

**Европейская экономическая комиссия****Конференция европейских статистиков****Группа экспертов по переписям населения
и жилищного фонда****Двадцать первое совещание**

Женева, 18–20 сентября 2019 года

Пункт 2 предварительной повестки дня

**Результаты тестирования с точки зрения методологии,
технологии, участия и других аспектов****Обзор результатов испытаний, проведенных для
повышения качества переписи населения и жилищ****Записка Центрального статистического управления Польши*****Резюме*

В процессе подготовки к переписи в рамках раунда переписей 2020 года в Польше был проведен ряд организационных, методологических и технологических испытаний с целью разработки оптимальных решений для применения в ходе переписи, которая будет проводиться в 2021 году. Во внимание были приняты следующие аспекты:

- a) оптимизация применяемых методов и инструментов сбора данных;
- b) объем ресурсов и средств (в том числе финансовых), необходимых со стороны как исполнителей переписи, т. е. официальных статистических органов, так и учреждений, которые обязаны предоставлять переписные данные;
- c) удобство применяемых решений для пользователей;
- d) коммуникационная и технологическая эффективность разработанных ИТ-приложений;
- e) безопасность собранных и обработанных данных переписи;
- f) оптимизация сроков проведения отдельных этапов получения данных от респондентов;
- g) полнота и качество результатов переписи;

* Переиздано по техническим причинам 14 августа 2019 года.

** Документ подготовили Януш Дыгашевич и Магдалена Янчур-Кнапек.



h) оптимизация, прозрачность и эффективность сообщений относительно популяризации переписи, ориентированных на переписные учреждения.

Каждый из этих аспектов был рассмотрен с точки зрения максимального упрощения методики проведения переписи наряду с поддержанием высокого уровня качества полученных результатов. В настоящем документе представлены результаты проведенных испытаний и соответствующие выводы.

I. Введение

1. Перепись населения и жилищ – это исключительное по своим масштабам и значению статистическое обследование. Оно не сравнимо с любым другим статистическим обследованием, поскольку охватывает все население страны с целью сбора и подготовки полной информации о положении и социодемографической структуре населения и жилищных ресурсов страны. Оно проводится каждые 10 лет. За это время динамично развиваются как методика переписей, так и методы сбора данных у респондентов. Активно меняется рынок технологических средств и систем, основанных на информационно-коммуникационной технологии (ИКТ), которые существенно влияют на разработку методов сбора данных и на участие респондентов в переписи. Все более широкий спектр возможностей для прямой рассылки информационных сообщений респондентам, в том числе через цифровые инструменты, способствует формированию социально и статистически ответственного поведения (воздействуя на уровень общественного восприятия официальной статистики).

2. С учетом вышеупомянутых аргументов можно утверждать, что с годами в системе проведения переписей населения и жилищ происходят постоянные изменения, позволяющие повышать качество этих переписей. Неизменным остается только принцип сохранения сопоставимости результатов и дальнейшего формирования временных рядов.

II. Анализ ситуации

3. С самого начала проведения переписей населения их первоочередной целью всегда являлись расчет численности населения и анализ его социодемографической структуры. Со временем тематический охват собираемой у населения информации был расширен. За несколько десятилетий проведения переписей спектр переписных данных эволюционировал, благодаря чему в настоящее время производится сбор международного комплекса данных. При необходимости в этот комплекс данных вносятся изменения усилиями группы международных экспертов, действующих в составе целевых и рабочих групп. Однако с точки зрения респондента уменьшение бремени, связанного с обязанностью участвовать в переписи, до настоящего времени было менее значительным. Но теперь ситуация изменилась: общество стало более разборчивым и не всегда соглашается с абсолютным требованием предоставлять данные для переписи при отсутствии каких-либо устройств, позволяющих сократить расход времени для подачи информации, и возможности выбрать современные формы участия в переписи.

4. В странах с развитыми административными регистрами были отмечены преимущества использования регистровых данных для замены информации, получаемой непосредственно от респондентов. Этот процесс проходил постепенно в течение нескольких лет, начавшись с частичного устранения некоторых данных путем их получения из регистров с последующим восполнением остающейся части информации за счет прямых опросов, а затем перешел в русло проведения переписи исключительно на основе регистровых данных. Необходимым условием является абсолютное заверение в надежности административных данных, т. е. дефиниционная совместимость понятий в регистровых и статистических данных, действительность

данных на момент проведения переписи и исчерпывающий охват данных, соответствующий охвату информации, собираемой в ходе переписи. Следствием такой деятельности в странах, проводящих перепись только на основе административных источников, является полное исключение участия респондентов или отмена обязанности участвовать в переписи. Таким образом, посредством минимизации расходов на перепись была обеспечена дополнительная, весьма значительная выгода, которая является еще одним аргументом в пользу изменения порядка проведения переписей.

5. Странам, не имеющим регистров, которые можно использовать для переписи или которые находятся на таком этапе разработки, который не гарантирует полного замещения переписных данных, необходимо изыскивать другие способы ограничения содержания электронных переписей и тем самым сократить число вопросов, чтобы уменьшить нагрузку на респондентов. При этом никогда не следует забывать, что меры содействия респондентам не должны влиять на качество переписи.

III. Обзор условий проведения переписи в Польше

6. Административные регистры, действующие в Польше, пока не достигли такого уровня качества, при котором для получения необходимых результатов переписи можно было бы пользоваться только их материалами. Тем не менее они постоянно совершенствуются, и в 2011 году уже было принято решение частично использовать регистровые данные и дополнить недостающую информацию за счет ответов респондентов. Для уверенности в качестве результатов переписи все вопросы были включены в переписной лист, в том числе те, которые охватывались административными источниками. Цель такой меры состояла в том, чтобы показать респондентам, какая информация о них содержится в регистрах (в переписном листе, включенном в форму главного регистра (ГР)). Их задача заключалась в исправлении данных, если они оказывались устаревшими. Число исправлений, внесенных респондентами, использовалось для оценки качества данных в регистрах.

