statistics, United Kingdom, “Data Science Campus: Data science for public good”. Доступно по <https://www.ons.gov.uk/aboutus/whatwedo/datasciencecampus> (accessed on 8 January 2019).

342 Data Science Campus, “Building Capability”. Доступно по <https://datasciencecampus.ons.gov.uk/capability> (accessed on 8 January 2019).

343 Там же. STEM stands for science, technology, engineering and mathematics.

344 Richard Heeks, “Measuring Barriers to Big Data for Development”, *ICTs for Development*, 9 August 2016. Доступно по <https://ict4dblog.wordpress.com/2016/08/09/measuring-barriers-to-big-data-for-development/>.

сделать что-то);

5. **Навыки и знания** – Охватывает качественные и количественные аспекты человеческих компетенций для выполнения процессов, связанных с данными;
6. **Системы и структуры управления** – Более развернутые системы управления, необходимые для взаимодействия внутри и между организациями и сетями, связанными с данными, а также структурная организация системы как формально, так и неформально;
7. **Другие ресурсы** – Время и деньги.

Используя модель для анализа барьеров на пути использования больших данных в целях развития в Колумбии, исследователи из Манчестерского университета отметили, что существуют серьезные пробелы во всех аспектах.

Эти исследователи разработали рекомендации для Колумбии, которые могут быть полезны для других развивающихся стран, стремящихся запустить проекты больших данных для развития, и они включают в себя следующее:

1. **Информация** – Активнее использовать существующие большие массивы данных, например, доступные через социальные сети и подробные записи о мобильных вызовах, а также «темные данные», существующие в государственном секторе;
2. **Технологии** – Выполнять поэтапные инвестиции в системы для цифрового сбора данных из существующих источников;
3. **Процессы** – Уделить первоочередное внимание принятию мер, связанных с созданием ценности, а не процессов генерации данных;
4. **Цели и ценности** – Принять или изменить законодательство, чтобы именно решать проблемы больших данных, такие как конфиденциальность данных, права интеллектуальной собственности и коммерциализация;
5. **Навыки и знания** – Включить больших данных в учебные планы для государственных учебных заведений и программ на получение степени;
6. **Системы и структуры управления** – Определить активных сторонников и целевые группы по большим данным в каждом крупном государственном учреждении и включить активных сторонников в предлагаемый центр передового опыта, чтобы обеспечить возможность обмена опытом между ведомствами;
7. **Другие ресурсы** – Развивать возможности финансирования за счет государственно-частного партнерства для инициатив в области больших данных.

Также полезным является «Руководство по инновациям данных в целях развития: от идеи до подтверждения концепции», разработанное ПРООН и инициативой «Глобальный пульс» Организации Объединенных Наций, в котором содержится пошаговое руководство по использованию новых источников данных в целях развития.³⁴⁵

345 UNDP and United Nations Global Pulse, *A Guide to Data Innovation for Development: From Idea to Proof-Of-Concept*, version 1 (New York, 2016). Доступно по <http://www.undp.org/content/undp/en/home/librarypage/development-impact/a-guide-to-data-innovation-for-development---from-idea-to-proof-.html>.

7. Создание культуры данных в государственном секторе

В данном разделе рассматриваются вопросы формирования культуры данных на государственной службе

Культура данных, по словам Элизабет Данли, означает «повсеместное использование данных» в организации.³⁴⁶ В государственном секторе культура данных означает «уверенный уровень комфортности для всей организации по применению показателей для достижения максимального социального воздействия».³⁴⁷

Шесть принципов, которые лежат в основе здоровой культуры данных:

1. **Культура данных – это культура принятия решений** – Основной целью сбора, анализа и развертывания данных является принятие более эффективных решений.
2. **Культура данных и лидерство** – Обязательство со стороны высших должностных лиц имеет важное значение, но это обязательство должно проявляться не только в периодических заявлениях высокого уровня; должен быть постоянный, информированный разговор с высшими лицами, принимающими решения, и теми, кто руководит инициативами в области данных во всей организации.
3. **Демократизация данных** – Предоставьте данные людям, и они будут рады, но проведение крутых экспериментов или навязывание инструментов сверху вниз недостаточно. Для создания конкурентного преимущества необходимо стимулировать спрос на данные от населения.
4. **Культура данных и риск** – Действующая культура данных ставит в основу риск. Хотя организации должны знать свои ограничения и соблюдать их, управление рисками должно действовать как интеллектуальный ускоритель, вводя ответственно аналитику в ключевые процессы и взаимодействия.
5. **Катализаторы культуры** – Чтобы действительно обеспечить поступательное движение, кто-то должен взять на себя ответственность и приложить значительные усилия для достижения перемен. Требуются люди, которые могут соединить оба мира – науку о данных и выполнение действий на местах. И, как правило, наиболее эффективные агенты изменений не являются цифровыми по образованию.
6. **Сочетание таланта и культуры** – Это включает в себя установление соответствующего баланса для любого учреждения между введением новых сотрудников и преобразованием существующих. Необходим более широкий взгляд

346 Elizabeth Dunlea, “The Key to Establishing a Data-Driven Culture”, *Gartner*, 30 November 2015. Доступно по <https://www.gartner.com/smarterwithgartner/the-key-to-establishing-a-data-driven-culture/>.

347 Kathleen Kelly Janus, “Creating a Data Culture”, *Stanford Social Innovation Review*, 2 March 2018. Доступно по https://ssir.org/articles/entry/creating_a_data_culture.

на источники и более четкий взгляд на навыки, необходимые вашей команде, имеющей дело с данными.³⁴⁸

Чтобы создать культуру данных, важно знать препятствия, которые необходимо преодолеть.

Рахул Бхаргава определяет следующие ключевые препятствия для культуры данных:

- **Путаница** – Большинство представлений о данных сбивают с толку и чрезмерно технические. Цель состоит не в том, чтобы превратить всех в организации в ученых по обработке данных. Культура данных означает, что люди распознают данные и могут определить новые возможности для получения знаний и их понимания.
- **Не совсем знать свои данные** – Иногда вы даже не знаете данные, которые у вас есть. Трудно отслеживать массивы данных в организации, которые могут быть связаны друг с другом. Необходимо определить человека (офис) и технологию, которая может стать главным центром обмена данными.
- **Организационные разрозненные системы** – Большинство организаций страдают от разрозненных систем – независимых функциональных подразделений, которые стараются контролировать часть всей работы. Необходимо признать эти подразделения, чтобы устранить их.
- **Технологическое мышление** – Люди не должны обращаться в отдел ИКТ, чтобы получить последние цифры, которые им нужны.³⁴⁹

Для формирования культуры данных потребуется план, который включает: определение внутренних сторонников/экспертов; обнаружение передового опыта; выстраивание внешних связей, ведущих вверх и вниз; и делать небольшие («детские») шаги.³⁵⁰

Как и в других инициативах по трансформации, лидерство является ключевым фактором в создании культуры данных. Ниже приведены способы, которыми лидеры могут привести пример с данными:

- **Ежедневное использование** – Один из наиболее эффективных способов демонстрации о том, что данные имеют значение в вашей организации, заключается в том, что руководители активно используют эти данные.
- **Решения** – Если данные действительно важны для вашей организации, все решения ваших руководителей должны основываться на данных – без исключений. Когда руководители запрашивают данные для принятия ключевых решений, они усиливают роль данных в качестве важного стратегического актива, который служит неотъемлемой частью процесса принятия решений.

