



SOLO PARA PARTICIPANTES
DOCUMENTO DE REFERENCIA

LC/CEA.11/DDR/1
18 de noviembre de 2021

ORIGINAL: ESPAÑOL

21-00888

Undécima Reunión de la Conferencia Estadística de las Américas
de la Comisión Económica para América Latina y el Caribe

Reunión virtual, 23 a 25 de noviembre de 2021

LINEAMIENTOS GENERALES PARA LA CAPTURA DE DATOS CENSALES

Revisión de métodos con miras a la ronda de censos 2020



Este documento se elaboró en el marco de las actividades del Programa Bianual 2020-2021 del Grupo de Trabajo para la elaboración del documento sobre Lineamientos generales para la captura de datos censales con miras a la ronda de censos 2020, de la Conferencia Estadística de las Américas de la Comisión Económica para América Latina y el Caribe, CEPAL, coordinado por el Instituto Nacional de Estadísticas de Chile (INE), con el especial apoyo de los Institutos Nacionales de Estadísticas del Brasil, Colombia, Costa Rica, el Ecuador, México y la República Bolivariana de Venezuela, junto con el Centro Latinoamericano y Caribeño de Demografía (CELADE) - División de Población (CEPAL), en su calidad de secretaría técnica. El documento se elaboró bajo la coordinación general de Lorena Espinosa, Coordinadora Nacional del Censo del INE Chile, y de Fabiana Del Popolo, Jefa del Área de Demografía e Información sobre Población del CELADE; la supervisión técnica estuvo a cargo de Daniel Allende, del INE Chile, y de Lenin Aguinaga, del CELADE. En la preparación del documento participaron: Lenin Aguinaga, Carolina Cavada, Amalia Plana y Ari Silva del CELADE y Daniel Allende y Grace Palma, del INE Chile. Asimismo, se contó con la colaboración de Fernando Baeza y Olga Barquero del INE Chile.

Las opiniones expresadas en este documento, que no ha sido sometido a revisión editorial, son de exclusiva responsabilidad de los autores y pueden no coincidir con las de la Organización o las de los países que representa.

Índice

Siglas y acrónimos	7
Resumen	9
Introducción	11
I. Contenido y uso de los lineamientos	15
A. Objetivo y estructura	15
B. Términos o conceptos utilizados	17
II. Antecedentes sobre la captura de datos censales en la región	20
III. Métodos para la captura de datos censales	31
A. Captura digital durante el operativo	34
1. Dispositivos Móviles de Captura (DMC)	35
2. Transmisión de datos	37
3. Dato primario	37
4. Aplicativos de Gestión	37
B. eCenso (Autoenumeración o autoempadronamiento vía web)	38
1. Diseño del formulario	39
2. Seguridad	40
3. Conectividad	40
4. Estrategia de contacto a la población	40
C. Captura desde papel	40
1. Digitación	41
2. Reconocimiento óptico (Escaneo)	43
D. Ventajas y desventajas de los diferentes métodos de captura	47

E. Combinación de métodos de captura	49
1. Desarrollo	50
2. Parcialidad (mode effect o bias).....	50
3. Operacionalización	50
4. Costo	51
5. Integración de datos	51
F. Métodos complementarios de recolección	51
IV. Matriz de decisión–criterios de evaluación	53
A. Criterios de evaluación.....	54
B. Evaluación cualitativa	56
C. Evaluación cuantitativa	59
V. Monitoreo del proceso de recolección censal	63
A. Monitoreo de la transmisión e integración de los datos	66
B. Monitoreo de indicadores del operativo	68
1. Indicadores de gestión operativa	68
2. Visualización geoespacial	69
C. Indicadores de seguimiento demográficos.....	70
VI. Consideraciones de gestión y planificación vinculadas con la captura de datos censales	71
A. Consideraciones para la gestión institucional y del censo al incorporar un método de captura electrónico.....	72
1. Estructura organizacional	73
2. Gestión del cambio	74
B. Impacto del método de captura de datos en el diseño de los procesos censales.	74
1. Calendario censal.....	75
2. Habilidades informáticas y desarrollo de capital humano	75
C. Gestión de la inversión en Tecnologías de Información	76
D. Consideraciones relacionadas con adquisiciones, desarrollos y externalizaciones	78
E. Plan de pruebas.....	78
VII. Conclusiones.....	81
Bibliografía.....	83
Anexos	85
Anexo 1	86
Anexo 2	88
Cuadros	
Cuadro 1 Matriz de métodos de recolección censal	18
Cuadro 2 Censos de población y vivienda en América Latina, según tipo de censo	22
Cuadro 3 Censos de población y vivienda en América Latina, según método principal y secundario de captura.....	23
Cuadro 4 Método de enumeración y momento de captura de datos.....	32
Cuadro 5 Características del desarrollo del aplicativo para DMC según países seleccionados....	36
Cuadro 6 Uso de doble digitación	42
Cuadro 7 Porcentaje de población cubierta según el número de personas por cuestionario, en censos seleccionados de la década de 2010	46
Cuadro 8 Páginas por cuestionario censal según módulo de preguntas y total, en censos seleccionados de la década de 2010.....	46

Cuadro 9	Características relativas entre los métodos de captura	47
Cuadro 10	Criterios de evaluación de los métodos de captura de datos	55
Cuadro 11	Ejemplo de criterios a considerar	56
Cuadro 12	Ejemplo de planilla para cálculo de costos fijos y variable según método de captura	56
Cuadro 13	Ejemplo de matriz cualitativa para elegir métodos de captura de datos censales	57
Cuadro 14	Ejemplo de matriz de decisión basada en 5 criterios globales	59
Cuadro 15	Matriz de decisión con ponderadores	61
Cuadro 16	Funciones mínimas que debiera contener un SIGO para hacer un censo con método de captura mixto	65
Cuadro A.1	Ejemplo de Matriz cualitativa para Censo de Perú 2017	86
Gráficos		
Gráfico 1	Métodos de captura utilizados en la década de 2010 y los que se planifican para la década de 2020 en América Latina	24
Gráfico 2	Sistemas operativos móviles que han planificado usar los países en la ronda censal de 2020	25
Gráfico 3	Cantidad de países que han declarado que van a utilizar sistemas operativos Android y otro sistema operativo móvil en los censos de la década 2020	25
Gráfico 4	Lenguaje de programación para la captura con DMC que se usaría en censos de la década 2020	26
Gráfico 5	Tipo de software para la captura de datos en los censos de la década de 2020	26
Gráfico 6	Plataformas tecnológicas para almacenamiento de datos mediante captura vía web en censos de la década de 2020	27
Gráfico 7	Cómo se desarrollará el sistema automatizado de monitoreo y control de cobertura en censos de la década del 2020	28
Gráfico 8	En qué se basa el sistema automatizado de monitoreo y control de cobertura en censos de la década del 2020	28
Gráfico 9	Sistemas de almacenamiento de datos en los censos de la década 2020	29
Gráfico 10	Software de procesamiento de datos en los censos de la década 2020	29
Recuadros		
Recuadro 1	Uso de registros administrativos en el Censo de Colombia 2018	33
Recuadro 2	Características de los resultados del eCenso 2018 de Colombia	39
Recuadro 3	Ecuador, fases y actividades del próximo censo de población y vivienda	44
Recuadro 4	Síntesis del proceso de decisión en la elección de uso de DMC como método para la recolección censal para el censo 2021, Costa Rica	54
Recuadro 5	Forma de monitoreo de la recolección con dispositivos móviles de captura, Brasil, Censo de 2010	65
Recuadro 6	Control y monitoreo de transmisión de datos, México, censo 2020	67
Recuadro 7	Monitoreo de la recolección con dispositivos móviles de captura, México, Censo 2020	68
Recuadro 8	Control de cobertura, Colombia, censo 2018	69
Recuadro 9	República Bolivariana de Venezuela 2011	70
Recuadro 10	Recomendación a otros países que deseen incorporar el DMC como método de captura Brasil censo 2021	73
Recuadro 11	Recomendaciones a países que deseen incorporar el DMC como método de captura, México censo 2020.	77
Recuadro 12	Pruebas realizadas durante la preparación de la recolección con DMC, Departamento Administrativo Nacional de Estadística (DANE), censo 2018	79
Recuadro 13	Prueba de funcionalidad de dispositivos móviles, México, censo 2010	79