7. После переписи 2011 года была продолжена работа по повышению качества регистровых данных, чтобы они в большей степени соответствовали требованиям, предъявляемым к официальной статистике. Эта работа всегда ведется в рамках системы официальной статистики, в сотрудничестве с администраторами регистров и при поддержке Министерства цифровых технологий. Представители органов официальной статистики, участвующие в деятельности межведомственных рабочих групп, стремятся уделять особое внимание важнейшей роли регистров и потребности в их согласованности. Кроме того, в исправленном Законе о государственной статистике была усилена роль начальника Центрального статистического управления Польши и его влияние на изменения, вносимые в государственные регистры.

8. При проведении переписи 2021 года описанная в пункте 6 система будет сохранена, но в измененной форме, которая предполагает сопоставление данных из регистров с данными, полученными от респондентов. Однако, учитывая положения Общего регламента по защите данных (ОРЗД), данные из регистров не будут предоставляться респондентам в заполняемой ими электронной форме. В результате непрерывного процесса повышения качества данных из административных источников в конечном счете появится возможность при проведении последующих переписей исключать вопросы, в которых содержатся данные из регистров, а затем полностью отказаться от необходимости привлекать респондентов.

9. В последние годы все более высокую значимость приобретает использование альтернативных источников, в том числе «больших данных». Проводится экспериментальная работа по вопросам наличия и качества этих источников, а также областей для использования «больших данных» при проведении переписи населения и жилищ.

10. Вне зависимости от использования регистров и альтернативных источников продолжается работа по модернизации сбора данных у респондентов, включая применение методов на основе ИКТ-систем.

IV. Выявление областей для внесения изменений

A. Методика формирования инструментов для сбора данных с помощью каналов САХІ: САWІ (компьютеризованного веб-опроса), САТІ (компьютеризованного телефонного опроса) и САРІ (компьютеризованного личного опроса)

11. Методика формирования инструментов для сбора данных предполагает соответствие обследований способу сбора данных. При этом учитываются такие аспекты, как объем вопросника, блоки вопросов, порядок вопросов, фильтрация вопросов и функциональность.

12. Проведение переписи с применением методов САХІ зависит от возможности:

a) для респондента самостоятельно заполнить переписной лист, в котором респонденту будут предложены только личное мнение о содержании вопросов и возможные варианты ответа;

b) подсказки со стороны счетчика переписи в ходе телефонного опроса, когда респондент не будет видеть содержания вопросов или возможных вариантов ответа;

c) для указаний и пояснений, предоставляемых счетчиком в ходе непосредственного опроса.

13. Электронный переписной лист должен быть универсальным инструментом, скорректированным в соответствии с интеллектуальными способностями среднего респондента и с базовыми навыками использования компьютера и интернет-приложений.

14. Существует ряд следующих основных аспектов, которые нужно учитывать при разработке электронного вопросника для переписи:

a) составление вопросов для электронного вопросника: вопросы нужно формулировать в транспарентном и легко читаемом для заполняющего вопросник лица виде. Они должны иметь простую и доходчивую формулировку, чтобы респондент точно знал, о чем его спрашивают. В противном случае респондент, не понимая сути вопроса, даст ошибочный ответ или пропустит вопрос;

b) объем электронного вопросника: следует избегать включения в формуляр чрезмерного количества вопросов. Вряд ли стоит ожидать, что респондент добросовестно заполнит, например, документ объемом 20 страниц. Слишком объемный вопросник отобьет у респондентов охоту к его заполнению. Электронный формуляр должен заполняться примерно за 15 минут и содержать необходимый минимум вопросов (в среднем 30 вопросов);

c) блоки вопросов: рекомендуется распределить вопросы по тематическим блокам, чтобы они затем составили одно компактное и логическое целое, а вопросник был упорядочен для обеспечения большей доходчивости. Каждый блок следует начинать с общего вопроса, а затем переходить к вопросам более детального характера;

d) порядок вопросов: вопросы в формуляре должны располагаться в правильном порядке. Перед вопросами детального характера следует располагать общие вопросы, а сложные вопросы должны предваряться более простыми;

e) фильтрация: весьма полезно производить фильтрацию вопросов, поскольку она дает возможность исключить темы, которые не касаются определенных групп населения. Тогда респонденту будет проще заполнять вопросник, избегая не затрагивающих его/ее вопросов;

f) указания, пояснения и словари: практические указания и пояснения, используемые в ходе обследования, позволяют респонденту ознакомиться с

определенным аспектом и упрощают выбор правильной категории ответа, тем самым повышая качество получаемых результатов;

g) функциональность: структура и графическое оформление вопросника должны способствовать изложению ответов. Электронный вопросник должен быть понятным и привлекательным и побуждать респондентов к участию в опросе.

15. В порядке экспериментов, направленных на повышение качества переписи, были введены следующие допущения:

a) сбор данных производится с применением смешанных методов опроса, т. е. CAWI, CATI и CAPI;

b) целевой группой является все население страны;

c) содержание информации, полученной при обследовании, отвечает потребностям национальных органов и Евростата;

d) спектр собираемых данных аналогичен при всех методах сбора данных;

e) оптимальный объем электронного вопросника, позволяющий заполнить его в течение 15 минут, и ограничение количества вопросов электронной переписи необходимым минимумом (в среднем – 30 вопросов);

f) вопросы объединяются в тематические блоки;

g) качественное, привлекательное графическое оформление электронного формуляра, в первую очередь – при использовании метода CAWI;

h) фильтрация вопросов: данные вводятся на полях, предусмотренных для заполнения электронного смарт-формуляра при использовании метода CAWI;

i) ввод данных на активных полях, которые указаны для заполнения в смарт-формуляре, или выбор интервьюером блока/вопроса/лица при использовании методов CATI и CAPI;

j) при использовании всех методов CAWI, в частности CATI и CAPI, обеспечивается возможность отхода от порядка заполнения электронного формуляра с автоматическим возвращением в ту или иную точку опроса (возможность дополнить информацию, полученную в процессе опроса по предыдущим тематическим блокам);

k) вопросы в электронном формуляре составлены в доходчивом для среднего респондента виде (не требуют дополнительных пояснений);

l) названия и определения, включенные в электронный формуляр, являются читабельными и понятными для каждого респондента (не требуют дополнительных пояснений);

m) используемые словари и указания появляются после ввода ряда символов применительно к тем вопросам, для которых это возможно и когда это обеспечит высокое качество данных;

n) в электронных вопросниках используются данные из административных источников (например, о зданиях, жилищах, зарегистрированных лицах).