348 Alejandro Díaz, Kayvaun Rowshankish and Tamim Saleh, “Why data culture matters”, *McKinsey Quarterly*, September 2018. Доступно по <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/why-data-culture-matters>.

349 Rahul Bhargava, “You Don’t Need a Data Scientist, You Need a Data Culture”, *Data Therapy*, 6 December 2017. Доступно по <https://datatherapy.org/2017/12/06/building-a-data-culture/>.

350 Rahul Bhargava, “Architectures for Building a Data Culture”, *Data Therapy*, 20 July 2015. Доступно по <https://datatherapy.org/2015/07/20/architectures-for-building-a-data-culture/>.

- **Коммуникации** – Каждое электронное письмо, презентация или обсуждение встречи дает возможность поделиться идеями в отношении эффективности бизнеса, содействовать выигрышным ситуациям, основанным на данных, и подчеркнуть важность данных для организации.
- **Встречи** – Если большой упор делается на рассмотрение ключевых показателей и разработку планов действий на основе результатов, данные могут затем определять повестку дня собрания и делать встречи более целенаправленными, продуктивными и полезными для всех участников.
- **Обучение** – Где занятые руководители предпочитают проводить время, может показать, насколько важно для них что-то. Если руководители выделяют время для участия в обучении навыкам работы с данными, это дает наглядный пример их команде, что эти навыки будут иметь решающее значение для успеха организации.
- **Цифровые дисплеи** – Когда руководители выбирают отображение ключевых показателей в видных общедоступных местах с помощью цифровых дисплеев, они демонстрируют сотрудникам то, что показатели и цели играют важное значение в команде.³⁵¹

351 Brent Dykes, "Creating a Data-Driven Culture: Why Leading By Example is Essential", *Forbes*, 26 October 2017. Доступно по <https://www.forbes.com/sites/brentdykes/2017/10/26/creating-a-data-driven-culture-why-leading-by-example-is-essential/#58e6672b6737>.

ЗАМЕТКИ ДЛЯ ИНСТРУКТОРА

ЧАСТЬ 1

Советы по преподаванию модуля:

Ключевые моменты, упражнения и список литературы для каждого раздела

Раздел 1 Революция данных и справедливость данных

В этом вводном разделе раскрывается контекст для интенсивного и широкого использования данных обществами с целью определения и достижения своего общего будущего.

Ключевые моменты

1. Революция в области информационных и коммуникационных технологий (ИКТ) привела к:

- **Революция данных** – взрыв объема данных, скорость, с которой производятся данные, количество производителей данных, распространение данных и количество предметов, на которых существуют данные.
- **Переход от отдельных лиц в качестве «потребителей информации» к индивидуумам как «производителей информации».** Как производители информации, мы создаем цифровые следы – сумму всех данных, которые мы производим в результате нашей онлайн-деятельности. Это включает в себя данные, которые мы создаем, когда мы используем социальные сети (такие как Facebook и Twitter), а также когда мы ищем в Google. Мы также совершаем производство данных – побочный продукт нашей онлайн-деятельности. Наши произведенные данные предоставляют значительную информацию о нашем поведении в Интернете, и при обработке могут привести к ценной информации о наших привычках и предпочтениях.
- **Интернет вещей (IoT)** – Подключенные к Интернету устройства, которые создают новые данные без вмешательства человека. К 2025 году IoT будет создавать более 2 зетабайтов данных, в основном, генерируемых устройствами бытовой электроники.

2. Мы также являемся свидетелями эволюции от «оцифровки» к «цифровизации», «цифровому преобразованию» и «датафикации»:

- **Оцифровка** – процесс преобразования аналогового формата в цифровой;

- **Цифровизация** – превращение взаимодействий, коммуникаций, бизнес-функций и бизнес-моделей в (более) цифровые. Также интеграция цифровых технологий в повседневную жизнь;
- **Цифровая трансформация** – использование цифровых технологий во всех аспектах деятельности предприятия для фундаментального изменения способов создания и доставки добавленной стоимости;
- **Датафикация** - превращение ранее невидимых процессов/деятельности в данные, которые можно мониторить, отслеживать, анализировать и оптимизировать. Под «datafy» понимается превращение в данные многих аспектов окружающего мира, которые никогда не были определены количественно.

3. Обеспечение справедливости на основе данных является ключевой проблемой в мире, управляемом данными.

Сегодня почти половина населения мира все еще «невидима» – они не создают цифровые следы, потому что у них нет доступа к Интернету и другим цифровым технологиям. Некоторые среди тех, кто уже «виден», представлены только как потребители и клиенты, которыми не занимаются «ведомства».

Справедливость на основе данных означает такой принцип, при котором люди становятся видимыми, представленными и обслуживаемыми в результате производства своих цифровых данных.

Три принципа справедливости данных:

- **Видимость** - обеспечение конфиденциальности и представление тех, кто живет за пределами социально принятых норм;
- **Взаимодействие с технологиями** – (1) свобода от необходимости использовать конкретные цифровые технологии; (2) как не стать частью коммерческих баз данных (как побочный продукт активности в целях развития); и (3) свобода по контролю условиям взаимодействия кого-либо с рынками данных;
- **Недискриминация** - способность выявлять и оспаривать предвзятость в использовании данных, а также свобода от предвзятого обращения.

Справедливость на основе данных необходима для инклюзивного управления с помощью данных.

Предлагаемые упражнения

1. Оцифровка и цифровизация в государственном секторе

- Разделите участников на четыре группы.
- Две группы для обсуждения инициатив по оцифровке в своих ведомствах.

- Другие две группы для обсуждения инициатив по цифровизации в своих ведомствах.
- Каждая группа представит результаты своих обсуждений.

2. Принципы справедливости на основе данных

- Разделите участников на три группы.
- Группа 1 для обсуждения и приведения примера видимости из принципов справедливости на основе данных.
- Группа 2 для обсуждения и приведения примера взаимодействия с технологиями из принципов справедливости на основе данных.
- Группа 3 для обсуждения и приведения примера недискриминации из принципов справедливости на основе данных.

Литература для чтения

United Nations Secretary-General's Independent Expert Advisory Group on the Data Revolution for Sustainable Development, "A World that Counts: Mobilizing the Data Revolution for Sustainable Development", November 2014, p. 11. Доступно по <http://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf>.

Margaret Rouse, "Data Exhaust", WhatIs.com, April 2015. Available at <https://whatis.techtarget.com/definition/data-exhaust>.

I-Scoop, "Digitization, digitalization and digital transformation: the differences". Доступно по <https://www.i-scoop.eu/digitizationdigitalization-digital-transformation-disruption/> (accessed on 8 January 2019).

