Организация Объединенных Наций ECE/CES/GE.41/2018/19

Экономический и Социальный Совет

Distr.: General
18 July 2018
Russian
Original: English

Европейская экономическая комиссия

Конференция европейских статистиков

Группа экспертов по переписям населения и жилищного фонда

Двадцатое совещание
Женева, 26–28 сентября 2018 года
Пункт 5 предварительной повестки дня
Технология

Подготовка к переписи населения и жилищного фонда 2020 года

Записка Национального института статистики и географии (НИСГ) Мексики

Резюме

Национальный институт статистики и географии (НИСГ) проводит в Мексике перепись населения и жилищного фонда. В целях содействия развитию Мексики НИСГ, входящий в Национальную систему статистической и географической информации, должен предоставлять обществу и государству качественную, актуальную, четкую и своевременную информацию.

Цель настоящей записки заключается в информировании о компьютерных средствах и технологиях, которые применяются или которые планируется применять в ходе работы, связанной с проведением переписи населения и жилищного фонда 2020 года, начиная с ее разработки, планирования и проведения и кончая представлением ее результатов.

GE.18-11933 (R) 270718 270718

1811933

Просьба отправить на вторичную переработку

I. Введение

1. На основании положений статьи 26, раздела B Конституции Мексиканских Соединенных Штатов и во исполнение полномочий, которыми наделяет его Закон о Национальной системе статистической и географической информации, Национальный институт статистики и географии (НИСГ) проводит переписи населения и жилищного фонда. В целях содействия развитию Мексики НИСГ, входящий в Национальную систему статистической и географической информации, должен предоставлять обществу и государству качественную, актуальную, четкую и своевременную информацию.

2. Цель настоящей записки заключается в информировании о компьютерных средствах и технологиях, которые применяются или которые планируется применять в ходе работы, связанной с проведением переписи населения и жилищного фонда 2020 года, начиная с ее разработки, планирования и проведения и кончая представлением ее результатов.

3. Определение концептуальных, технических, методологических аспектов и соответствующих мер административной поддержки, а также осуществление этапов сбора и обработки информации происходит в соответствии с требованиями, установленными на этапах разработки обследования, и завершается на этапе представления результатов, в рамках которого распространяются продукты и услуги со статистическими данными.

4. Для повышения эффективности и контроля в ходе описанного выше процесса в рамках переписи населения и жилищного фонда 2020 года (переписи 2020 года) ИНЕГИ планирует широко использовать информационно-коммуникационные технологии, поскольку он имеет предыдущий опыт статистических мероприятий; помимо использования мобильных устройств, осуществляется стратегия оперативных процедур с использованием ряда компьютерных инструментов для субъектов оперативной структуры, т. е. аналитиков и лиц, принимающих решения, в целях автоматизации ряда процессов на этапах планирования, проведения переписи, последующей работы и обработки информации.

5. Работа по общему планированию переписи населения и жилищного фонда 2020 года началась в 2017 году с проведения общественных консультаций в период с 21 августа по 30 ноября, что позволило определить информационные потребности пользователей, а также включала в себя другие мероприятия, такие как проведение в 2017 году тестирования оперативной стратегии, целью которой являлось внедрение технологий и компьютерной техники в процесс сбора информации.

II. Основные характеристики переписи населения и жилищного фонда 2020 года

6. Переписные данные весьма полезны, поскольку предоставляемая ими информация позволяет проанализировать произошедшие в Мексике изменения и дает возможность выявить социальную отсталость, условия жизни уязвимых групп, информацию о жилье, образовании и здравоохранении, а также потребности в водопроводной воде, электричестве и канализации, помимо других аспектов, которые имеют важное значение для принятия решений в политических вопросах и для подготовки планов и программ, направленных на борьбу с нищетой и улучшение условий жизни населения.

7. Та информация, которая станет предметом переписи, должна удовлетворить информационные потребности трех уровней государственного управления и различных слоев населения. Кроме того, в результате должна быть получена информация, необходимая для отчета о прогрессе Мексики на пути достижения Целей устойчивого развития и выполнения государственных планов и программ, а также для покрытия потребностей в информации, касающейся исполнения мер, предусмотренных законами, регулирующими развитие страны. Важно отметить, что в

процессе определения круга тем учитывались международные принципы и рекомендации, которые позволяют обеспечить сопоставимость данных о ситуации в этой стране с данными других стран.

III. Методологические основы

8. Как и предыдущие переписи, перепись 2020 года будет переписью де-юре, т. е. подсчет населения будет вестись по месту обычного жительства.