1. Влияние новых инструментов для сбора данных на качество результатов

16. Методы CAWI, применяемые для сбора данных, оказывают большое влияние на качество результатов. Вне зависимости от способа проведения переписи удобный для пользования и информативный электронный формуляр повышает заинтересованность респондентов и их готовность участвовать в переписи. Недвусмысленный характер вопросов позволяет избегать ошибок, связанных с их неправильным пониманием. Ясные ответы на четко сформулированные вопросы также влияют на число ошибок, которые предстоит исправить при обработке данных, а также на долю полученных, обработанных и представленных результатов. Поскольку подготовка результатов переписи является кульминацией всего процесса ее проведения, такую подготовку нужно осуществлять качественно и быстро. Применение современных методов сбора данных значительно ускоряет их обработку и уточнение. Распространение данных

переписи в форме микросовокупностей, совокупностей и наборов результатов должно осуществляться новаторскими методами, с предоставлением доступа к аналитической базе данных. Лица, которые по разным причинам пользуются результатами переписей, наверняка будут проявлять к ним повышенный интерес, если электронный переписной лист будет оформлен в более привлекательном для пользователя виде. Информация, полученная с применением методов SAxI, позволяет проводить широкий спектр анализов (включая пространственный анализ) и дает получателям возможность составлять индивидуальные отчеты. Эти отчеты также станут качественнее, если формулировки в переписном листе будут более точными и недвусмысленными.

a) Влияние методов сбора данных

17. В настоящем разделе рассматривается вопрос о том, как различные методы сбора данных влияют на показатели ответа по ключевым переменным и на качество исходных данных.

18. При оценке влияния методов сбора данных, т. е. смешанных приемов опроса, на показатели ответа по ключевым переменным и на качество исходных данных устанавливались следующие допущения:

- a) метод сбора данных влияет на:
 - i) качество исходных данных: виды ошибок, шкалу ошибок, точность, полноту, своевременность сбора данных;
 - ii) объем затрат по сбору данных: со стороны производителя данных и со стороны респондента;
 - iii) организацию сбора данных;
 - iv) форму электронного вопросника;
 - v) уровень стабильности сбора данных;

b) необходимо достижение компромисса между качеством собранных данных и расходами на сбор данных.

19. При оценке влияния методов сбора данных – смешанных способов опроса респондентов – на ключевые показатели ответа большое значение приобретают следующие аспекты:

- a) определение того, какие способы сбора данных можно использовать для обеспечения сочетания значительной доли ответов с низкими затратами на получение данных (сбор данных через Интернет – CAWI – не может являться единственным способом сбора данных из-за низкой доли ответов);
- b) установление последовательности способов сбора данных;
- c) обеспечение соответствия способа сбора данных тем различным группам населения, которые по плану должны быть охвачены переписью.

В. Методы измерения влияния способов сбора данных на качество исходных данных и методы смягчения нежелательных факторов такого влияния

20. При разработке методов измерения влияния способов сбора данных на качество исходных данных были приняты следующие допущения:

- a) оценка влияния методов сбора данных на качество исходных данных производится:
 - i) с учетом стандартного метода, определенного в рамочной системе качества (пункт С);

ii) при наличии установленного эталона для сопоставлений (например, с данными традиционных обследований с использованием статистических методов);

iii) с определенными методами исправления ошибок (например, при формировании выборки);

b) Нужно иметь возможность для количественного определения воздействия способов сбора данных на их качество. Это означает, что в рамках каждого способа следует установить число и тип ошибок, допущенных по аналогичным вопросам.

21. При разработке методов смягчения нежелательного влияния способов сбора данных необходимо прежде всего рассмотреть все факторы, которые могут исключаться вне зависимости от применяемого способа сбора данных; к ним относятся:

a) «коммуникативность» электронного формуляра (независимо от способа сбора данных). Недостаток коммуникативности электронного вопросника в равной степени затруднителен при любом способе сбора данных. Вне зависимости от канала сбора данных (и даже если вопросы имеют двусмысленную формулировку) при необходимости в электронный формуляр следует включать указания, которые помогут респонденту или счетчику заполнить вопросник, а также сократить время, которое требуется для его заполнения. При этом автоматически возрастет уровень качества данных;

b) эффективность мобильного приложения (электронного формуляра). При каждом из способов сбора данных функциональность мобильного приложения очень важна, так как она позволяет быстро переходить от одного вопроса к другому, контролировать правильность логики ответа, подтверждение ответов, показ сообщений об ошибке и т. п. Удобство приложения влияет на качество данных и на отношение лица, заполняющего электронный вопросник, будь то счетчик переписи, статистический интервьюер или сам респондент;

c) свойства информационно-коммуникационной среды, в которой проводится сбор данных. Для отдельных каналов ИКТ-среду нужно формировать таким образом, чтобы все операции, выполняемые с электронным формуляром (чаще всего – на базовом уровне, т. е. незаметно для пользователя), были максимально быстрыми и доходчивыми на интуитивном уровне. Проблемы с применением мобильного приложения или с необходимостью ожидания открытия либо закрытия электронного формуляра, с появлением последующих вопросов и т. п. вызывают у пользователя разочарование и желание прекратить начатую процедуру.

22. После выполнения рекомендаций, изложенных в пунктах а)–с), наверняка будет сделан вывод о том, что большинство факторов, снижающих качество данных, было устранено. При этом останутся причины, связанные непосредственно с лицом, заполняющим электронный формуляр, например «эффект счетчика», или же факторов, которые нельзя предвидеть: спешка, нерасположенность в момент заполнения вопросника (стресс, болезнь и т. п.). Если такие факторы воздействия можно сгладить благодаря коммуникативности электронного формуляра, то, возможно, это позволит надлежащим образом подготовиться к непредвиденным ситуациям.