Margarita Shilova, "The Concept Of Datafication; Definition & Examples", Apiumhub, 15 June 2017. Доступно по <https://apiumhub.com/tech-blog-barcelona/datafication-examples/>.

Linnet Taylor, "What is data justice? The case for connecting digital rights and freedoms globally", 16 February 2017. Доступно по https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2918779.

Раздел 2 Данные и управление

Этот раздел начинается с обсуждения управления, а затем фокусируется на управлении на основе данных и администрировании на основе данных.

Ключевые моменты

1. Управление определяется как организация общества для определения и достижения своего общего будущего.

- Управление включает в себя «политику спроса» и «политику предложения».
- **Управление происходит на разных уровнях** - местном, национальном и глобальном – и имеет много действующих лиц, включая правительство, гражданское общество и бизнес-группы.
- **Управление тесно связано с правительством** – формальные структуры государственного сектора и группы субъектов, осуществляющих государственную власть, или материальное образование, специализирующееся на обеспечении управления. Правительство является наиболее важным субъектом управления.

2. Управление на основе данных – это интенсивное использование данных в деятельности общества для определения и достижения общего будущего.

- Правительство, осуществляющее управление на основе данных – это то место, где для всех критических решений доступна, когда и где это необходимо, действенная информация.
- В государственном секторе более широкое использование данных может привести к: более эффективному управлению и контролю; оптимизированному обнаружению мошенничества и ошибок, уменьшению и предотвращению; улучшенному предоставлению услуг, основанному на информации, полученной от тех, кто обслуживается; и повышению эффективности, что может привести к снижению затрат.

Различные формы управления на основе данных включают:

- **Принятие решений на основе данных** – Процесс, который включает сбор данных, извлечение закономерностей и фактов из этих данных, и использование этих фактов для вывода заключения, влияющее на принятие решений;
- **Разработка политики на основе фактов** - Использование наилучших доступных исследований и информации о результатах исполнения программ для принятия решений на всех этапах политического процесса и в каждом подразделении правительства;
- **Администрирование, ориентированное на результаты** - Стратегия управления, с помощью которой все участники, прямо или косвенно способствуя достижению ряда результатов, обеспечивают условия того, чтобы их процессы, продукты и услуги способствовали достижению желаемых результатов (результаты, итоги и цели или воздействие более высокого уровня).

3. Сложности в управлении на основе данных хорошо прослеживаются при мониторинге достижения Целей ООН в области устойчивого развития (ЦУР).

В целях обеспечения эффективной реализации Генеральная Ассамблея Организации Объединенных Наций приняла показатели ЦУР – надежную систему показателей и статистических данных для мониторинга прогресса, разработки политики и обеспечения подотчетности всех заинтересованных сторон. ЦУР имеют 232 показателя для измерения прогресса.

Показатели ЦУР разбиты на категории с использованием трехуровневой системы, основанной на методологии и данных:

- Уровень 1 – Существует установленная методология, и данные широко доступны;
- Уровень 2 – Существует установленная методология, но данные не всегда доступны;
- Уровень 3 – Согласованная на международном уровне методология еще не разработана.

По состоянию на 2017 год:

- Треть показателей имеют доступные данные, которые можно использовать для мониторинга ЦУР;
- Около четверти имеют методологию, но не имеют данных;
- Тридцать восемь процентов не имеют согласованной методологии и данных.

В Азии проблемы с данными, с которыми сталкиваются правительства, включают следующее:

- Необходимость укрепления национальных статистических систем для отчетности по ЦУР и поощрения нововведений в сборе данных для упрощения достижения поставленных задач;
- Необходимость для стран выбора показателей, которые являются наиболее значимыми для их приоритетов реализации (все 232 будут актуальны не во всех случаях);
- Устранение мониторинга, которое обусловлено наличием данных, а не их актуальностью для национальных приоритетов.

Предлагаемые упражнения

Оценка модели зрелости принятия решений на основе данных в правительстве

- Участники делятся на группы.
- Каждая группа будет оценивать зрелость принятия решение на основе данных в правительстве (в целом), используя шкалу NESTA:

- **Зарождающийся** – Богат данными, обделен интеллектом. Данные не являются ключевой частью процессов принятия решений.
- **Базовый** – Данные используются в отчетах, но, как правило, поверхностно и практически без ссылок на решения, которые необходимо принять.
- **Средний** – Анализ данных обычно запрашивается для принятия решения, но может быть не отвечающим требованиям, потому что анализ не высокого качества, не нацелен на решение, которое должно быть принято, или нужных данных нет.
- **Продвинутый** – Некоторые решения основываются на данных как с низового, так и старшего уровней, но они не согласованы в рамках всей организации.
- **Датавар (Datavore)** – Богатый интеллект данных и понимание. Данные анализируются специально для принятия ключевых решений последовательно во всей организации. Данные доступны своевременно для поддержки принятия решений.

Все группы представляют результаты.

Литература для чтения

B. Guy Peters and Jon Pierre *Comparative Governance: Rediscovering the Functional Dimension of Governing* (Cambridge, Cambridge University Press, 2016)

Terence Lutes, “Data-driven government: Challenges and a path forward”, IBM Analytics White Paper, 2015. Доступно по <https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=GQW03008USEN>

Tom Symons, “Wise Council: Insights From the Cutting Edge of Data-Driven Local Government”, NESTA, November 2016. Доступно по https://media.nesta.org.uk/documents/wise_council.pdf.

Joel Schwartz, “Data-Driven Decision Making: A Primer For Beginners”, Northeastern University Blogs, 14 December 2017. Доступно по <https://www.northeastern.edu/graduate/blog/data-driven-decision-making/>.

Justin Parkhurst, *The Politics of evidence: From evidence-based policy to the good governance of evidence* (Oxford, Routledge, 2017).

UNDG, “Results-Based Management Handbook: Harmonizing RBM concepts and approaches for improved development results at country level”, October 2011. Доступно по <https://undg.org/wp-content/uploads/2016/10/UNDG-RBM-Handbook-2012.pdf>.

UNDP, “Guidance Note: Data for Implementation and Monitoring of the 2030 Agenda for Sustainable Development”, September 2017. Доступно по <http://www.undp.org/content/undp/en/home/librarypage/poverty-reduction/guidance-note--data-for-implementationand-monitoring-of-the-203.html>.

Раздел 3 Мир переполнен данными

Несмотря на революцию в области данных, многие правительства в развивающихся странах по-прежнему сталкиваются с проблемой нехватки качественных данных. Поэтому важно изучить возможности улучшения традиционных источников данных и использование новых источников данных для управления. В этом разделе рассматриваются официальная статистика, большие данные, данные в реальном времени, малые данные и данные, созданные гражданами.

Ключевые моменты

1. Официальная статистика - это числовые наборы данных, производимые официальными государственными органами, в основном, для административных целей.

Ее целью является: «Получение и распространение авторитетных результатов, предназначенных для надежного отражения экономически и социально значимых явлений сложной и динамичной реальности в определенной стране»

В отличие от данных частного сектора, официальная статистика основана на общих принципах, стандартах, методологиях и технологиях, установленных в соответствии с профессиональным этическим кодексом.