Diagramas

Diagrama 1	Factores involucrados en la decisión sobre la tecnología a usar para la captura de datos censales	34
Diagrama 2	Plataforma Opera del INEGI	38
Diagrama 3	Modelos de recolecciones de datos censales	49

Siglas y acrónimos

ALC	América Latina y el Caribe
CANCEIS	Canadian Census Edit and Imputation System
CATI	Computer assisted telephone interview
CEA	Conferencia Estadística de las Américas
CELADE	Centro Latinoamericano y Caribeño de Demografía (CELADE)-División de Población de la Comisión Económica para América Latina y el Caribe (CEPAL)
CEPAL	Comisión Económica para América Latina y el Caribe
CRTC de Venezuela	Centros de Recolección y Transmisión Censal, del INE de la República Bolivariana de Venezuela
CSPro	Census Processing System
DANE	Departamento Administrativo Nacional de Estadística
DMC	Dispositivo móvil para la captura de datos
eCenso	Autoempadronamiento utilizando Internet (Web)
GPS	Sistema de Posicionamiento Global
GTC CEA-CEPAL	Grupo de Trabajo Censos de la Conferencia Estadística de las Américas (CEA) de la CEPAL
IBGE	Instituto Brasileño de Geografía y Estadística
INE	El nombre genérico de INE hace referencia a institutos, oficinas y direcciones generales o nacionales de estadísticas en la región.
INEC	Instituto Nacional de Estadística y Censos del Ecuador
INEGI	Instituto Nacional de Estadística y Geografía de México
OPERA	Sistema para el seguimiento a operativos censales del INEGI
REDATAM	Software de procesamiento, análisis y disseminación de microdatos censales, encuestas y otras fuentes de datos.
RRAA	Registros administrativos
SEN	Sistema Estadístico Nacional
SAMC	Sistemas automatizados de monitoreo y control de cobertura
SIG	Sistema de Información Geográfica
SIGO	Sistema Integral de Gestión Operativa
SIGPC	Sistema de Gerenciamiento del Puesto de Recolección
SPOT	Dispositivo con tecnología GPS, cuya funcionalidad consiste en enviar reportes sobre la posición geográfica del personal asignado al método de rutas del DANE.
STATCAN	Statistics Canada
TCO	Costo total de propiedad

TI	Tecnologías de la Información
TIC	Tecnologías de la Información y la Comunicación
UNSD	United Nations Statistics Division
USCB	United States Census Bureau

Resumen

El presente documento se enfoca en la mejora de los procesos asociados a la recolección de datos censales mediante la incorporación de mayor tecnología en el método de captura. En los países de América Latina y el Caribe, la tendencia ha sido, hasta ahora, transitar hacia censos con metodología de derecho, usando dispositivos móviles de captura (DMC), lo cual se ha visto impulsado en un contexto de pandemia ante los desafíos planteados para los Institutos Nacionales de Estadísticas¹ (INE) de la región en la recolección de datos. Otra tendencia, ha sido implementar métodos mixtos con el uso del autoempadronamiento por medio del censo electrónico o virtual, denominado eCenso, como complemento al método de captura principal, sea este en papel o con DMC. Por último, los países que planean usar cuestionarios de papel avanzan también hacia el uso de escáner para la captura óptica de los datos censales.

En cuanto a la toma de decisiones para elegir el método de captura más adecuado para un censo exitoso, se propone un modelo compuesto por matrices cualitativas y cuantitativas que contribuye a la evaluación de cada método, basándose en los criterios que son más relevantes para cada país. Asimismo, se deben considerar las características de cada método, sus ventajas y desventajas, garantizando siempre las condiciones para su correcta implementación, evaluación y consistencia entre los procesos, tanto técnicos como de gestión y de administración.

Cuando los países determinan más de un método de recolección (mixto), en modo secuencial o simultáneo, se complejiza el proceso censal y surgen diversos tipos de errores originados desde cada uno de los métodos. Indiscutiblemente, resulta necesario desarrollar más de un sistema, además de unificar e integrar la información. En dicho contexto, constituye un desafío la operacionalización en terreno, el control cuantitativo de la recolección, el o los sistemas de análisis y monitoreo de la información recolectada para generar reportes oportunos por niveles geográficos, minimizando así el riesgo de generar resultados distintos por la utilización de un método u otro.

¹ Cabe recordar que el nombre genérico de INE hace referencia, en general, a los institutos y oficinas nacionales o direcciones generales o nacionales de estadísticas

En la actualidad, el monitoreo y control de la recolección censal ha evolucionado hacia sistemas automatizados, integrados y dinámicos, algo que ha llevado a mejoras en la supervisión del avance en terreno, gestión de personal y monitoreo de indicadores demográficos y de desempeño. En este sentido, las innovaciones tecnológicas efectuadas en geoespacialidad figuran hoy como herramientas imprescindibles para el monitoreo de la recolección censal y como parte del Sistema Integral de Gestión Operativa, SIGO. Al mismo tiempo es importante señalar que, en los censos de derecho mediante entrevistas por DMC, hay mayores posibilidades de reaccionar con un plan de contingencia ante una alerta temprana en el operativo y mayores oportunidades para supervisar y asegurar la calidad del proceso.