1. Сильные и слабые стороны многоканального сбора данных

23. Многоканальный сбор данных заключается в одновременном получении данных в ходе переписи по различным каналам; в этом случае одновременность означает продолжительность фазы сбора данных в ходе переписи, а не ссылку на какой-либо конкретный момент времени (определенную дату). Поэтому можно применять разные способы обеспечения активации определенных каналов:

a) одновременный старт в начале проведения переписи и ее одновременное завершение; это означает одновременность действия всех каналов сбора данных. Такая организация сбора данных эффективна только в случаях, когда респонденты настолько хорошо профилированы, что их с высокой степенью вероятности можно

ориентировать на соответствующий канал. Кроме того, у респондентов должна быть высокая заинтересованность в выполнении процедуры самоисчисления, так как в противном случае не исключено, что многие могут отказаться от телефонного или прямого опроса, ссылаясь на выбор самоисчисления, после чего они не пройдут эту процедуру, и из-за недостатка времени к ним не будет повторно отправлен счетчик, и они в конечном счете окажутся обойденными переписью. Такой метод активирования канала не рекомендуется, поскольку он не обеспечивает высокого качества данных;

b) последовательный запуск отдельных каналов без дублирования. Такое происходит в случаях, когда каналы полностью разделены во времени; это означает, что одновременно активным является только один канал. Последний, замыкающий канал должен относиться к системе САРІ, поскольку именно такой канал обеспечивает самый высокий уровень контроля над респондентом. Это – весьма эффективный метод активации канала, позволяющий достичь значительной полноты переписи. Однако его недостатком является необходимость продления времени действия отдельных каналов, вследствие чего весь период сбора данных оказывается длиннее, чем в других организационных вариантах. В результате, хотя это не оказывает негативного влияния на качество данных (более того, возможно даже повышение качества), расходы на проведение переписи возрастают;

c) последовательный запуск каналов с возможным частичным дублированием действий. Важно сохранить порядок активации отдельных каналов и продолжительность их эксклюзивного действия. В особенности важно, чтобы в первую очередь активировался канал онлайн-самоисчисления, поскольку он наиболее экономичен и дает респондентам неограниченную возможность сообщать о себе информацию (по деликатным вопросам), особенно когда речь идет о лицах замкнутого склада. Не менее важно, чтобы в начальный период сбора данных канал САWІ являлся единственным способом проведения переписи. Остальные каналы могут открываться параллельно при наличии комплекса линий телефонной связи с респондентами и возможности назначить этим респондентам метод САТІ. В этом случае счетчиков переписи можно направлять к другим респондентам. Если проведение переписи, включая канал САWІ, было широко популяризировано среди населения, то можно предположить с высокой вероятностью, что большинство граждан выполнит процедуру самоисчисления в течение первых двух недель переписи. Исходя из этого предположения, через две недели рекомендуется активировать последующие каналы. При одновременной активации каналов нет необходимости продлевать срок проведения переписи. Это даст возможность рационально использовать выделенные на перепись финансовые ресурсы. Кроме того, необходимо помнить о скорейшей деактивации самоисчисления, чтобы не блокировать проведение переписи с использованием других методов. Сценарий активации каналов сбора данных, описанный в этом пункте, является наиболее предпочтительным благодаря качеству и полноте данных и рентабельности переписи.

24. Многоканальный сбор данных также дает респонденту возможность выбрать наиболее подходящий вариант для представления информации о себе. Каждый способ имеет свои достоинства и недостатки, которые, однако, не являются аналогичными для всех пользователей. Важно, чтобы каждый респондент выбрал себе наиболее удобный для него канал. Далее описываются характеристики каналов сбора данных, которые могут определить или повлиять на решение респондента относительно выбора того или иного канала:

a) *Онлайн-самоисчисление.* Респонденту назначается определенный срок для заполнения электронного вопросника. В его распоряжении также имеется мобильное приложение, с помощью которого процедура самоисчисления проводится быстро и без затруднений. Респондент должен выбрать правильный момент для работы с приложением независимо от времени суток. Проблемы возникнут в случае, если обнаружится, что в приложении имеются технические или существенные дефекты, которые препятствуют работе с ним. При надлежащей работе приложения дополнительным преимуществом является анонимность, которая не обеспечивается, когда респонденту приходится передавать информацию счетчику, т. е. неизвестному лицу, чего некоторым респондентам хотелось бы избежать. Кроме того, онлайн-самоисчисление

самоисчисление – это приложение, которое после надлежащей подготовки может обеспечить участие в переписи для лиц с ограниченными возможностями, например для глухих, немых или слепых. Такой способ могут также выбрать лица, опасющиеся впускать к себе в дом неизвестного человека (счетчика) по соображениям безопасности.

b) *Телефонный опрос.* Этот метод участия в переписи удобен для лиц, не допускающих присутствия в доме посторонних людей (счетчиков) и не имеющих возможности выполнить процедуру самоисчисления. Возможную обеспокоенность этих лиц в связи с вопросом о личности интервьюера можно развеять, предоставив в их распоряжение «Инфолинию» с изложением всех сведений о проводимой переписи. Телефонный опрос, характеризующийся отсутствием визуального контакта с собеседником, для одних людей является барьером (поскольку им удобнее физический контакт), а для других – преимуществом, так как они в некотором смысле остаются неизвестными. Одной из предпосылок успешного применения метода телефонного опроса является подготовка такого электронного формуляра, чтобы на его заполнение уходило не более 15 минут. Увеличение продолжительности опроса приводит к утомлению и утрате заинтересованности респондента.

c) *Личная беседа со счетчиком.* Этот метод особо рекомендуется лицам, которые не владеют цифровыми технологиями или являются одиночками, и для них ценен контакт с другим лицом, пусть даже с незнакомцем (счетчиком). Этот метод также подходит для респондентов, которым трудно мобилизоваться для выполнения процедуры самоисчисления: они должны предоставить счетчику информацию вне зависимости от того, является ли момент прихода счетчика удобным для респондента.