Традиционно официальная статистика состоит из данных переписи, данных опросов и административных данных.

Несмотря на попытки быть всеобъемлющими, отсутствуют официальные данные по таким важным вопросам, как гендер. В отчете «*Мир, который считает*» отмечается, что гендерное неравенство и недооценка деятельности и приоритетов женщин во всех сферах были воспроизведены в статистических данных.

2. Большие данные – Массивы данных, размер которых превышает возможности обычных программных средств по администрированию баз данных для сбора, хранения, управления и анализа. Они характеризуются тремя V (volume-объем, variety-разнообразие и velocity-скорость), где «объем» относится к количеству данных, «разнообразие» относится к числу типов данных, а «скорость» относится к скорости обработки данных.

Другие предлагают, что существует не только три, а пять V больших данных, добавляя «*veracity-правдивость*» (надежность или достоверность данных) и «*value-значение*» (выгода или ценность) к объему, разнообразию и скорости.

Некоторые сторонники определяют большие данные как:

- **Исчерпывающие данные** – Пассивно собираемые данные об использовании людьми цифровых услуг, таких как мобильные телефоны, финансовые транзакции или результаты интернет-поисков.
- **Данные зондирования** – Активно собираемые данные с датчиков, например, в

умных городах или от носимых устройств, а также с помощью дистанционного зондирования и спутниковых снимков.

- **Цифровой контент** – Открытый интернет-контент, активно создаваемый людьми: социальные сети, новостные статьи, блоги или объявления о работе. В отличие от исчерпывающих и данных зондирования цифровой контент специально кем-то редактируется, поэтому он субъективен или даже обманчив, в зависимости от намерений автора.

3. Данные в режиме реального времени – данные, которые передаются и используются сразу после сбора. Однако в области развития это также относится к информации, которая создается и становится доступной за относительно короткий и соответствующий период времени, и информации, которая становится доступной в течение срока, позволяющего предпринять ответные действия.

Данные в режиме реального времени включают в себя сообщения в социальных сетях, спутниковые снимки, данные об уровне осадков и потоков, отслеживаемые датчиками, а также данные о местонахождении смартфонов.

Данные в режиме реального времени позволяют принимать своевременные решения, а также предотвращают мошенничество.

4. Малые данные - человеко-ориентированная альтернатива большим данным. Они также характеризуются в качестве небольших наборов данных, которые могут влиять на решения в режиме реального времени.

Малые данные связывают людей со своевременной, полезной информацией (полученной из больших данных и/или «локальных» источников), организованной и оформленной – часто визуально – чтобы быть доступной, понятной и применимой для повседневных задач.

По словам Джона Спейси, малые данные ранее просто назывались данными. Современный термин используется для различия между традиционными конфигурациями данных и большими данными.

5. Данные, генерируемые гражданами – Данные, которые люди или их организации производят для непосредственного мониторинга, запроса или стимулирования изменений по затрагивающим их вопросам.

Данные, сгенерированные гражданами, могут привлечь внимание к вопросам, важным для людей, и представить свое мнение во время политических дискуссий на более высоком уровне.

Применения:

- Проверка официальной информации и наборов данных;
- Обеспечение подотчетности процессов развития через независимый мониторинг;
- Обеспечение прямого представительства граждан и предоставление им возможности напрямую отслеживать, требовать или инициировать изменения по

вопросам, которые их касаются.

Примерами могут служить данные, созданные гражданами о качестве воздуха в Пекине, карта сексуальных домогательств в Египте (см. Вставку 16) и обновленные статусы по пунктам водоснабжения в Танзании.

Предлагаемое упражнение

Данные, генерируемые гражданами

- Участники делятся на группы.
- Каждая группа должна определить и обсудить возможную инициативу по данным, генерируемым гражданами, в своей стране для оказания помощи в мониторинге Целей в области устойчивого развития.
- Все группы представляют результаты обсуждений.

Литература для чтения

Walter J. Radermacher, "The Future Role of Official Statistics", *Power from Statistics Outlook Report*. Доступно по https://www.researchgate.net/publication/320616460_The_Future_Role_of_Official_Statistics

Soenke Ziesche, *Innovative Big Data Approaches for Capturing and Analyzing Data to Monitor and Achieve the SDGs* (Bangkok, ESCAP, 2017). Доступно по <https://reliefweb.int/sites/reliefweb.int/files/resources/Innovative%20Big%20Data%20Approaches%20for%20Capturing%20and%20Analyzing%20Data%20to%20Monitor%20and%20Achieve%20the%20SDGs.pdf>

United Nations Global Pulse, "Big Data for Development: Challenges and Opportunities", May 2012. Доступно по <http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseMay2012.pdf>.

Andreas Pawelke and others, *Data for development: What's next? – Concepts, trends and recommendations for German development cooperation* (Bonn and Eschborn, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, 2017). Доступно по http://webfoundation.org/docs/2017/12/Final_Data-for-development_Whats-next_Studie_EN-1.pdf.

Roger Dooley, "Small Data: The Next Big Thing", *Forbes*, 16 February 2016. Доступно по <https://www.forbes.com/sites/rogerdooley/2016/02/16/small-data-lindstrom/#1811e5fb7870>.

Development Initiatives, "Citizen-generated data and sustainable development: Evidence from case studies in Kenya and Uganda report", March 2017. Доступно по <https://hivos.org/sites/default/files/publications/15-citizen-generated-data-and-sustainable-developmentevidence-from-case-studies-in-kenya-and-uganda.pdf>.

Раздел 4 Управление и администрирование данными

Управление и администрирование данными обеспечивают возможность того, что данные легко доступны, легко обмениваются и легко используются повторно. Это промежуточный вид деятельности / этап между сбором / генерацией данных и анализом данных для возможных закономерностей.

Ключевые моменты

1. Управление данными – это комплексный процесс контроля целостности, использования, доступности, удобства использования и безопасности всех данных, принадлежащих или контролируемых организацией.

Это система прав на принятие решений и ответственности за процессы, связанные с информацией, которые выполняются в соответствии с согласованными моделями, описывающими, кто может, какие действия предпринимать, с какой информацией и когда, при каких обстоятельствах, какими методами.

Цели управления данными состоят в том, чтобы:

- Предоставить организации возможность управлять своими данными как активом;
- Определить, утвердить, сообщить и реализовать принципы, политики, процедуры, показатели, инструменты и обязанности по управлению данными;
- Проводить мониторинг и осуществлять руководство политикой на предмет соответствия, предоставлять услуги по использованию и управлению данными.

2. Администрирование данными - Разработка и выполнение процессов, архитектур, политик, практик и процедур для управления информацией, генерируемой организацией.

Оно включает в себя получение, проверку, хранение, защиту и обработку данных, чтобы обеспечить их доступность, надежность и своевременность для своих пользователей.

Администрирование данными помогает обеспечить доступность данных, когда и где они необходимы. Это особенно необходимо, когда государственные данные хранятся в различных форматах и системах, которые поддерживают деятельность конкретных программ, отделов и организаций.