9. Единицами наблюдения этой переписи являются постоянно проживающие в стране лица, частные и коллективные жилища.

10. Основным респондентом будет глава жилища или, в его отсутствие, лицо 18 лет и старше, которое является постоянным жителем этого жилища и обладает информацией о жилище и проживающих в нем лицах.

11. Перепись будет проведена в марте 2020 года.

IV. Метод сбора данных

12. Сбор информации будет главным образом осуществляться с помощью личного опроса с использованием компьютерных приложений на электронных устройствах. Кроме того, при необходимости с помощью интернета и телефонных опросов будет проводиться самоисчисление.

13. Если говорить об инструментах сбора информации, то для переписи населения и жилищного фонда будет использовано 2 переписных листа: короткая форма, которая будет применяться при сплошной регистрации и содержать около 30 вопросов, и длинная форма для вероятностной выборки, которая будет содержать примерно 75 вопросов.

14. Кроме того, в рамках переписи 2020 года будет собираться информация о городской среде кварталов в населенных пунктах, насчитывающих 5 000 и более жителей, а также об инфраструктуре и социально-экономических характеристиках населенных пунктов, насчитывающих менее 5 000 жителей. Планируется также провести перепись социального жилья (CAAS), которая отражает характеристики проживающих и работающих там людей, а также зданий и предоставляемых услуг.

15. Для правильной географической привязки информации, собранной в ходе переписи населения 2020 года, счетчик будет создавать перечень жилищ и наклеивать коды быстрого реагирования (QR-кодами). QR-код, а также географическое местоположение жилища, полученное с помощью Глобальной системы позиционирования (GPS), позволит обновить Национальный регистр жилищного фонда и приступить к статистической регистрации зданий страны.

16. Используемые устройства должны иметь следующие характеристики: сенсорный экран размером 6,4"–7" с разрешением 1 280 x 800 мегапикселей, операционная система Android 6 или выше; минимальные технические требования: 32 Гб внутренней памяти, 2 Гб RAM, минимальная длительность работы батареи – 8 часов в день, цифровая камера с разрешением 2 мегапикселя и возможность подключения USB OTG (On-The-Go). Другие необходимые элементы: GPS, порты для сим-карты и SD-карты, а также возможность подключения Wi-Fi и Bluetooth. Для целей спутниковой навигации и распознавания местоположения в большинстве случаев используются системы GPS и ГЛОНАСС, используемые устройства поддерживают технологии GPS и A-GPS.

V. Использование компьютерных инструментов на разных этапах переписи 2020 года

A. Концептуальная основа

17. Здесь рассматриваются мероприятия, в ходе которых выявляются потребности в информации, используемые для определения и выработки концептуальной основы привязки данных, схемы представления результатов, инструментов их получения и критериев подтверждения, а также возможностей устранения несоответствий.

18. В целях реализации задач по разработке концептуальной основы были разработаны два веб-приложения: первое из них, «Общественные консультации с пользователями переписи 2020 года», призвано облегчить задачу выявления информационных потребностей пользователей. Оно позволяет пользователю в простой и систематизированной форме изложить цель и значимость его потребностей в информации, а ответственным сотрудникам Института – проследить выполнение каждого поступившего от населения запроса. Для этой системы были использованы платформы Visual Studio .Net и JavaScript AngularJS, а для управления базой данных использовалась система Oracle 12c.

19. Второе приложение, «Концептуальная инфраструктура переписи и подсчетов», призвано стать хранилищем концептуальных основ для переписей населения и жилищного фонда, с тем чтобы простыми методами обеспечить обзор и обновление концептуальной и оперативной информации о переменных, которые собираются в переписных листах, а также обзор концептуальных изменений переменных во времени. Для разработки этого приложения также использовалась программа Visual Studio .Net 2013 и база данных на основе SQL Server 2016.

B. Оперативная структура

20. На этом этапе определяются процедуры и технические схемы деятельности по сбору данных, а также административные и организационные аспекты и вопросы, связанные с контролем и последующими действиями. Начиная с переписи 1990 года НИСГ накапливает опыт в разработке систем для выполнения этих задач, и в 2020 году предлагается использовать три системы, описанные ниже.

21. Система конфигурации зон ответственности позволяет равномерно распределить рабочую нагрузку в рамках оперативной структуры и автоматически выдает группы или сегменты городских кварталов или сельских населенных пунктов, сформированные, помимо прочих факторов, на основе номеров их жилищ и коммуникационных маршрутов, образуя таким образом зоны оперативной ответственности и порядок проведения переписи с помощью алгоритма имитации отжига. Она позволяет корректировать конфигурации вручную с использованием полевого опыта и знаний оперативного персонала. Эта система была разработана с помощью Delphi XE4 и сервера данных MS SQL SERVER.