2. Достоинства и недостатки многоканального сбора данных: анализ

25. Метод многоканального сбора данных, бесспорно, имеет целый ряд как достоинств, так и недостатков. Он основан на стремлении придать переписи более удобный и менее трудоемкий для респондентов характер. Важнейшим аргументом в пользу многоканальности на основе ИКТ-решений является гарантированное повышение качества данных благодаря алгоритмам, обеспечивающим регулярную проверку правильности данных (подтверждение в процессе заполнения электронного вопросника). Наряду с этим незамедлительная проверка данных открывает возможность для значительно более быстрой обработки результатов переписи по сравнению, например, с применением традиционных письменных методов, т. е. бумажных формуляров. Сбор данных с помощью различных ИКТ-решений снижает расходы на перепись как минимум до того момента, когда появится возможность проводить перепись только на основе регистров или путем онлайн-ового самоисчисления.

26. Недостатком многоканального сбора данных является необходимость в создании затратной информационной системы для обеспечения сбора данных в различных ИКТ-средах с последующим объединением данных, полученных по разным каналам. Тем не менее расходы на создание такой информационной системы будут понесены только один раз; затем ее можно использовать для проведения не только переписей, но и текущих обследований.

C. Система обеспечения качества многоканального сбора данных

27. Система обеспечения качества многоканального сбора данных должна обеспечивать стандартный метод оценки качества данных, собранных по различным каналам; в частности, в этой системе нужно учитывать следующие аспекты:

- a) точность;
- b) своевременность;
- c) расходы на сбор данных;
- d) трудоемкость для респондентов.

28. В системах обеспечения качества, утверждаемых к применению для проведения переписи населения в Польше в 2021 году, следует найти компромисс между вышеупомянутыми аспектами качества. На этапе разработки методики переписи следует провести отбор наиболее важных аспектов и выстроить их иерархию. Согласно Типовой модели производства статистической информации (ТМПСИ), разработка методики сбора данных начинается с выявления потребностей пользователя. Эти потребности и их приоритеты весьма разнообразны. Лица, ответственные за разработку методики сбора данных, учитывают выявленные потребности пользователей, особенно в части точности и своевременности данных, обращая особое внимание на ограничения, включая объем затрат, трудоемкость для респондента, технологии и новаторские методы сбора данных. Важным вопросом является воздействие растущих показателей отсутствия ответов на точность оценочных данных по выборке, а в конечном счете – на точность итоговых переписных данных. Метод сбора данных влияет на типы ошибок (охват, измерение, отсутствие ответов) и на размер ошибок. Таким образом, проверку и оценку качества следует проводить отдельно для каждого канала.

29. Система обеспечения качества многоканального сбора данных определяет требования к уровню качества; она необходима для управления качеством при регулировании многоканального сбора данных. Система обеспечения качества охватывает этап разработки методики переписи и этап сбора данных; для каждого канала она должна содержать следующие аспекты:

- a) проблемные факторы;
- b) планирование качества: цели, параметры качества, типы ошибок, пороговые величины;
- c) стандарты, требования к качеству;
- d) инструменты контроля качества: оценка и измерение;
- e) способы контроля качества;
- f) методы исправления ошибок, отклонения от принятых норм;
- g) стандартный и системный многоэтапный контроль качества собранных данных: мониторинг качества (контрольные перечни, отчеты о качестве, информация об ошибках);
- h) оценка качества, анализ и интерпретация значения переменных;
- i) эффективность процессов сбора данных;
- ii) эффективность мер, принятых после обнаружения ошибок;
- iii) сравнение показателей качества для данных, собранных по разным каналам;
- iv) оценка воздействия методов сбора данных на качество собранных данных.

1. Оценка качества многоканального сбора данных

30. Для получения возможности оценивать качество многоканального сбора данных необходимо составить контрольный перечень многоканального сбора данных, охватывающий параметры качества, показатели и их пороговые значения.

31. Контрольный перечень для методики многоканального сбора данных должен включать в себя пороговые значения отдельных показателей. Поскольку в ходе переписи 2021 года будет применяться метод смешанного сбора данных, для каждого канала необходимо определить минимальные стандарты качества. Роль минимального стандарта качества будет возложена на пороговое значение. Установленный стандарт следует использовать в процессе оценки сбора данных. Стандарты должны быть известны и определены при разработке методики, а также на этапе сбора данных, включая подготовку электронного вопросника и организацию всего процесса в целом. Опираясь на систему обеспечения качества, следует составить вопросник по оценке

качества с учетом установленных стандартов для всех аспектов, которые оказывают влияние на качество собранных данных. Требования, изложенные в этом вопроснике, должны располагаться по степени их значимости. В процессе мониторинга качества многоканального сбора данных следует учесть риски и определить последствия несоблюдения установленных стандартов качества.

D. Система управления сбором данных

32. Целью системы управления многоканальным сбором данных является:

- a) сбор данных требуемого качества в рамках имеющихся сроков и бюджета;
- b) обеспечение высокой эффективности процессов сбора данных;
- c) ориентированность на потребности респондентов;
- d) уделение особого внимания проблемным областям, включая инновации в сфере сбора данных.

33. Вопросы, касающиеся сбора данных, нужно рассматривать с надлежащей тщательностью, а управление должно осуществляться в более широком контексте. При разработке системы управления многоканальным сбором данных нужно учитывать процессы планирования, организации и мониторинга.

34. При анализе процесса планирования были выявлены следующие важные аспекты:

- a) проблемные факторы:
 - i) методологические;
 - ii) технические; применение новых технологий;
 - iii) организационные;
- b) способ реагирования на проблемные факторы; задачи для осуществления;
- c) риски и реагирование на них;
- d) принципы сбора данных: качество, реальные затраты и предполагаемые расходы респондентов; формирование атмосферы сотрудничества и доверия в отношениях между производителями статистических данных и респондентами;
- e) планирование ресурсов;
- f) сметные предположения по строительству хранилища для сбора метаданных о процессе сбора данных. Необходима информация для мониторинга и решения проблем. Информация также требуется для оценки хода работы в соответствии с принятыми сроками и допущениями. Для этого необходима эффективно функционирующая система.