Администрирование данными сосредоточено на следующих областях знаний:

- **Архитектура данных** - Определяет потребности организации в данных (независимо от структуры), а также разрабатывает и поддерживает основной план для удовлетворения этих потребностей;
- **Моделирование и дизайн данных** - Процесс выявления, анализа и определения требований к данным, а затем представления и передачи этих требований к данным в точной форме, называемой моделью данных;

- **Хранение и операции с данными** - Включают в себя проектирование, реализацию и поддержку хранимых данных, чтобы обеспечить их максимальную пользу в течение всего жизненного цикла: от создания/сбора до утилизации. Они включают два элемента деятельности – обслуживание баз данных и техническая поддержка данных;
- **Безопасность данных** – Планирование, разработка и выполнение политик и процедур безопасности для обеспечения надлежащей аутентификации, авторизации, доступа и аудита массивов данных и информации;
- **Интеграция и совместимость данных** – Процессы, связанные с перемещением и консолидацией данных внутри и между хранилищами данных, приложениями и организациями;
- **Управление документацией и контентом** – Контроль сбора, хранения, доступа и использования данных и информации, хранящихся вне реляционных баз данных;
- **Справочные и базовые данные** – Управление общими данными для достижения целей организации, снижения рисков, связанных с избыточностью данных, обеспечения более высокого качества и снижения стоимости интеграции данных;
- **Хранение данных и бизнес-аналитика** – Процессы планирования, внедрения и контроля для предоставления данных в целях принятия решений и поддержки работников умственного труда, занимающихся отчетностью, запросами и анализом;
- **Метаданные** – Содержат информацию о технических и бизнес-процессах, правилах и ограничениях данных, а также логических и физических структурах данных;
- **Управление качеством данных** – Планирование, внедрение и контроль действий, в которых применяются методы управления качеством к данным, чтобы обеспечить их пригодность для использования и удовлетворения нужд потребителей данных.

3. Прогресс в управлении данными можно измерить с помощью модели зрелости управления данными. Ниже приведены уровни указанной модели:

- **Уровень 0 – Нет способности:** Нет организованных методов управления данными или формальных корпоративных процессов для управления данными.
- **Уровень 1 – Начальный/Специальный:** Практика управления данными общего назначения с использованием ограниченного набора инструментов, практически без управления.
- **Уровень 2 – Повторяемый:** Появление согласованных инструментов и определение роли для поддержки выполнения процесса. Организация начинает использовать централизованные инструменты и обеспечивает больше контроля над управлением данными.
- **Уровень 3 – Определенный:** Новые возможности управления данными. Внедрение и институционализация масштабируемых процессов управления

данными и представление управления данными в качестве организационного инструмента.

- **Уровень 4 – Управляемый:** Стандартизированные инструменты для управления данными от настольных компьютеров до инфраструктуры в сочетании с хорошо сформированными функциями централизованного планирования и управления. Достижение этого уровня представляют собой ощутимое повышение качества данных и возможностей всей организации, таких как сквозной аудит данных.
- **Уровень 5 – Оптимальный:** Распространение данных контролируется для предотвращения ненужного дублирования. Хорошо понятные метрики используются для управления и измерения качества данных и процессов.

Предлагаемое упражнение

Оценка зрелости управления данными

- Участники делятся на группы.
- Каждая группа обсуждает и определяет уровень зрелости управления данными правительства (от 0 до 5).
- Все группы представляют результаты обсуждений.

Литература для чтения

DAMA International, Data Management Body of Knowledge, second edition (DAMA-DMBOK 2) (New Jersey, Technics Publications, 2017).

Data Governance Institute, "Definitions of Data Governance". Доступно по http://www.datagovernance.com/adg_data_governance_definition/.

Pulse Lab Jakarta, "Identifying Opportunities to Strengthen Data Governance in the Public Sector", Medium, 8 October 2017. Доступно по <https://medium.com/pulse-lab-jakarta/identifying-opportunities-to-strengthen-data-governance-in-the-public-sector-92d9bc7db5ad>.

Matthew Chase, "The state of data management in the public sector in 2018", Experian, 8 February 2018. Доступно по <https://www.edq.com/blog/the-state-of-data-management-in-the-public-sector-in-2018/>

Раздел 5 Аналитика данных, алгоритмическая необъективность и интуиция

Чтобы получить представление о данных, они должны быть проанализированы.

Также необходимо следить за необъективностью и назначать правильную роль интуиции в принятии решений.

Ключевые моменты

1. Аналитика данных – Стремление извлекать смысл из необработанных данных с использованием специализированных компьютерных систем ... которые преобразуют, систематизируют и моделируют данные, чтобы сделать выводы и определить закономерности.

Рассмотрим четыре типа аналитики:

- **Описательная аналитика** – Дает ответ на вопрос: Что произошло?
- **Диагностическая аналитика** – Сравнивает исторические данные с другими данными, чтобы ответить на вопрос: Почему что-то случилось?
- **Прогностическая аналитика** – Использует статистические модели и методы прогнозов, чтобы ответить на вопрос: Что может случиться?
- **Предписывающая аналитика** – Подсказывает, какие действия предпринять.

Использование аналитики предполагает применение алгоритмического или механического процесса для получения информации.

Алгоритмы являются эффективными и действенными инструментами для анализа и решения проблем. Тем не менее, они не обязательно являются нейтральными или объективными.

2. Алгоритмическая необъективность

Алгоритм представляет собой набор математических инструкций или правил, которые, особенно с помощью компьютера, помогут вычислить ответ на проблему.

Алгоритмическая необъективность имеет место, когда человеческие предрассудки и предвзятость присутствуют на этапе проектирования. Следовательно, дискриминация (неодинаковое отношение) встраивается в модель.

Алгоритмическая необъективность может возникнуть в результате проблем на разных этапах. Они включают предвзятые или, по-другому, исходные данные низкого качества, плохо определенные правила, недостаток контекстной осведомленности и петли обратной связи.

Алгоритмическая необъективность может привести к двум типам ущерба:

- **Распределительный вред** – Когда алгоритм использует или удерживает определенные возможности или ресурсы на основе предвзятых предположений;
- **Репрезентативный вред** – Когда системы усиливают зависимость некоторых групп по линии идентичности.

Еще одна проблема, которая возникает при использовании аналитики и алгоритмов – это роль интуиции в принятии решений и управлении.

3. Интуиция и принятии решений

Малкольм Гладуэлл в своей книге «*Озарение. Сила мгновенных решений*» 2005 года отстаивает непреходящее значение интуиции.

Исследования выявили следующие условия, когда можно положиться на интуицию:

- **Наличие опыта** – Чем больше опыта в определенной области, тем больше шансов использовать интуицию для принятия решений по конкретным направлениям;
- **Характер проблемы** – Интуиция может быть использована для неструктурированных проблем – тех, в которых нет четких правил принятия решений или мало объективных критериев для принятия решения;
- **Доступное время** – В ситуациях, которые требуют быстрого решения, и для подробного анализа остается мало времени.

Интуиция НЕ должна использоваться в ситуациях, имеющих четкие правила принятия решений, объективные критерии и значительный объем данных, с помощью которых можно выполнить анализ.