22. Система обучения, разработанная на языке JavaScript на основе стандарта Ecmascript6 и с использованием базы данных SQLite, направлена на обеспечение оперативного персонала необходимыми навыками и инструментами для выполнения своих функций с помощью учебных модулей, включающих учебные материалы, возможности дистанционных рекомендаций и отправки выполненных заданий, и контроля за продвижением обучения. Для разработки приложения для Windows использовалась система Electron, а для версии для операционной системы Android – мобильная среда Apache Cordova.

23. С другой стороны, система Opera, которая используется на нескольких этапах процесса подготовки информации, позволяет систематизировать деятельность по найму персонала, начиная с регистрации заявителей через Интернет и заканчивая их наймом, а также отслеживать прогресс, достигнутый в разработке оперативной структуры, в целях мониторинга оперативной логистики путем отслеживания

материалов, которые необходимы для полевых и управленческих мероприятий (транспортные средства, компьютеры, вопросники, брошюры и т. д.). Для разработки этой системы использовалась программа Visual Studio.Net, а для управления базой данных использовалась среда AngularJS JavaScript с Oracle 12c.

C. Сбор информации

24. Этот этап охватывает комплекс мероприятий, направленных на получение данных по каждому элементу населения на основе программы и установленных процедур работы с использованием оперативной структуры и механизмов контроля, обеспечивающих эффективность каждой из мер. Это предполагает реализацию схем сбора предварительно определенных данных, включая подготовку и распространение вспомогательных материалов (карт, пособий, инструкций и каталогов), интеграцию людских ресурсов, коммуникацию и согласованные действия в рамках подробной программы мероприятий, организационную структуру и механизмы контроля, обеспечивающие эффективность каждой из мер.

25. Для этого этапа в настоящее время разрабатывается система «Администратор переписи», в которой планируется предусмотреть функции визуализации и регулирования рабочей нагрузки и при необходимости ее перераспределения; в ней имеется «модуль учета», используемый для регистрации данных перечня зданий, вопросника по городской среде, вопросника по местности и вопросника по заселенным жилищным единицам; она дает возможность получения конкретных отчетов о прогрессе и охвате и включает в себя «модуль контроля», в котором реализована функция зональной выборки с целью обзора классификации зданий и жилищных условий; она также позволяет произвольно выбирать жилища для повторных опросов руководителем счетчиков и проверять качество информации, регистрируемой счетчиком с помощью механизмов проверки качества.

26. Система «Администратор переписи» включает «картографический модуль», цель которого заключается в том, чтобы регистрировать обновленную картографическую информацию на уровне квартала, населенного пункта и улицы, позволяя оцифровывать полигоны в новых зонах, устранять те из них, которые больше не существуют, а также наглядно отражать изменения, такие как слияние и разделение зон, благодаря чему происходит оперативное обновление картографии и осуществляется планирование сбора информации.

27. Для переписи 2020 года будет введен «модуль для оперативной проверки охвата», целью которого является облегчение деятельности по обзору, осуществляемой независимой от счетчиков оперативной структурой, обеспечивающей надзор за определенной выборкой зон, и в случае необходимости в целях улучшения охвата переписи населения и жилищного фонда будет собираться недостающая демографическая информация. С другой стороны, в ней также будут содержатся адаптированные необходимые компьютерные программы для проверочного обследования и модуль для сопоставления данных, полученных в ходе переписи и проверочного обследования.

28. Система «Администратор переписи» позволяет осуществлять передачу информации от счетчиков к их руководителю, а затем в центральную базу данных, управляемую системой Opera, где имеется модуль, позволяющий контролировать объединение данных, собранных каждым оперативным субъектом, а также проводить анализ и последующую обработку определенных показателей охвата, скорости и производительности переписи; это имеет важное значение для принятия управленческих решений в ходе оперативного подсчета. «Администратор переписи» будет использоваться на этапе проведения переписи, а также на этапах проверки и послепереписного обследования и был разработан с помощью программы Embarcadero RAD Studio 10.2.2 Tokyo в версии Architect, которая представляет собой стандартный инструмент разработки для нескольких платформ и устройств с операционными системами Windows/Android/iOS, включающий в себя модуль FireMonkey,

позволяющий создавать один и тот же программный код для платформ Windows и Android.