35. При анализе организационного процесса были выявлены следующие важные аспекты:

- a) штат сотрудников, участвующих в процессе сбора данных. Организационная структура подразделений, вовлеченных в процесс сбора данных:
 - i) функции;
 - ii) задачи;
 - iii) степень приоритетности задач;
- b) общение со счетчиками и персоналом, участвующим в процессе сбора данных;
- c) виды и последовательности методов сбора данных;

- d) инструменты для сбора данных: электронные формуляры;
- e) сбор данных из внешних административных систем, которые предназначены для использования в ходе переписи в качестве источника данных;
- f) поставка электронного формуляра с данными из внешних административных систем;
- g) составление для переписи списков в формате «адрес/жилище/лицо»;
- h) правила сбора данных индивидуальными методами;
- i) система управления действиями на местах;
- j) правила отбора счетчиков;
- k) обучение и подготовка вспомогательных материалов для обучения;
- l) стандартизация методов и способы осуществления;
- m) методы устранения ошибок, недостаточный охват;
- n) порядок рассылки напоминаний и действий при низких показателях ответа; подход для применения в случае отсутствия ответа или недостаточного контакта с респондентом;
- o) мероприятия по сбору данных:
 - i) количество и содержание напоминаний о переписи, метод распространения информации: телефон, электронная почта, письмо (традиционное бумажное), посещение;
 - ii) график посещений, беседы счетчика/интервьюера с респондентом (количество, дата);
 - iii) количество повторных посещений, попытки контакта при отсутствии ответа в период планируемого сбора данных;
 - iv) методы взаимодействия с респондентами, с которыми сложно установить контакт;
 - v) обработка отказов в предоставлении информации;
 - vi) коммуникация с респондентами;
 - vii) продолжительность беседы;
 - viii) время сбора данных по каждому каналу;
- p) техническая инфраструктура ИКТ, которая поддерживает процесс сбора данных и адаптирована к управлению сбором данных, управлению действиями счетчиков и интервьюеров, мониторингу, отчетности и сбору данных;
- q) мониторинг сбора данных.

36. При анализе процесса мониторинга сбора данных были выявлены следующие ключевые аспекты:

- a) методы ведения мониторинга (включая отслеживание маршрута перемещений счетчиков с помощью инструментов ГИС);
- b) инструменты мониторинга и отчеты о мониторинге:
 - i) виды отчетов: о результатах работы (отчеты о качестве по каждому из каналов), о ходе работы (показатели выполнения работы интервьюерами);
 - ii) структура отчетов;
- c) анализ отчетов и оценка процессов сбора данных;
- d) решение проблем: процедуры действий в конкретных ситуациях, методы урегулирования проблем, оптимальные решения;
- e) меры по исправлению положения;

- f) оценка эффективности мер по исправлению положения.

Е. Новые решения для сбора данных с использованием передовых технологий, полученных в результате развития ИКТ

37. Задача по разработке новых организационных решений для сбора данных на основе методов САХІ предполагает следующие аспекты:

- a) определение условий, при которых удастся эффективно применять передовые ИКТ-решения для сбора данных;
- b) выявление всех рисков, которые могли бы негативно повлиять на результаты сбора данных;
- c) установление минимального порога для получения результатов переписи, которые можно считать достаточными для признания в качестве достоверных;
- d) спецификация любых мер по исправлению положения, которые следует немедленно принять в случае угрозы, препятствующей продолжению или завершению переписи согласно предусмотренному плану.

38. Кроме того, в ходе организационного планирования важно учитывать:

a) *Последовательность мероприятий и их продолжительность.* Составление графиков осуществления (общего и подробного) является одним из основных мероприятий по подготовке переписи независимо от того, какой метод будет использоваться на этапе сбора данных. Однако при использовании метода многоканального сбора данных на основе САХІ такой график будет намного сложнее, поскольку в нем необходимо учесть все аспекты взаимозависимости как внутри индивидуальных мер, так и между такими мерами (строго в отношении каналов сбора данных), принятыми в ходе переписи. В общем графике указываются все ключевые этапы переписи и основные задачи, которые нужно решать на каждом из них. В нем дается общий обзор хода переписи, и на его основе составляется последовательность мероприятий, которые должны осуществляться эффективно и в оптимальные сроки, исключая чрезмерное нагромождение задач или неоправданные перерывы. Подробный график дает возможность планировать все задачи и подзадачи, предусмотренные внутри основных этапов. Кроме того, во всех графиках необходимо указывать лиц, ответственных за выполнение поставленных задач, особенно если штат лиц, занимающихся проведением переписи, имеет сложную структуру.

b) *Закрепление мероприятий за ответственными лицами по четко установленной процедуре.* Для применения современных технологических решений при проведении сбора данных в рамках переписи населения требуется подготовка целого ряда процессов (нередко состоящих из subprocesses) на основе всеобъемлющего подхода, охватывающего индивидуальные задачи, которые предстоит выполнить на каждом этапе проведения переписи. Чем сложнее и разнообразнее используемая технология, тем больше процессов нужно подготовить. За каждым процессом следует закрепить лицо, ответственное за его выполнение, а также непосредственных исполнителей. Чтобы оптимизировать процесс выполнения задач, нужно провести четкое разделение между обязанностями и самими задачами. Наряду с этим следует разработать подробные процедуры для применения в стандартных ситуациях, а также на случай возникновения нештатных ситуаций. Такие процедуры должны включать в себя описание ситуации и индивидуальные меры, которые следует принять в случае инцидента, а также описание функций и обязанностей.

c) *Конкретный объем людских и материальных ресурсов.* Важную роль в ходе разработки метода сбора данных на основе современных ИТ-технологий, равно как и самой методики проведения переписи населения, приобретает точный расчет необходимого количества сотрудников (включая разделение функций и назначение определенного штата сотрудников для их выполнения), объема ИТ-оборудования для