Важно отметить, что:

- Имеются убедительные доказательства того, что всякий раз, когда доступны варианты, использование одних только данных и алгоритмов обычно приводит к лучшим решениям и прогнозам, чем полагаться на мнение даже опытных людей и «экспертов».

Предлагаемые упражнения

Аналитика данных

- Разделите участников на четыре группы.
- Распределите темы: Группа 1 – описательная аналитика, Группа 2 – диагностическая аналитика, Группа 3 – прогностическая аналитика и Группа 4 – предписывающая аналитика.
- Всем группам необходимо обсудить фактическое или потенциальное

использование для деятельности правительства назначенной им типовой аналитики.

- Все группы представляют результаты обсуждения.

Пределы интуиции

- Разделите участников на группы.
- Каждая группа определяет и обсуждает определенные / конкретные примеры / случаи, когда существуют три условия (наличие опыта, характер проблемы, доступное время) для оправдания использования интуиции.
- Все группы представляют результаты обсуждения.

Литература для чтения

Anushka Mehta, "Four Types of Business Analytics to Know", *Analytics Insight*, 13 October 2017. Доступно по <https://www.analyticsinsight.net/four-types-of-business-analytics-to-know/>.

Sid Frank and Traci Gusher, "Better data, better government: Effective use of data and analytics at all levels deliver improved citizen services and outcomes", *KPMG Government Institute*, June 2016. Доступно по <https://assets.kpmg/content/dam/kpmg/pdf/2016/06/cogv-6-better-data,-better-government.pdf>

Keith Kirkpatrick, "Battling Algorithmic Bias" *Communications of the ACM*, vol. 59, no. 10 (October 2016). Доступно по <https://cacm.acm.org/magazines/2016/10/207759-battling-algorithmic-bias/abstract>.

World Wide Web Foundation, *Algorithmic Accountability: Applying the concept to different country contexts* (Washington D.C., 2017). Доступно по http://webfoundation.org/docs/2017/07/Algorithms_Report_WF.pdf.

United Kingdom Department for Digital, Culture, Media and Sport, "Guidance: Data Ethics Framework (updated 30 August 2018)". Доступно по <https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework>.

Malcom Gladwell, *Blink: The Power of Thinking Without Thinking* (New York and Boston, Blackbay Books, 2005).

Connson Chou Locke, "When It's Safe to Rely on Intuition (and When It's Not)", *Harvard Business Review*, 30 April 2015. Доступно по <https://hbr.org/2015/04/when-its-safe-to-rely-on-intuition-and-when-its-not>.

Раздел 6 Благоприятная политика

Управление на основе данных не возникнет самопроизвольно в результате революции данных. Для его реализации должны быть приняты или реализованы политики и программы. Они включают в себя укрепление процесса сбора данных; институционализация управления данными; повышение общественного доступа и доверия, а также готовность к большим данным.

Ключевые моменты

1. Укрепление процесса сбора данных

Существует три аспекта для усиления сбора данных: улучшение национальных статистических систем, улучшение гендерной статистики и стимулирование генерирования гражданами данных.

Кейптаунский глобальный план действий в области данных по устойчивому развитию является полезным руководством по улучшению сбора данных на национальном уровне.

Он обеспечивает планирование необходимых действий для регулярного получения качественных и своевременных данных для информирования об устойчивом развитии на требуемом уровне разбивки и охвата населения, включая наиболее уязвимых и труднодоступных групп.

Он определяет шесть стратегических направлений для действий:

- Координация и стратегическое лидерство в области данных для устойчивого развития;
- Инновации и модернизация национальных статистических систем;
- Укрепление основных статистических мероприятий и программ с особым акцентом на удовлетворение потребностей в мониторинге Повестки дня на период до 2030 года;
- Распространение и использование данных для обеспечения устойчивого развития;
- Многосторонние партнерства для получения данных по устойчивому развитию;
- Мобилизация ресурсов и координация усилий по наращиванию статистического потенциала.

Также необходимо улучшить гендерную статистику – научное представление и интерпретация статистических данных, которые соответствующим и полным образом отражают условия жизни и положение женщин и мужчин во всех сферах и областях политики.

Ключевым ресурсом для данной области деятельности является Минимальный набор гендерных показателей Статистической комиссии ООН. Этот набор включает 52 количественных показателя и 11 качественных показателей, сгруппированных в пять категорий: (1) экономические структуры и доступ к ресурсам; (2) образование; (3)

здравоохранение и сопутствующие услуги; (4) общественная деятельность и принятие решений; и (5) защита прав женщин и детей.

Другим ключевым ресурсом является основной набор гендерных показателей ЭСКАТО для Азиатско-Тихоокеанского региона, который является дополнением к глобальному Минимальному набору гендерных показателей.

Азиатско-тихоокеанский базовый набор гендерных показателей состоит из следующего:

- Шесть основных категорий;
- Пять дополнительных категорий;
- Приоритетная целевая группа (сельские женщины);
- Качественные показатели, связанные с национальными критериями.

Третий компонент в усилении сбора данных – это данные, генерируемые гражданами.

Данные, создаваемые гражданами, дополняют официальную статистику. Они также напрямую касаются вопроса справедливости на основе данных, поскольку данные, созданные гражданами, могут сделать видимыми отдельных лиц и группы, которые невидимы для правительства из-за отсутствия данных.

Для правительства действия по стимулированию генерации данных гражданами включают следующее:

- Поиск путей для взаимодействия, признания и поддержки этих инициатив.
- Изучение моделей финансирования и обеспечения устойчивости для данных, генерируемых гражданами, и данных гражданского общества.
- Поддержка процедур консультаций и участия, позволяющих гражданам и гражданскому обществу вносить вклад в сбор государственных данных, в том числе посредством проведения мероприятий, оперативных каналов коммуникации и процессов совместной разработки.
- Поддержка дальнейших исследований и разработки ресурсов в этой области, которые могут быть использованы для того, чтобы позволить инфраструктурам государственных данных лучше учитывать интересы и проблемы гражданского общества.

2. Институционализация управления данными

Правительствам следует рассмотреть возможность разработки стратегии управления данными, которая включает:

- драйверы, видение, миссия и принципы управления данными, включая оценку готовности, выявление внутренних процессов, а также текущих проблем или критериев успеха;
- структуры и обязанности для управления данными;
- временные рамки для осуществления политики и директив, бизнес-гlossарий, архитектура, оценки активов, стандартов и процедур, ожидаемых изменений в бизнес-процессах и технологических процессах, а также результатов для поддержки аудиторской деятельности и соответствия нормативным требованиям;
- конечное состояние устойчивой деятельности по управлению данными.

3. Повышение общественного доступа и доверия к данным

Включает в себя два действия - принятие политики открытых данных и усиление конфиденциальности данных.

- **Открытие государственных данных**

Открытые данные (или открытые данные правительства) представляют собой данные, произведенные или введенные в действие правительством или организациями, контролируемые государством, которые могут свободно и повторно использоваться и распространяться любым лицом. Это также философия – все большее число политик, которые способствуют прозрачности, подотчетности и созиданию, делая государственные данные доступными для всех.