29. Кроме того, в ходе следующей переписи населения планируется значительно расширить самоисчисление через Интернет и телефонный центр, поэтому населению будет доступно приложение, которое позволяет ответить на вопросы в краткой форме, и следует отметить, что эта система будет адаптирована для сбора информации в коллективном жилье и среди сотрудников дипломатической службы Мексики. Инструмент для самоисчисления был разработан с помощью Visual Studio 2017 и базы данных Oracle 12c.

D. Использование QR-кодов

30. Для обеспечения надлежащего охвата жилищ была разработан уникальная наклейка с QR-кодом (код быстрого реагирования), которая будет наклеиваться на все здания. Этот инструмент будет служить в качестве основы для полевого контроля и мероприятий по проверке.

31. Кроме того, он будет использован в процессе самоисчисления с помощью интернета и телефона, т. е. люди смогут ответить на вопросы переписного листа без помощи счетчика, но будет осуществляться контроль охвата и правильной географической привязки собранной информации, поскольку счетчик составит перечень жилищ и наклеит QR-коды, которые в свою очередь будут увязаны с распечатанными приглашениями, содержащими уникальное имя пользователя и пароль для каждого жилища, а также необходимые инструкции для регистрации через Интернет и телефонный номер для консультаций, причем этот пользователь будет одним и тем же для обоих вариантов.

E. Способы применения GPS

32. QR-код, а также географическое местоположение жилища, полученное с помощью Глобальной системы позиционирования (GPS), позволит обновить Национальный регистр жилищного фонда и приступить к статистической регистрации зданий страны. В городских районах планируется установить географические координаты каждого жилища, в то время как в сельской местности это будет сделано в определенном месте населенного пункта.

F. Обработка информации

33. В ходе этого процесса ведется подготовка файлов с данными, обеспечивается их согласование и структурное соответствие целям статистического применения; в случае сбора информации через мобильные устройства она включает кодирование, подтверждение, обновление базы данных географических обозначений и публикацию базы данных. Определяются также стратегии обработки информации, включая разработку соответствующих систем кодирования, подтверждения и использования результатов с инструментами контроля качества для каждой из них.

34. Для обработки будет использована система обработки, отслеживания и контроля, которая включает в себя механизмы автоматического кодирования, работающие с помощью алгоритмов, определяющих текст открытых переменных и присваивающих ключ к заранее определенным каталогам, после чего распределяется рабочая нагрузка автоматизированного кодирования. После того, как присваивается ключ, проводится подтверждение записей. Эта система также включает тестирование критериев подтверждения, с тем чтобы обеспечить правильное применение концептуальных принципов.

35. В ходе каждого этапа, описанного выше, это приложение предоставляет в режиме реального времени отчеты о ходе процессов и качестве информации и создает журнал учета изменений в данных. Параллельно с этим осуществляется контроль

присвоения географических обозначений в рамках национальной геостатистической системы[1].

36. И наконец, с ее помощью можно получить доступ к базе данных мероприятий по публикации данных о деятельности, которая представляет собой статистический обзор изменений социально-демографических показателей на государственном и муниципальном уровнях и в разбивке по размеру населенного пункта. Для этой деятельности будет использоваться система «Либера», которая позволяет проводить контрольный обзор по каждому показателю, а также служит хранилищем для документации, подтверждающей изменения некоторых статистических данных. «Либера» и система мониторинга и контроля обработки данных были также разработаны с помощью Visual Studio .Net 2017 и Oracle 12c.

G. Представление результатов

37. Эта деятельность предусматривает представление результатов, ранее определенных в концептуальной структуре, завершая работу над продуктами редакционной программы таким образом, чтобы различные потребности пользователей удовлетворялись наилучшим образом и при минимальных затратах.

38. Традиционно публиковались печатные данные в виде таблиц, однако начиная с переписи 2010 года они больше не публикуются, а вместо этого пользователям предоставляются электронные таблицы и таким образом преодолеваются ограничения бумажного документа и общественность обеспечивается большим объемом обработанных данных.

39. Кроме того, после этого события были также введены следующие три онлайн-системы, позволяющие опытным пользователям получить доступ к более подробной информации: система доступа к информации, полученной в результате переписи, которая основана на институциональной платформе MxSIG, которая предусматривает использование надежного и бесплатного программного обеспечения и к элементам с открытым исходным кодом, таким как Tomcat для публикации услуг, «Проект P» для статистического анализа, PostgreSQL для хранения и PostGIS для управления геопространственными данными, которая позволяет связать полученную в ходе переписи статистическую информацию с географическим пространством, к которому она принадлежит, облегчить толкование социально-демографических явлений путем создания тематических карт; система консультаций по вопросам территориальной интеграции и местности, разработанная с помощью Visual Studio .Net 2017 and Oracle 12c и позволяющая получить динамическую информацию об изменениях каждого населенного пункта страны; и социально-демографическая панорама Мексики, который представляет собой динамичную систему доступа к информации, которая объединяет и обобщает соответствующие данные, описывающие основные демографические, социальные и экономические характеристики населения и жилого фонда Мексики.