En resumen, la decisión de cambiar el método de captura para recolectar datos censales, mediante innovación tecnológica, tendrá un alto impacto en todo el proceso censal. De modo que, dicha decisión deberá incluirse tempranamente en la planificación y organización del proyecto censal, ya que implica, además, una transformación en el modo de gestionar el proyecto y el trabajo operativo en terreno. De manera análoga, la planificación censal deberá ser flexible y contemplar ciertas holguras para incorporar los cambios derivados de las pruebas. Es así como, en definitiva, durante el proceso de aprobar el método de captura que utilizará en un futuro censo, el Instituto de Estadísticas analice, debidamente, si están dadas las condiciones técnicas, humanas, organizacionales, presupuestarias y legales, entre otras variables, para adoptar dicha tecnología e implementarla de forma sostenible durante y posteriormente al censo.

Introducción

Un censo de población y vivienda es la actividad de recopilación de datos más grande, complicada y costosa que se lleva a cabo en cualquier sistema estadístico nacional del mundo. Su complejidad se explica, en parte, por la cantidad de esfuerzos y recursos en que se debe incurrir para identificar y secuenciar, adecuadamente, una multitud de procesos y actividades simultáneas e interrelacionadas, así como también ejecutar y monitorear su correcta implementación, considerando su gran escala. Se suma a ello el esfuerzo logístico requerido y las dificultades propias de recopilar información sobre toda la población de un país en un período limitado de tiempo.

En este contexto, la adopción de nuevas tecnologías para recolectar de datos en los censos de población y vivienda representan una gran oportunidad para mejorar, significativamente, la eficiencia de las operaciones y la calidad de los datos y productos. No obstante, su incorporación conlleva una serie de actividades, pruebas, riesgos y desafíos que, también, deben incluirse en la planificación y gestión del censo, con implicaciones para las áreas operativas, metodológicas y tecnológicas involucradas en este proceso.

Considerando dicho contexto, en noviembre de 2019, durante la Décima Reunión de la Conferencia Estadística de las Américas de la Comisión Económica para América Latina y el Caribe (CEA-CEPAL), se aprobó el funcionamiento del Grupo de Trabajo de Censo (GTC) para el bienio 2020-2021, con el objeto de fortalecer las capacidades técnicas de los institutos nacionales de estadística (INE) de la región para la etapa censal de procesamiento, consistencia y validación de datos. Al mismo tiempo y en el marco de la preparación de la ronda de censos 2020, el fin de dicho grupo de trabajo fue promover la coordinación y la cooperación en esta temática entre los países de la región. Es en dicho escenario que se planificó una serie de actividades para elaborar el documento titulado “Lineamientos para el procesamiento, validación y consistencia de datos censales: experiencia, buenas prácticas y desafíos para los países de América Latina”.

El grupo de trabajo fue liderado por el INE de Chile, con el Centro Latinoamericano y Caribeño de Demografía (CELADE)-División de Población de la Comisión Económica para América Latina y el Caribe (CEPAL) como secretaria técnica y la participación de Aruba, Bolivia, el Brasil, Colombia, Costa Rica, Cuba, el Ecuador, Nicaragua, el Paraguay, el Perú, la República Dominicana y el Uruguay.

Con la irrupción de la pandemia por COVID-19 y el impacto que generó en los INE de la región, el GTC aprobó la decisión de focalizar su labor exclusivamente en el desarrollo y redacción de los lineamientos para la captura de datos censales.

Para dicho estudio, uno de los objetivos específicos fue reunir las experiencias de los países respecto de la utilización de algunos métodos de recolección y captura de datos censales, para lo cual se seleccionaron siete países bajo el criterio del tipo de captura de datos utilizado en su último censo. La recolección de las experiencias se efectuó mediante la aplicación de una ficha cuyo objetivo fue recabar y, posteriormente, sistematizar la experiencia de los países acerca del método de captura de datos censales. Los países participantes en este estudio fueron: Brasil, Chile, Colombia, Costa Rica, Ecuador, México y Venezuela (República Bolivariana de)². Así fue como, durante el año 2020, se realizaron cinco reuniones virtuales de trabajo con los países ya mencionados donde se compartieron las experiencias de cada uno de ellos referente a tres métodos de captura empleados en sus respectivos censos; a saber: papel, eCenso y DMC³.

Por tanto, dado el inicio de la ronda de censos de la década 2020, en la región de América Latina y el Caribe, las necesidades de asistencia técnica declaradas por los INE⁴ y considerando el marco de las actividades del GTC CEA-CEPAL, resulta oportuno contar, hoy, con un documento orientador que brinde lineamientos para la captura de datos censales con gran énfasis en el uso de mayor tecnología, teniendo en cuenta la amplitud y complejidad de la preparación y ejecución de los censos de población y vivienda, acentuado por el contexto que impone la pandemia por COVID-19.

Sin duda, los países de la región han ido avanzando hacia una mayor utilización de tecnología en sus procesos censales en una constante búsqueda por hacer más eficiente y eficaz un proceso de por sí costoso, apuntando a mejorar la calidad y oportunidad de los datos censales. Dicho proceso ha decantado en una permanente conversación sobre las bondades y desafíos que representa para los INE de la región avanzar en este sentido, toda vez que el proceso no es simple ni está exento de problemas, pues no se trata solo de una cuestión de recursos financieros, sino de la capacidad institucional de innovar, sostenidamente y sin arriesgar el operativo censal ni la calidad de los resultados, algo que demanda determinados perfiles profesionales, así como habilidades gerenciales y de gestión que antes no eran tan relevantes para asegurar el censo.

Si bien, la presente publicación se inició antes del comienzo de la pandemia, la crisis sanitaria solo ha hecho más tangible la necesidad de avanzar hacia un mayor uso de tecnologías en la recolección de datos, a fin de minimizar la interacción humana sin afectar la calidad de los datos. Aquí es donde las alternativas de innovación, usando dispositivos móviles de captura (DMC) y entrevistas autoadministradas por internet (eCenso), tienen un gran potencial y se hacen aún más necesarias, dado que permiten reducir los tiempos de entrega de los resultados, en momentos cuando se requiere información oportuna para la toma de decisiones y hacer frente a la pandemia, como también para la implementación de políticas públicas que permitan superar la actual crisis económica y social.

El tipo de método de captura es una de las definiciones cruciales que se deben considerar al momento de analizar y planificar un censo de población y vivienda, entendiendo por método de captura

² Véase en el Anexo 2, las fichas utilizadas.

³ Las reuniones se realizaron el 5 de noviembre (reconocimiento y lectura óptica), el 19 y 26 de noviembre (DMC), y el 10 y 18 de diciembre (eCenso).

⁴ El nombre genérico de INE hace referencia a los institutos, oficinas y direcciones generales o nacionales de estadísticas.

el proceso de entrada, ingreso o digitalización de los datos censales, proceso que puede realizarse durante o después de la enumeración, dependiendo del método utilizado durante el operativo: electrónico con DMC, eCenso o papel con digitación y escaneo (lectura óptica).