обеспечения эффективного проведения переписи населения и необходимого объема финансовых ресурсов. Однако подлинным отличием современной, технологичной переписи населения от традиционной является поддержание соответствующих пропорций между расходами на создание каналов сбора данных и остальными расходами. Это особенно важно в случаях, когда при внедрении новых методов требуется также учитывать использование ИТ-инфраструктуры, созданной для данной переписи, которую после завершения переписи можно применять для проведения текущих статистических обследований. В ходе последующих переписей, когда каналы сбора данных уже будут созданы, соотношение расходов снова изменится, поскольку тогда речь уже пойдет не о формировании каналов, а только об их обновлении сообразно технологическому прогрессу в сфере статистики.

d) *Повышение квалификации лиц, которых планируется привлекать на каждом этапе составленной программы переписи.* Изменение технологий сбора данных требует приобретения новых навыков и квалификаций. Поэтому необходимо составить программы профессиональной подготовки сотрудников по проведению переписи, которые будут ориентированы на выполнение определенных переписных функций. Не следует также забывать, что курс такой подготовки должен завершаться экзаменом для проверки степени усвоения участниками новых знаний, поскольку успех проведения переписи будет зависеть от эффективности применения новой технологии.

1. Описание новых организационных решений

39. *Определение вероятного уровня эффективности и сферы применения методов.* В ходе разработки новых технологических решений для их применения при сборе данных, особенно если речь идет о многорежимном проекте, следует оценивать эффективность отдельных каналов. Это – отправная точка планирования дальнейшей организационной работы. Для каждого канала сбора данных есть определенные ограничения, которые нужно принимать во внимание, например:

- a) для каналов CAWI: доступ к Интернету, цифровая изоляция;
- b) для каналов CATI: объем переписного листа (количество вопросов), доступ к телефонной базе данных;
- c) для каналов CAPI: респонденты, проживающие в закрытых, защищенных домах; трудность застать того или иного респондента дома из-за того, что он работает в разное время суток, и т. п.

Учитывая вышеизложенное и целый ряд других факторов – социотехнических, культурных или связанных с положением домохозяйств в зависимости от населенности региона страны, – следует оценить ожидаемый процент ответов через онлайн-процедуру самоисчисления. В свою очередь, выбор каналов CATI и CAPI будет в большей степени зависеть от финансовых условий (объема средств в бюджете переписи) и от числа лиц, которых можно привлечь в качестве телефонных интервьюеров, если опросы по телефону проводятся с использованием ресурсов статистического управления.

40. *Профилирование респондентов для предварительного назначения канала.* Цель этой процедуры – оказать содействие в оценке эффективности сбора данных в рамках отдельных методов, а также в разработке эффективного метода охвата респондентов при надлежащем предварительном оповещении о переписи. Следует отобрать группы респондентов, которые, скорее всего, выберут онлайн-заполнение вопросника. В этом случае данные, полученные из источников информации и касающиеся использования ИТ-услуг, возрастной группы, профессии и т. д., могут оказаться весьма ценными. С другой стороны, для телефонного контакта следует намечать те домохозяйства, которые вряд ли выберут самоисчисление (например, престарелые, хронически больные и т. п.). Остальные лица будут регистрироваться счетчиками переписи, хотя по результатам профилирования уже можно выделить такую группу лиц методом исключения каналов CAWI и CATI.

41. *Получение данных.* Одним из ключевых мероприятий в поддержку сбора данных с применением методов САхI является получение соответствующих данных, например телефонных номеров или адресов электронной почты. Если такая практика не разрешена законодательством, нужно принять меры по обеспечению отступления от закона для целей переписи, учитывая значение проведения переписи и полезность ее результатов для широкого спектра важных получателей информации.

42. *Список мероприятий, которые могут быть выполнены штатными сотрудниками статистического управления, и объем внешнего подряда.* Применение многорежимного метода сбора данных требует привлечения сотрудников с надлежащим уровнем компетентности или же передачи поставленных задач для полного или частичного выполнения на внешнем подряде. Бесспорно, возложение всех функций по проведению переписи на внешнего подрядчика лишит производителей государственной статистики возможности приобретения знаний, навыков и практического опыта. С другой стороны, это усложнит контроль над ходом отдельных процессов и соблюдение конфиденциальности статистики, а также формирование или поддержание общественной репутации официальных статистических данных. Поэтому целесообразно выполнить хотя бы часть задач по проведению переписи внутренними силами, чтобы в будущем обеспечить полную независимость от внешних подрядчиков. Одна из важных задач заключается в том, чтобы задать требуемый уровень качества работы. Кроме того, необходимым условием, не зависящим от объема внешнего подряда, является сохранение ответственности за ход и результаты переписи и выполнение в связи с переписью контрольных и управленческих функций в рамках системы государственной статистики.

43. *Разработка ИТ-системы переписи (ИСП).* ИКТ-решения, основанные на передовых технологиях, особенно при использовании нескольких каналов сбора данных, требуют разработки и формирования соответствующей информационной среды. Зачастую речь идет о необходимости объединения отдельных систем в одну последовательную систему, чтобы обеспечить абсолютные гарантии получения полных, уникальных и высококачественных результатов переписи. Поэтому целесообразно обращать внимание не только на правильность функционирования отдельных систем или приложений, но и на коммуникацию внутри ИСП. Надежно разработанная ИСП позволяет обеспечивать всестороннюю поддержку всех переписных мероприятий, а также, где это возможно, их автоматизацию. В установленных пределах (на основе оценочных данных, введенных в систему) можно выполнять часть работы в ручном режиме или корректировать последствия автоматических операций (например, присвоение адресов счетчикам при использовании метода CAP или замена канала сбора данных для определенного адреса в ходе переписи).