Три составных элемента открытых данных для достижения прозрачности, расширения прав и возможностей граждан и инноваций:

- публикация открытых данных правительствами;
- преобразование данных в полезную информацию организациями-посредниками;
- использование данных гражданами, государственными должностными лицами и другими заинтересованными сторонами для достижения результатов развития.

Фонд Sunlight предоставило правительствам полезное руководство по разработке и реализации политики открытых данных. Руководство состоит из трех частей:

- Какие данные должны быть общедоступными (7 рекомендаций);
- Как сделать данные общедоступными (16 рекомендаций);
- Как реализовать политику открытых данных (8 рекомендаций).

- **Укрепление конфиденциальности данных**

Это включает принятие национального законодательства о конфиденциальности данных, которое усиливает контроль человека над тем, как его/её личная информация собирается, используется, передается и удаляется.

В 1980 году Организация экономического сотрудничества и развития (ОЭСР) опубликовала следующие принципы защиты данных:

- Принцип ограничения сбора;
- Принцип качества данных;
- Принцип спецификации цели;
- Принцип ограничения использования;
- Принцип обеспечения безопасности;
- Принцип открытости;
- Принцип индивидуального участия;
- Принцип ответственности.

Данные принципы были приняты в различных национальных законодательствах и

типовых кодексах по всему миру.

Недавним событием является использование этики в качестве важного дополнения к юридическим обязательствам для достижения целей обеспечения конфиденциальности.

Нобелевский лауреат Джозеф Стиглиц считает, что существует необходимость принятия правил для следующего: «Какие данные могут хранить технологические фирмы; какие данные они могут использовать; могут ли они объединять разные наборы данных; цели, для которых они могут использовать эти данные; и какую степень прозрачности они должны предоставить при вопросе о том, что они делают с данными».

4. Готовимся к большим данным

Создание благоприятной среды для больших данных требует:

- Лидерства – Обязательство руководить на всех уровнях и во всех областях;
- завоевание и сохранение общественного доверия – Это означает: (1) персональные данные, собранные и хранящиеся в ведомстве, используются надлежащим образом и эффективно; и (2) персональные данные в безопасности, особенно когда они передаются между различными ведомствами.
- Нарастивание потенциала государственной службы по сбору, хранению, анализу, обмену и использованию данных.

Также необходимо рассмотреть семь аспектов использования больших данных в целях развития:

1. Информация – Включает в себя как информацию, так и данные, предшествующие созданию информации;
2. Технологии – Основное внимание уделяется ИКТ, которые обрабатывают данные;
3. Процессы – Действия, предпринятые для производства, сбора, анализа, представления и использования данных;
4. Цели и ценности – Компонент «цели» охватывает вопросы достижения собственных интересов и реализации политики, а также неформальные и формальные стратегии; и компонент «ценности» охватывает культуру (например, что заинтересованные стороны считают правильным и неправильным способом сделать что-то);
5. Навыки и знания – Охватывает качественные и количественные аспекты человеческих компетенций для выполнения процессов, связанных с данными;
6. Системы и структуры управления – Более развернутые системы управления, необходимые для взаимодействия внутри и между организациями и сетями, связанными с данными, а также структурная организация системы как формально, так и неформально;
7. Другие ресурсы – Время и деньги.

Предлагаемое упражнение

Краткое представление управления на основе данных

- Разделите участников на 4 группы.
- Каждая группа должна оценить, используя трехбалльную шкалу, готовность правительства к управлению на основе данных, оценивая развитие в следующих направлениях: (1) усиление сбора данных; (2) институционализация управления данными; (3) повышение общественного доступа и доверия; и (4) готовность к большим данным.
- Для облегчения обсуждения и составления отчета, пожалуйста, используйте матрицу ниже.

	0 Не начался	1 Некоторый прогресс	2 Завершено
1) Сбор данных <ul style="list-style-type: none"> • Совершенствование национальной системы статистики • Улучшение гендерной статистики • Минимальный набор гендерных показателей • Азиатско-тихоокеанский базовый набор • Стимулирование генерации данных гражданами 			
2) Институционализация управления данными			
3) Повышение общественного доступа и доверия к данным <ul style="list-style-type: none"> 3. Открытие государственных данных 4. Укрепление конфиденциальности данных 			
4) Готовность к большим данным <ul style="list-style-type: none"> • Лидерство • Общественное доверие • Наращивание потенциала государственной службы 			

Все группы представляют результаты обсуждения.

Литература для чтения

High-level Group for Partnership, Coordination and Capacity-Building for Statistics for the 2030 Agenda for Sustainable Development, "Cape Town Global Action Plan for Sustainable Development Data", adopted by the United Nations Statistical Commission at its 48th

Session, March 2017, p. 2. Доступно по https://unstats.un.org/sdgs/hlg/Cape_Town_Global_Action_Plan_for_Sustainable_Development_Data.pdf

United Nations Statistics Division, “Minimum Set of Gender Indicators”. Доступно по <https://genderstats.un.org/#/home>.

ESCAP, “Core set of gender indicators for Asia and the Pacific: Note by the secretariat”, Committee on Statistics, Fourth Session, 23 January 2015 (E/ESCAP/CST(4)/10). Доступно по http://www.asiapacificgender.org/sites/default/files/pdf/statstics_documents/Core_Set_Gender_Indicators_Asia_Pacific.pdf.

Danny Lammerhirt, Shazade Jaeson and Eko Presetyo, “Making Citizen Generated Data Work: Towards a Framework Strengthening Collaborations Between Citizens, Civil Society Organisations, and Others”, Data Shift, March 2017. Доступно по http://civicus.org/thedatashift/wp-content/uploads/2017/03/Making-Citizen-Generated-Data-Work_short-report_.pdf.

NASCIO, “Data Governance Part II: Maturity Models – A Path to Progress”, March 2009. Доступно по <https://nascio.org/Portals/0/Publications/Documents/NASCIO-DataGovernancePTII.pdf>.

Sunlight Foundation, “Open Data Policy Guidelines”, Open Data Policy Hub. Доступно по <https://opendatapolicyhub.sunlightfoundation.com/guidelines/>.

Holvast, “History of Privacy”, in *The Future of Identity*, V. Matyáš and others, eds. (International Federation for Information Processing, 2009). Доступно по https://link.springer.com/content/pdf/10.1007/978-3-642-03315-5_2.pdf.

United Nations Global Working Group for Big Data, “Using Big Data for the Sustainable Development Goals”. Доступно по <https://unstats.un.org/bigdata/taskteams/sdgs/>.

Richard Heeks, “Measuring Barriers to Big Data for Development”, *ICTs for Development*, 9 August 2016. Доступно по <https://ict4dblog.wordpress.com/2016/08/09/measuring-barriers-to-big-data-for-development/>.

UNDP and United Nations Global Pulse, *A Guide to Data Innovation for Development: From Idea to Proof-Of-Concept, version 1* (New York, 2016). Доступно по <http://www.undp.org/content/undp/en/home/librarypage/development-impact/a-guide-to-data-innovationfor-development---from-idea-to-proof-.html>.