Es así como, el presente documento contiene recomendaciones para apoyar dicho análisis, sobre la base de la experiencia de América Latina y el Caribe en las últimas rondas censales y de la experiencia acumulada por los profesionales interdisciplinarios de los INE y del CELADE que participaron en los diversos intercambios ya mencionados. Al mismo tiempo, el documento se apoya, fuertemente, en las recomendaciones expuestas en el documento *Guidelines on the use of electronic data collection technologies in population and housing censuses* de Naciones Unidas (2019). Por lo cual, en los siguientes capítulos se abordarán los diferentes métodos de captura empleados para el proceso de recolección de datos y su interacción con etapas previas y posteriores (planificación o procesamiento), a fin de levantar alertas o llamados de atención sobre aspectos que deben ser considerados al momento de hacer el análisis decisional respecto del método de captura que utilizarán las instituciones y organismos correspondientes en sus próximos censos. Adicionalmente, se bosquejan los desafíos en la captura de datos durante la etapa de empadronamiento, en un escenario donde, cada vez más, los países complementan los mecanismos de recolección y buscan el aprovechamiento óptimo de la tecnología.

I. Contenido y uso de los lineamientos

Este capítulo describe el objetivo y la estructura del documento “Lineamientos generales para la captura de datos censales” y ofrece una introducción básica a los métodos de captura y a las definiciones de los principales términos utilizados en el presente documento.

A. Objetivo y estructura

El objetivo de los lineamientos es ofrecer una guía práctica y general que ayude a orientar la toma de decisiones respecto al método de captura que se utilizará en los futuros censos en América Latina y el Caribe, teniendo en cuenta tanto la amplitud y complejidad de la preparación y ejecución de las mediciones de población y vivienda, como las complicaciones que ha agregado la pandemia por COVID-19. Al mismo tiempo, considerando que ya se ha iniciado la ronda censal 2020 —y dado el creciente interés de los países de la región por avanzar hacia el uso de DMC e Internet en los procesos de recolección—, resulta oportuno contar con un documento que sintetice algunas experiencias regionales interesantes, contribuyendo a promover y facilitar una mayor incorporación de tecnología en la captura de datos censales.

De esta manera, los lineamientos tienen por objeto aportar al mejoramiento de la calidad de los datos censales en la región al ofrecer orientación respecto de las ventajas y desventajas de cada método de captura, junto con examinar el contexto orgánico e institucional que enfrenta cada INE al momento de decidir avanzar hacia el uso de mayor tecnología para la captura de datos. Asimismo, se revisa la complejidad de algunos procesos que requieren nuevas habilidades y competencias en los equipos técnicos y profesionales al agregar mayor tecnología en la captura de datos, incluyendo su gran potencial para mejorar la calidad y oportunidad de los resultados censales.

En ningún caso, los lineamientos propuestos tienen por objeto sustituir los manuales o guías ya existentes sobre los métodos de captura, planificación o implementación de un censo de población y vivienda. Más bien, las indicaciones persiguen acercar las directrices —sobre el uso de tecnologías de recopilación de datos electrónicos en los censos de población y vivienda (Naciones Unidas, 2019)—, a la

realidad de los censos y oficinas de estadísticas de la región, por medio de una integración, interpretación y priorización de los lineamientos comprendidos en ella, con la “*Guía para elaborar un proyecto censal*” (CELADE-CEPAL, 2011), sumada a la experiencia transmitida por los países durante las distintas videoconferencias de revisión de las fichas de sistematización de experiencias para los países de América Latina y el Caribe⁵, realizadas por el Grupo de Trabajo Censos (CEA-CEPAL) para el acopio de antecedentes en la construcción de este documento, y a la experiencia acumulada en el Centro Latinoamericano y Caribeño de Demografía (CELADE)- División de Población. Debido a esto, algunos textos son frases o párrafos con traducciones libres de los documentos originales y en otros se fusionan o reinterpretan en función de la experiencia regional.

Todos los capítulos de este documento proporcionan diferentes aristas al tema de los métodos de captura de los datos censales y permiten obtener información relevante tanto para el INE y la gerencia o administración del proyecto, como para los encargados del desarrollo de las tecnologías de información de los censos, referente a la toma de decisiones del método de captura y sus principales alcances en el proceso censal:

- **Capítulo II. Antecedentes sobre la captura de datos censales en la región:** en este acápite se exponen los métodos de captura que han utilizado los países en sus últimos censos y las principales innovaciones que están pensando integrar en sus futuros proyectos censales.
- **Capítulo III. Métodos para la captura de datos censales:** el apartado se centra en las características principales de cada método de captura, sus principales ventajas, desventajas y aportes potenciales al mejoramiento de la calidad y oportunidad de los datos censales.
- **Capítulo IV. Matriz de decisión – criterios de evaluación:** en esta sección se definen matrices de decisión cualitativas y cuantitativas que, en conjunto, pueden ser utilizadas por los países para hacer una completa evaluación y optar por un método de captura u otro, de acuerdo a sus contextos específicos.
- **Capítulo V. Monitoreo del proceso de recolección censal:** en el capítulo se abordan los principales aspectos que debieran considerarse al diseñar un sistema de monitoreo para el proceso de recolección de datos censales, conteniendo una serie de indicadores que permiten generar alertas tempranas para asegurar la calidad y cobertura de la información que se está relevando.
- **Capítulo VI. Consideraciones de gestión y planificación vinculadas con la captura de datos censales:** en este apartado se plantean los aspectos y requisitos que se imponen en la planificación y gestión del proyecto al implementar métodos de captura de mayor tecnología, como DMC o sistemas de eCenso. Para esto, se basa en los aspectos de mayor importancia para la región, incluidos en Naciones Unidas (2019), al tiempo que se muestran ejemplos de algunos países y algunos instrumentos existentes de apoyo para la planificación.

⁵ Para más detalle, en el Anexo 2 se adjuntan las Fichas de Sistematización de Experiencias para Países de América Latina y el Caribe, del Grupo de Trabajo Censos de la Conferencia Estadística de las Américas-CEPAL, para la elaboración del documento “Lineamientos generales para la captura de datos censales” durante el año 2020.

B. Términos o conceptos utilizados

Censo tradicional⁶: implica la completa enumeración de campo; es decir, la información censal de personas, hogares y viviendas que se recopila mediante un cuestionario, aplicado directamente a las personas por los enumeradores (encuestadores) u otros modos de recopilación, por ejemplo, entrevista telefónica, envío/devolución por correo, Internet o una combinación de estos⁷.

Censo basado en registros⁸: la información, a nivel individual, sobre personas y hogares se recopila total o parcialmente desde fuentes administrativas. Es decir, la enumeración se constituye a partir de un registro de población, en el cual los campos para los diferentes atributos del censo se completan (total o parcialmente) con los datos disponibles en registros subsidiarios relacionados con temas específicos.

Métodos de enumeración⁹: en un contexto de censos tradicionales, existen dos métodos principales de enumeración: la entrevista cara a cara (con enumerador) y la autoenumeración. En el primer caso, la información para cada persona (en un censo de población) y para cada vivienda y sus ocupantes (en un censo de vivienda) se recopila y se ingresa en el cuestionario mediante un censista designado para realizar esta operación en una zona específica. En el caso del método de autoenumeración, la responsabilidad de ingresar la información recae en una persona de la misma unidad que se enumera (generalmente, la persona de referencia del hogar), aunque el cuestionario puede ser distribuido, recopilado y verificado por un funcionario del censo.