44. *Порядок активации каналов сбора данных.* Это – одно из стратегических решений, от которого зависят конкретные организационные мероприятия. Зная о влиянии методов сбора данных на качество результатов, на ожидаемый спектр отдельных каналов и их предполагаемую эффективность, можно с высокой вероятностью рассчитать оптимальную дату для использования каждого канала. Кроме того, обеспечив профилирование групп респондентов перед их распределением по соответствующим каналам сбора данных, следует определить последовательность методов и наметить дублирование (частичное либо полное) или разделение конкретных используемых методов. Поскольку вторым из наиболее желаемых последствий сбора данных – помимо достижения высокого качества результатов как главного итога переписи – является обеспечение значительного процента ответов по каналам CAWI, онлайн-овое самоисчисление должно быть стартовым каналом, а в начальный период переписи – даже эксклюзивным каналом сбора данных.

45. *Пункты доступа респондентов к Интернету.* Для респондентов, которые желают пройти регистрацию через Интернет, но не имеют доступа к сети, в общественных местах должны быть оборудованы так называемые пункты доступа, гарантирующие безопасность процедуры самоисчисления как для самих респондентов (с точки зрения неприкосновенности их частной жизни), так и для представляемых

ими данных (сохранение статистической конфиденциальности). Наиболее подходящими для этого объектами являются государственные административные здания (муниципалитеты, статистические бюро), учреждения и центры социальной помощи, почтовые отделения, школы и т. п. Принятие мер в этом направлении может обеспечить существенное увеличение доли ответов через Интернет.

46. *Разработка мероприятий по управлению, мониторингу, контролю и отчетности.* В процессе сбора данных по нескольким каналам, который дополнительно затрудняется тем, что при проведении переписи для одного и того же респондента можно менять канал сбора данных (в случае, если первоначальный канал неэффективен), необходимы:

a) создание структур управления каналами с возложением на них конкретных функций, т. е. для систем САТІ и САРІ;

b) тесное сотрудничество между диспетчерами обоих каналов для непрерывного обмена информацией;

c) наличие соответствующих инструментов для контроля межканального потока электронных формуляров;

d) наличие надлежащих инструментов для содействия управлению, мониторингу и контролю. Текущее отслеживание хода переписи по отдельным каналам дает возможность выявлять нежелательные ситуации и незамедлительно устранять их;

e) наличие адекватных механизмов отчетности как для отслеживания прогресса в сборе данных по каждому каналу и в целом, так и для текущей проверки коэффициентов ошибок при проведении переписи на уровне воеводств и в масштабе всей страны.

47. *Формирование среды для сбора данных: популяризация переписи.* Особой задачей в ходе переписи является распространение знаний о переписи и поощрение участия в ней. При умелой популяризации переписи респондентам можно «предписать» такой канал сбора данных, который наиболее выгоден ее организаторам. В то же время нужно уделять особое внимание выбору и содержанию соответствующего медийного средства, ориентированного на целевые группы получателей. Не менее важно разработать информационную кампанию, обеспечивающую соблюдение планов оповещения о соответствующих этапах переписной деятельности и надлежащую интенсивность мероприятий по популяризации переписи. Цель кампании состоит в том, чтобы респондент был воодушевлен популяризацией переписи и не испытывал утомления от ее интенсивности, что может привести к обратному эффекту.

48. *Прогнозирование мероприятий, сопутствующих сбору данных, с учетом потребностей респондентов.* Опыт показывает, что даже весьма тщательная организация работы по проведению переписи не позволяет исключить ряд событий, которые могут произойти при сборе данных на местах. Поэтому целесообразно заранее готовить решения даже для маловероятных случаев, если это не сопряжено с чрезмерными финансовыми расходами. Одним из таких решений является предоставление респонденту возможности зарегистрироваться в системе САТІ по его/ее просьбе, выраженной в телефонном обращении на номер «Инфолинии». Это может потребоваться в ситуации, когда респондент в силу специфики своей работы, семейного положения, намеченной длительной поездки и т. п. не сможет выполнить свою обязанность по участию в переписи в какой-либо другой момент ее проведения. Тогда следует использовать вариант САТІ, т. е. перевести телефонный звонок с «Инфолинии» на канал САТІ.

2. Влияние новых организационных решений на качество результатов

49. На каждом этапе проведения организационных мероприятий независимо от метода проведения переписи последствия таких мероприятий нужно проверять и реагировать соответствующим образом, если обнаружится, что они приводят к

нежелательным последствиям. Это особенно важно в условиях применения новых методов, таких как методы сбора данных САХІ.

50. Воздействуя на респондентов, методы сбора данных САХІ могут косвенно влиять на качество результатов переписи. При этом имеется в виду не система планирования качества при многоканальном сборе данных, которая описана в пункте С «Система обеспечения качества при многоканальном сборе данных»; речь идет о потенциальном получении респондентами информации о новом порядке сбора данных. Людям свойственно проявлять недоверие к изменениям. Поскольку общественное восприятие – это один из определяющих факторов успеха переписи, крайне важно уделять особое внимание реакции со стороны респондентов. Средствами выражения их реакции могут являться:

- a) показатели ответов по отдельным каналам;
- b) обратная связь по «Инфолинии»;
- c) ответы на телефонные обращения интервьюеров и посещения счетчиков переписи;
- d) частотность посещения мест, где установлены компьютеры для самоисчисления;
- e) тональность откликов в прессе и в других широко доступных и используемых средствах массовой информации;
- f) позитивное или негативное отношение к официальной статистике.

51. Каждый из вышеупомянутых факторов следует тщательно проанализировать, чтобы затем сформулировать выводы и разработать стратегии для поддержания нужных тенденций и противодействия негативным ситуациям. Особенно важна при этом роль групп, которые ответственны за популяризацию переписи и действуют в консультации с группами по организационным вопросам.

V. Заключение

52. Казалось бы, в эпоху информационного общества будет ошибкой не воспользоваться той возможностью, которую предоставляют современные технологии. Тем не менее для повышения качества переписи населения требуется дальнейшая исследовательская работа.

53. Мир ИКТ переживает период бурного развития. То, что сегодня является инновацией, через несколько лет уже устареет. Поэтому постоянная работа над повышением качества переписи позволит сформировать механизмы и модели, которые можно будет использовать для обеспечения дальнейшего роста качества переписей населения в будущем, когда появятся новые инструменты и источники данных для проведения таких переписей.