40. Для переписи 2020 года планируется обновление этих систем доступа к данным, после чего будет опубликована всеобъемлющая система доступа к данным переписи с целью предоставления пользователям максимально возможного объема информации, касающейся включенных в охват проектов переписи переменных, начиная от этапа концептуализации и заканчивая полученными статистическими результатами, включая описание операционной структуры, процесса обработки информации и ее оценки.

[1] INEGI. Marco Geoestadístico Nacional. http://www.inegi.org.mx/geo/contenidos/geoestadistica/m_geoestadistico.aspx.

VI. Выводы по вопросам технологической архитектуры и информационной безопасности

41. В ходе тестирования переписи 2020 года была реализована стратегия передачи информации, где используются упакованные файлы с паролями безопасности и внутренняя интеграция с базами данных, закодированными с помощью алгоритма кодирования по методу AES256 (расширенный стандарт шифрования с 256-битным ключом), который является одним из наиболее безопасных и широко используемых открытых алгоритмов, используемых, в частности, для целей военной безопасности. Эти пакеты закодированных данных помещаются в память USB с помощью функции передачи информации, содержащейся в «Администраторе переписи», и отсылаются с помощью модуля интеграции информации системы Opera, который размещен на безопасном институциональном веб-сайте, а затем переносятся на серверы массового хранения данных и записываются на сервер Oracle центральных баз данных.

42. Компьютерные телекоммуникационные услуги, серверы корпоративной сети и приложений, базы данных и массовое хранилище (NAS) внутри корпоративной сети (DMZ и Интранет) предоставляются за счет областей компьютерной инфраструктуры, которые остаются доступными этим серверам. Система доступности услуг управляется с помощью механизма распределения нагрузки запросов на базе веб-домена Opera, который равномерно распределяет каждый запрос на обработку данных среди серверов приложений и баз данных. Различные стандарты безопасности и тесты доступности услуг позволяют добиться планируемой стабильности всего комплекса услуг.

43. Для разработки веб-приложений используются современные технологические платформы, причем иногда осуществляется их полное перепрограммирование, что способствует повышению результативности их работы. Основным инструментом разработки является система Visual Studio .NET, которая представляет собой полный набор инструментов для создания веб-приложений, в дополнение к системе AngularJS, которая обеспечивает технологии для создания одностраничных приложений, что улучшает производительность, так как поступающие на сервер запросы становятся менее сложными, поскольку требуют доступа только к определенной части страницы. С другой стороны, благодаря использованию JWT (JSON Web Token), обеспечивается более безопасная коммуникация и облегчается управление пользователями с помощью распределения ролей пользователей и разрешений. Для управления базой данной используется Oracle 12c, что дает более высокий уровень безопасности и доступности информации.

44. Что касается компьютерного инструментария для сбора информации на мобильных компьютерных устройствах, то в качестве языков программирования главным образом используются Javascript и HTML5. Применяемые технологии позволяют задействовать цифровую картографию в формате GeoJSON с использованием системы работы с картами Leaflet, в которой предусматриваются особые функции для мобильных устройств. Стандартные программы используются для формирования вопросников, загружаемых в SQLite, облегченную мобильную базу данных и через сервер приложений, программируемый с помощью Delphi Tokyo of Embarcadero Rad Studio; они форматируются для Android и для Windows, в результате чего обеспечиваются многоплатформенные интерактивные интерфейсы. Кроме того, для улучшения производительности и внешнего вида этих приложений используются такие системы, как jQuery и Bootstrap.

Литература

Cámara de Diputados del H. Congreso de la Unión 2017 - Constitución Política de los Estados Unidos Mexicanos

Cámara de Diputados del H. Congreso de la Unión 2008 - Ley del Sistema Nacional de Información Estadística y Geográfica

Instituto Nacional de Estadística y Geografía (INEGI) 2017 - Aspectos Generales del Censo de Población y Vivienda 2020. Consulta Pública 2017. Aguascalientes, México, INEGI. (Internal document)

Instituto Nacional de Estadística y Geografía (INEGI) 2010 - Censo de Población y Vivienda 2010. Aguascalientes, México, INEGI