Método de captura: en el presente documento, se entenderá por método de captura la forma por la cual los datos de cada unidad enumerada ingresan al registro electrónico de información censal: por digitación manual, por sistema de lectura óptica en el caso de un cuestionario en papel o por el registro en un cuestionario electrónico mediado por un censista (en el caso de los DMC) o por la misma persona encuestada por autoenumeración (en el caso del eCenso)¹⁰.

Recolección de los datos censales (en un censo tradicional): se entenderá por este procedimiento la combinación entre la forma en que se hace contacto con los encuestados (método de enumeración) mediante entrevista cara a cara o autoempadronamiento y el modo de ingresar o digitalizar sus respuestas (método de captura), por digitación, lectura óptica (escáner) o cuestionario electrónico (DMC o eCenso), como se puede apreciar en el Cuadro 1.

eCenso: se denomina al método de autoenumeración mediante el uso de un formulario electrónico disponible en la Web.

⁶ Véase párrafos 1.69 al 1.72 de United Nations (2017).

⁷ En contraposición a ello, se encuentran los censos que usan total o parcialmente los registros administrativos. Esto implica que los datos de los registros se emplean no sólo como marco o para respaldar las operaciones de campo, sino directamente como fuente de datos para cierta información del censo.

⁸ Ídem párrafos 1.80 al 1.84.

⁹ Ídem párrafo 3.120.

¹⁰ Para fines de este documento, el sistema CATI (Computer Assisted Telephone Interview) no se incluye como método de recolección o de captura de datos, dado que las experiencias en la región, no siempre se refieren a un sistema telefónico de entrevistas asistido por computadoras, sino que se refiere a un sistema en que un entrevistador o censista aplica el censo a la persona de referencia y luego ingresa la información en sistemas que no fueron diseñados especialmente para esto, como en los mismos DMC que utilizan los censistas en terreno o en el aplicativo para el eCenso, incluso en cuestionarios en papel que, luego, son capturados por digitación o escáner.

Cuadro 1
Matriz de métodos de recolección censal

Método de Recolección	Método de captura			
	Posterior a la recolección		Simultáneo a la recolección	
Método de enumeración	Digitación	Escaneo	DMC	eCenso
Entrevista cara a cara	x	x	x	
Auto respuesta (correo postal)	x	x		
Auto respuesta (eCenso)				x
Auto respuesta (correo electrónico)				x

Fuente: elaboración propia.

Calidad: “es el grado en que un conjunto de características inherentes a un objeto cumple los requisitos” (ISO 9000:2015 s.f.). En el contexto de las organizaciones estadísticas, el objeto es la producción o producto estadístico, el proceso, el entorno institucional o el sistema estadístico en su conjunto. Una definición simple de calidad es "apto para el uso" o "apto para la finalidad". Son las necesidades de los usuarios las que definen la calidad. Los distintos usuarios pueden tener necesidades diferentes que deben equilibrarse, entre sí, para dotar al concepto de calidad de un contenido concreto.

En el Manual de censos¹¹ se plantea que el sistema de aseguramiento de la calidad debe desarrollarse como parte del programa general del censo e integrarse transversalmente en todas las fases censales, mencionando que, al final, la calidad tiene que ver con la satisfacción de las necesidades de los usuarios y que —si bien, en estadística se tendía a asociar la calidad, principalmente, con la precisión—, ahora se reconoce como un concepto multidimensional que es el resultado de los procesos y sus deficiencias, más que de las acciones de las personas que trabajan en ellos. Por lo tanto, los procesos deben mostrar, al menos, solidez metodológica y adherencia a métodos profesionales y normas acordadas internacionalmente y eficiencia, entendida como el grado en que, en la compilación, se reducen los costos y la carga a los usuarios en relación con la producción.

Asimismo, la calidad en la producción estadística sugiere que el resultado de cualquier ejercicio estadístico debe poseer alguno o todos los siguientes atributos:

- Pertinencia, entendida como el grado en que las estadísticas satisfacen las necesidades de los usuarios y sugiere la necesidad de evitar la producción de datos irrelevantes; es decir, datos para los que no se encontrará ningún uso;
- Completitud, es el grado en que las estadísticas cubren completamente el fenómeno que se supone que describen;
- Precisión, refiere a la distancia entre el valor estimado y el valor verdadero (desconocido);
- Comparabilidad, el grado en que las estadísticas son comparables en el espacio (entre países) y en el tiempo (entre diferentes períodos de tiempo);
- Coherencia, alude al grado en el que los datos de un solo programa estadístico y los datos reunidos entre programas estadísticos están lógicamente conectados;
- Oportunidad, es el tiempo transcurrido entre el período de referencia y la publicación de los datos;

¹¹ Ídem párrafos 2.170 a 2176.

- Puntualidad, entendida como el grado de cumplimiento de las fechas de liberación de los resultados anunciados previamente;
- Claridad, se refiere al grado en que las estadísticas son comprensibles para usuarios no expertos;
- Accesibilidad, alude a la facilidad con la que los usuarios pueden obtener datos estadísticos;
- Interpretabilidad de los datos del censo, incluidos los metadatos, disponibilidad de información que describa fuentes, definiciones y métodos; y,
- Confianza, en términos de la confianza del público en la autoridad estadística nacional y sus entregas.

Calidad en referencia a los métodos de captura: a lo largo del documento, se menciona la forma en que los métodos de captura pueden mejorar la calidad del censo y sus datos. En este sentido, se hace hincapié en algunas de las dimensiones de la calidad, dado su carácter multidimensional. Es decir, se hará referencia al potencial de la tecnología para favorecer la calidad de los datos, mejorando su completitud, precisión, oportunidad y puntualidad. Esto se traduce en la reducción de errores durante el proceso de recopilación de la información censal como en la digitación, la pérdida de material censal y la inconsistencia en las respuestas. También, al incluir alertas tempranas sobre sesgos u omisiones acortando los tiempos para la publicación, junto con hacer viable el que la base de datos esté disponible para su procesamiento, casi inmediato, al término del trabajo de campo.

II. Antecedentes sobre la captura de datos censales en la región

La realización de un censo de población y vivienda es la operación estadística más compleja a la que se enfrenta un INE debido a su volumen, extensión y cantidad de procesos que se deben realizar, algunos de forma simultánea y otros secuencialmente. Esto implica la colaboración de un gran número de personas en un tiempo limitado y en tareas de muy diversa índole. En cada ronda censal se hacen recomendaciones tendientes a la mayor integración de tecnología en cada etapa censal, según los contextos y capacidades nacionales, con el objetivo de mejorar el proceso completo, reduciendo la intervención humana y las posibilidades de error en toda la cadena de valor, redundando, finalmente, en una mayor eficiencia, oportunidad, calidad y seguridad, entre otros atributos deseables.