Раздел 7 Создание культуры данных в государственном секторе

Data culture can accelerate that adoption of data-driven governance and amplify its power.

Ключевые моменты

Культура данных означает повсеместное использование данных в организации.

В государственном секторе культура данных означает уверенный уровень комфортности для всей организации по применению показателей для достижения максимального социального воздействия.

Её шесть принципов заключаются в следующем:

- Культура данных – это культура принятия решений – Основной целью сбора, анализа и развертывания данных является принятие более эффективных решений;
- Культура данных и лидерство – Обязательство со стороны высших должностных лиц имеет важное значение, но это обязательство должно проявляться не только в периодических заявлениях высокого уровня;
- Демократизация данных – Для создания конкурентного преимущества необходимо стимулировать спрос на данные от населения;
- Культура данных и риск – Действующая культура данных ставит в основу риск. Управление рисками должно действовать как интеллектуальный ускоритель, вводя ответственно аналитику в ключевые процессы и взаимодействия;
- Катализаторы культуры – Кто-то должен взять на себя ответственность и приложить значительные усилия для достижения перемен – требуются люди, которые могут соединить оба мира – науку о данных и выполнение действий на местах;
- Сочетание таланта и культуры – Это включает в себя установление соответствующего баланса для любого учреждения между введением новых сотрудников и преобразованием существующих.

Для формирования культуры данных потребуется план, который включает: определение внутренних сторонников/экспертов; обнаружение передового опыта; выстраивание внешних связей, ведущих вверх и вниз; и делать небольшие («детские») шаги.

Литература для чтения

Kathleen Kelly Janus, "Creating a Data Culture", Stanford Social Innovation Review, 2 March 2018. Доступно по https://ssir.org/articles/entry/creating_a_data_culture

Rahul Bhargava, "You Don't Need a Data Scientist, You Need a Data Culture", Data Therapy, 6 December 2017. Доступно по <https://datatherapy.org/2017/12/06/building-a-data-culture/>.

ЗАМЕТКИ ДЛЯ ИНСТРУКТОРА

ЧАСТЬ 2 **Варианты по проведению обучения**

В зависимости от аудитории, временных ограничений и целей существует, как минимум, пять способов преподавания данного модуля.

Курс подготовки преподавателей по управлению на основе данных 2,5 дня

День 1

08:00 - 9:00	Открытие
09:30 - 11:30	Раздел 1: Революция данных и справедливость данных
01:00 - 03:00	Раздел 2: Данные и управление
03:30 - 05:30	Раздел 3: Мир переполнен данными

День 2

08:00 - 10:00	Раздел 4: Управление и администрирование данными
10:30 - 12:30	Раздел 5: Аналитика данных, алгоритмическая необъективность и интуиция
01:00 - 02:30	Раздел 6: Благоприятная политика - 1
03:00 - 04:30	Раздел 6: Благоприятная политика - 2
04:30 - 05:30	Раздел 7: Культура данных

День 3

08:30 - 10:00	Разработка национальных планов преподавания модуля по осуществлению государственного управления на основе данных
10:30 - 11:30	Презентация (Страны / участника)
11:30	Заккрытие

(Обычный) курс по управлению на основе данных 2 дня

День 1

08:30 - 09:00	Открытие
09:30 - 10:30	Раздел 1: Революция данных и справедливость данных
10:30 - 11:30	Раздел 2: Данные и управление
01:00 - 02:30	Раздел 3: Мир переполнен данными
03:00 - 04:30	Раздел 4: Управление и администрирование данными

День 2

08:30 - 10:00	Раздел 5: Аналитика данных, алгоритмическая необъективность и интуиция
10:30 - 12:00	Раздел 6: Благоприятная политика - 1
01:00 - 02:30	Раздел 6: Благоприятная политика - 2
03:00 - 04:00	Раздел 7: Культура данных
04:00 - 05:00	Заккрытие

Курс по управлению на основе данных для руководящего состава 1 день

08:30 - 09:00	Открытие
09:30 - 10:30	Раздел 2: Данные и управление
10:30 - 11:30	Раздел 4: Управление и администрирование данными
01:00 - 2:30	Раздел 6: Благоприятная политика - 1
03:00 - 04:30	Раздел 6: Благоприятная политика - 2
04:30 - 05:00	Заккрытие

Курс по управлению на основе данных - большие данные 1 день

08:30 - 09:00	Открытие
09:30 – 10:30	Разделы 1 и 2: Революция данных и справедливость данных, данные и управление
11:00 – 12:00	Раздел 3: Большие данные (Мир переполнен данными)
01:00 – 02:30	Раздел 5: Аналитика данных, алгоритмическая необъективность и интуиция
03:00 - 04:00	Раздел 6: Готовимся к большим данным (Благоприятная политика)
04:00 - 05:00	Раздел 7: Культура данных
05:00	Заккрытие

Курс по управлению на основе данных - управление и администрирование данных 1 день

8:30 - 9:00	Открытие
9:30 - 10:30	Раздел 1: Революция данных и справедливость данных
11:00 – 12:00	Раздел 2: Данные и управление
01:00 - 02:30	Раздел 4: Управление и администрирование данными – 1
03:00- 04:30.	Раздел 4: Управление и администрирование данными – 2
04:30- 05:00	Раздел 7: Культура данных
05:00	Заккрытие

АТУЦ ИКТР / ЭСКАТО

Азиатско-тихоокеанский учебный центр информационно-коммуникационных технологий в целях развития (АТУЦ ИКТР) является региональным институтом Экономической и социальной комиссии ООН для Азии и Тихого океана (ЭСКАТО). Целью АТУЦ ИКТР является укрепление усилий стран-членов ЭСКАТО по использованию ИКТ в своем социально-экономическом развитии на основе наращивания человеческого и институционального потенциала. Работа АТУЦ ИКТР сосредоточена в трех направлениях: обучение, обмен знаниями, а также диалог и партнерство с участием многих заинтересованных сторон. Вместе они формируют комплексный подход к созданию человеческого потенциала ИКТ.

АТУЦ ИКТР находится в Инчхоне, Республика Корея.
<http://www.unapcict.org>

ЭСКАТО

Экономическая и социальная комиссия для Азии и Тихого океана (ЭСКАТО) является региональным отделением Организации Объединенных Наций по вопросам развития и служит главным центром экономического и социального развития Организации Объединенных Наций в Азиатско-Тихоокеанском регионе. Её задача заключается в развитии сотрудничества между 53 членами и 9 ассоциированными членами. ЭСКАТО обеспечивает стратегическую связь между глобальными и страновыми программами и вопросами. Она поддерживает правительства стран региона в деле укрепления региональных позиций и защищает региональные подходы в решении уникальных социально-экономических проблем региона в условиях глобализации в мире.

ЭСКАТО находится в Бангкоке, Таиланд.
<http://www.unescap.org>



APCICT/ESCAP

United Nations Asian and Pacific Training Centre for Information
and Communication Technology for Development
Bonbudong, 3rd Floor Songdo Techno Park 7-50
Songdo-dong, Yeonsu-gu, Incheon, Republic of Korea

www.unapcict.org