En el caso puntual de los procesos asociados a la recolección de datos, la combinación entre el método de enumeración; esto es, la forma en que se toma contacto con los encuestados (entrevista cara a cara o autoempadronamiento) y el método de captura; es decir, la forma en que se ingresan o digitalizan sus respuestas (digitación, lectura óptica, DMC o eCenso); hoy es de gran interés para los países de la región, quienes se interesan cada vez más en la incorporación de mayor tecnología para disminuir el trabajo manual, con el objetivo principal de mejorar la calidad y oportunidad de los resultados censales.

Sin embargo, se debe señalar que el mayor uso de tecnología en la captura de los datos censales ha incorporado un grado de confusión en la secuencialidad de algunos procesos, en comparación a cuando solo se utilizaba el papel, con captura por medio de la digitación o la lectura óptica. Esto se debe, principalmente, a que la inclusión de los DMC o el eCenso, necesariamente implican repensar la secuencialidad de los procesos relacionados con la captura de la información, tales como el diseño del cuestionario, el control de seguimiento, algunas validaciones y consistencias, entre otros procesos. Por ejemplo, al efectuar la captura por medio de DMC o eCenso, con seguridad en el momento de la recopilación de los datos, se realizarán algunas validaciones y consistencias en ese instante. En cambio, cuando se usa papel éstas se realizan posteriormente o se consideran parte del procesamiento.

Otro ejemplo de cómo se van mezclando los procesos y se dificulta la planificación, especialmente la primera vez que se avanza desde un censo en papel a uno con DMC, es el hito de la obtención de la base inicial (o sucia) de microdatos. Es decir, en el caso de la DMC y del eCenso, el momento en que se cuenta con las respuestas del censo en formato digital es simultáneo con el término de la etapa de recolección de información. En cambio, con el uso de papel, este hito se cumple varios meses después de finalizada la etapa de recolección. Dicha diferencia entre métodos de captura no es menor, puesto que esta primera base de microdatos es el insumo clave para las etapas siguientes de procesamiento, análisis y disseminación de los resultados, por lo tanto, es deseable para los países obtenerla lo más pronto posible.

De esta forma, el uso de tecnologías computacionales permite que este hito se cumpla e inclusive contribuya al mejoramiento de la calidad, oportunidad y universalidad del censo, a través de la reducción de la interacción humana (errores en la digitación), otorgamiento de una mayor seguridad a los datos, generación simultánea al operativo de terreno de archivos digitales y resguardo, ante riesgos de pérdida de información por eventos naturales o humanos, a los cuales los registros en papel están expuestos (inundaciones, incendios, robos u otros eventos).

La reducción de los costos de las tecnologías, de forma generalizada, ha propiciado una mayor tendencia a su incorporación e innovación en los censos; sin embargo, en los países de la región, el tipo de censo (de hecho o derecho) sigue siendo un factor determinante a la hora de definir el método de captura de los datos, a la par de cómo se proveerán los servicios tecnológicos que se requieren; es decir, si se abordará por medio de desarrollo propio de los INE (*in-house*) o por medio de un proveedor externo (tercerizado o *outsourcing*).

Un primer aspecto definitorio en la elección del método de enumeración y de captura de datos es el tipo de censo. A continuación, se presenta un resumen de las últimas cuatro décadas en que los países de la región han implementado censos de hecho o derecho.

Con el afán de conocer sobre el uso de la tecnología en los países de la región, el Grupo de Trabajo de Censos de la CEA-CEPAL aplicó —en 20 países de la región, entre julio y octubre de 2018 y con una tasa promedio de respuesta del 89%—, la “Encuesta sobre el uso de tecnología en los censos de población y vivienda de la ronda 2020¹²”, cuyo objetivo general fue: “Conocer el estado actual sobre el uso de tecnología en los censos de población y vivienda que permita orientar y fomentar el intercambio de buenas prácticas, cooperación sur-sur, la asistencia técnica y la cooperación internacional en tecnología”.

Sobre la base de esta encuesta y otras informaciones sistematizadas en el CELADE, se puede inferir el estado de avance de la región en cuanto a la utilización de mayor tecnología en los métodos de captura, por lo que, a continuación, se presentan algunos de los resultados y se recuerda al lector que los datos referentes a la década de 2020, emanados de esta encuesta, se refieren a las planificaciones e ideas que tenían los respectivos INE en 2018 sobre el futuro censo y no corresponden, necesariamente, a la solución final implementada.

Se observa, en el Cuadro 3, la tendencia en la región de ir avanzando desde censos realizados en papel hacia censos con dispositivos móviles de captura (DMC), pasando o no, por diferentes combinaciones entre censos de hecho o derecho con digitación, uso de escáner para la captura óptica de datos, DMC y mixtos con autoenumeración por medio de eCenso. A la vez, se observa que, más países están planificando una mayor combinación de métodos secundarios de captura, para mejorar la cobertura y acceder a un mayor número de viviendas al ofrecer diversas formas para censarse.

¹² Ídem.

Cuadro 2
Censos de población y vivienda en América Latina, según tipo de censo

País	Tipo de censo según década			
	1990	2000	2010	2020a
Argentina	Hecho	Hecho	Hecho	Derecho
Bolivia (Estado Plurinacional de)	Hecho	Hecho	Hecho	Hecho
Brasil	Derecho	Derecho	Derecho	Derecho
Chile	Hecho	Hecho	Hecho	Derecho
Colombia	Derecho	Derecho	Derecho	Derecho
Costa Rica	...	Derecho	Derecho	Derecho
Cuba	...	Derecho	Derecho	Derecho
Ecuador	Hecho	Hecho	Hecho	En proceso de definición ^b
El Salvador	Derecho	Derecho
Guatemala	Derecho	Derecho	Derecho	...
Haití	...	Derecho	...	Derecho
Honduras	...	Derecho	Derecho	Derecho
México	Derecho	Derecho	Derecho	Derecho ^c
Nicaragua	Derecho	Derecho
Panamá	Hecho	Hecho	Hecho	Derecho
Paraguay	Hecho	Hecho	Derecho	Hecho
Perú	Hecho	Hecho	Hecho	...
República Dominicana	Hecho	Derecho	Derecho	Derecho
Uruguay	Hecho	Derecho	Derecho	Derecho
Venezuela (República Bolivariana de)	Derecho	Derecho	Derecho	Derecho

Fuente: Elaboración propia sobre la base de: Principales cambios en las boletas de los censos latinoamericanos de las décadas de 1990, 2000 y 2010, CELADE-CEPAL - Serie Manuales N.º 80, 2012 y Encuesta sobre el uso de tecnología en los censos de población y vivienda de la ronda 2020³³ y declaraciones de los países en asistencias técnicas en reuniones internacionales.

^a corresponde al tipo de censo que tienen pensado realizar en la ronda 2020 al momento de realizar la encuesta.

^b corresponde a información actualizada, entregada por los países en Asistencias Técnicas o seminarios. Los cambios atienden a reprogramaciones producto de la pandemia.

^c Censo realizado en marzo de 2020.

³³ Algunos resultados de esta encuesta se presentaron en el Seminario "Aspectos conceptuales de los censos de población y vivienda: desafíos para la definición de contenidos incluyentes en la ronda 2020", Santiago de Chile, 6-8 de noviembre, 2018. <https://www.cepal.org/es/publicaciones/44944-aspectos-conceptuales-censos-poblacion-vivienda-desafios-la-definicion>. CEPAL (2019).

Cuadro 3
Censos de población y vivienda en América Latina, según método principal y secundario de captura.

País	Década de 2010		Década de 2020a	
	Método principal	Método secundario	Método principal	Método secundario
Argentina	Papel (escaneo)		Papel (escaneo)	eCenso
Bolivia (Estado Plurinacional de)	Papel (escaneo)		Papel (escaneo)	
Brasil	DMC	eCenso	DMC	eCenso
Chile	Papel (escaneo)	eCenso	DMC	eCensob
Colombia	DMC	eCenso y Papel (Digitación)	No aplica	No aplica
Costa Rica	Papel (escaneo)		DMC b	eCenso b
Cuba	Papel (digitación)		DMC b	
Ecuador	Papel (escaneo)		En proceso de definición b	
El Salvador			DMC b	Papel b
Guatemala	Papel (digitación)	eCenso		
Haití	...		DMC b	Papel
Honduras	Papel (digitación)		DMC	
México	Papel (digitación)	CATI, autoenumeración en papel, entregado y recolectado por el entrevistador, eCenso	DMC c	eCenso, papel (digitación) y CATI c
Nicaragua	...			
Panamá	Papel (digitación)		DMC b	Papel b
Paraguay	DMC	Papel	Papel (escaneo) b	
Perú	Papel (escaneo)			
República Dominicana	Papel (escaneo)		DMC b	Papel b
Uruguay	DMC	Papel	DMC b	eCensob
Venezuela (República Bolivariana de)	DMC	Papel	eCenso y DMC	

Fuente: Elaboración propia a partir de la base de la encuesta sobre el uso de tecnología en los censos de población y vivienda de la ronda 2020 y declaraciones de los países de la región en asistencias técnicas o en reuniones internacionales.

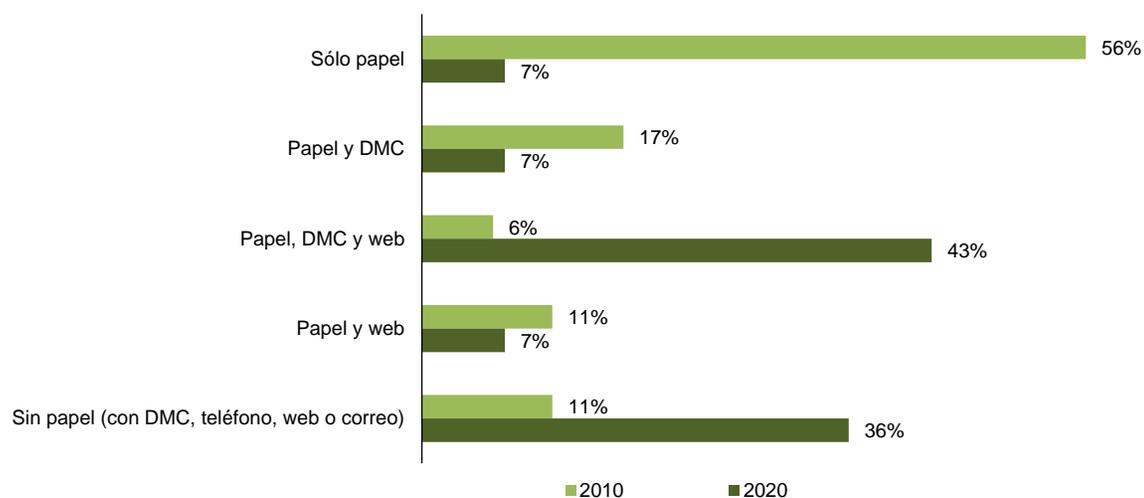
^a corresponde al tipo de censo que tienen pensado realizar en la ronda 2020 al momento de realizar la encuesta.

^b corresponde a información actualizada, entregada por los países en asistencias técnicas o reuniones.

^c censo realizado en marzo de 2020.

En la década de 2010, de 18 países que realizaron sus censos, 10 lo hicieron solo por medio de entrevistas cara a cara, utilizando formularios en papel, cifra que representó el 56%, al tiempo que sólo dos países de la región hicieron censos sin utilizar papel; es decir, la recolección se realizó utilizando mayor tecnología (11%). En cambio, para la década del 2020, de los 14 países que entregaron información sobre las herramientas o métodos de captura que estaban planificando para el futuro censo, solo uno ha planificado utilizar exclusivamente papel, mientras ocho declararon que pensaban utilizar métodos mixtos que mezclan entre papel, eCenso y DMC (57%) y cinco esperaban no utilizar papel (36%). En el gráfico siguiente, se puede observar cómo los países de la región están evolucionando hacia el uso de mayor tecnología en la captura de la información censal.

Gráfico 1
Métodos de captura utilizados en la década de 2010 y los que se planifican para la década de 2020 en América Latina
(En porcentajes)



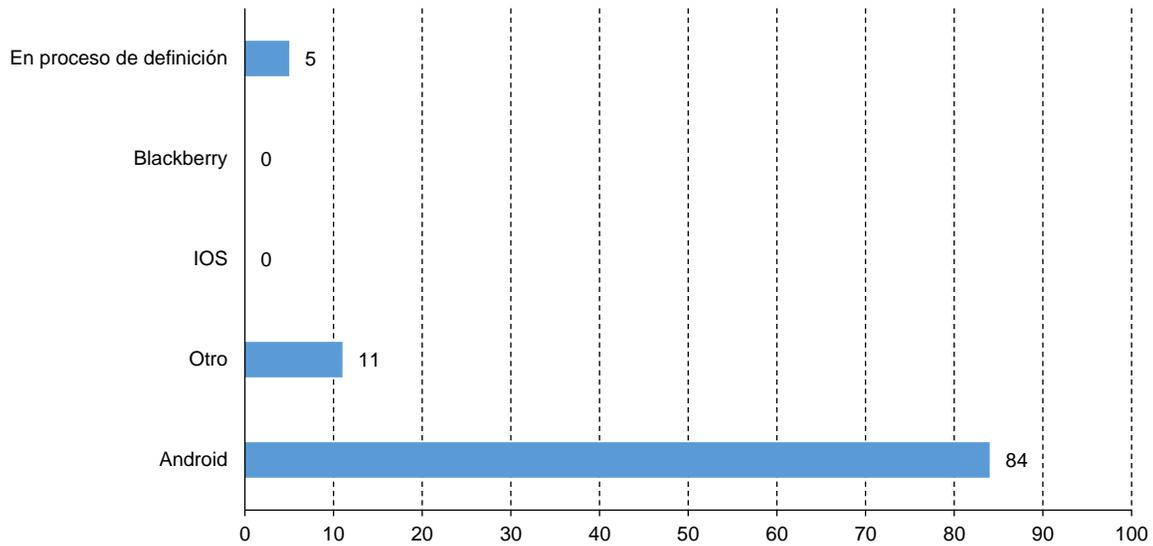
Fuente: Elaboración propia en base a la Encuesta sobre el uso de tecnología en los censos de población y vivienda de la ronda 2020, realizada por el GTC CEA-CEPAL en 2018.

En este contexto, es importante destacar que, algunos países incorporan métodos mixtos de captura, lo cual complejiza el proceso de enumeración y recopilación de información, al tener que controlar el avance del censo (cobertura territorial) en dicho escenario y convivir con fuentes de error que tienen diferente comportamiento, según el método de captura de origen, entre otros aspectos. En el uso de métodos mixtos se continúa utilizando papel, pero se ha ido incorporando la tecnología de manera creciente, a fin de facilitar la respuesta de los diferentes grupos poblacionales, para mejorar la calidad y cobertura.

De los mismos datos que dan origen al gráfico anterior, se desprende que, para la década de 2010, cinco de 14 países utilizaron DMC como método principal. En cambio, para la ronda censal de la presente década, un total de 11 países ha planificado utilizar DMC, por lo que adquieren mayor relevancia las decisiones que se deben tomar respecto al sistema operativo móvil que se planea implementar. El 84% de los países que contestaron esta pregunta declaró que pensaba utilizar, en el censo de la década de 2020, el sistema Android y 11% otro sistema. El 84% recién mencionado se compone de 16 países, de los cuales 13 declararon que utilizarían Android. En tanto, dos países usarían Android junto con el sistema Windows Mobile o Windows Phone; mientras que un solo país emplearía Android y otro sistema (Véase gráficos 2 y 3). Respecto de los tipos de DMC que estaban considerando utilizar, la mayor parte de los países se inclina por usar teléfonos móviles o Tablet.

Respecto de los países que han declarado la intención de utilizar DMC en su próximo censo, se observa en la misma encuesta una mayor intención de usar C# como lenguaje de programación para capturar datos, con un 37% del total, seguido por Java y por la alternativa "otro o ambos" con un 21% y un 11% que declara estar en proceso de definición sobre el tipo de lenguaje de programación que usarán.

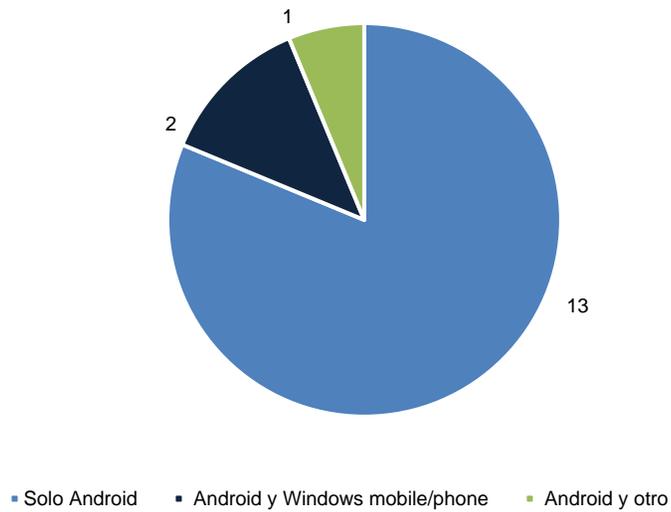
Gráfico 2
Sistemas operativos móviles que han planificado usar los países en la ronda censal de 2020
(En porcentajes^a)



Fuente: Elaboración propia en base a la Encuesta sobre el uso de tecnología en los censos de población y vivienda de la ronda 2020, realizada por el GTC CEA-CEPAL en 2018.

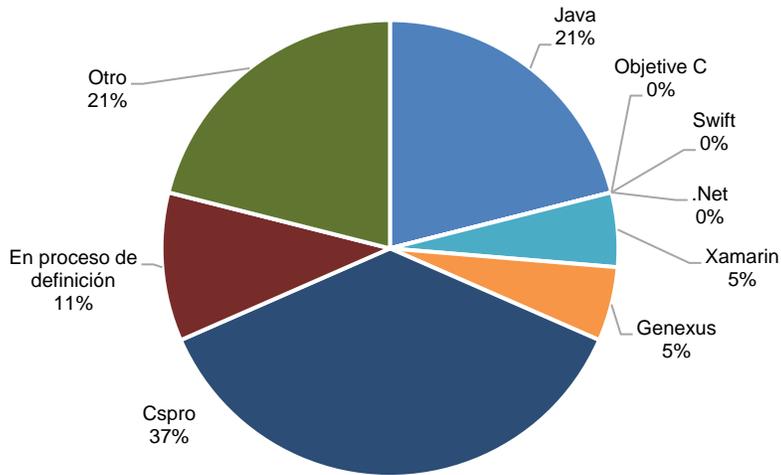
^a Sobre los 19 países que respondieron a la pregunta ¿Qué sistema operativo va a utilizar (Móvil)?

Gráfico 3
Cantidad de países que han declarado que van a utilizar sistemas operativos Android y otro sistema operativo móvil en los censos de la década 2020
(En cantidad de países)



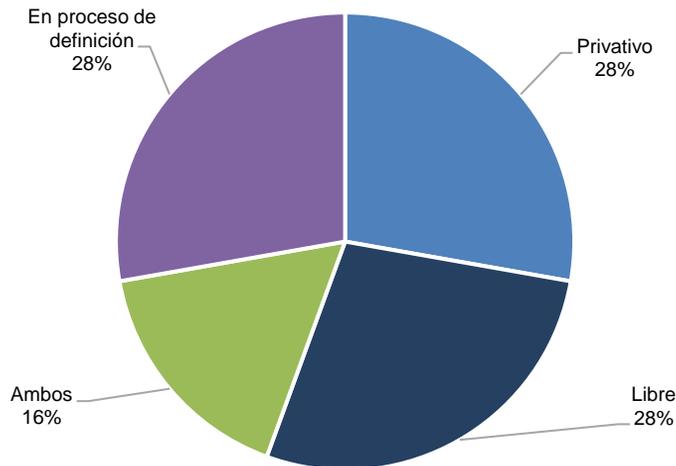
Fuente: Elaboración propia en base a la Encuesta sobre el uso de tecnología en los censos de población y vivienda de la ronda 2020, realizada por el GTC CEA-CEPAL en 2018.

Gráfico 4
Lenguaje de programación para la captura con DMC que se usaría en censos de la década 2020
(En porcentajes)



Fuente: Elaboración propia en base a la Encuesta sobre el uso de tecnología en los censos de población y vivienda de la ronda 2020, realizada por el GTC CEA-CEPAL en 2018.

Gráfico 5
Tipo de software para la captura de datos en los censos de la década de 2020
(En porcentajes)

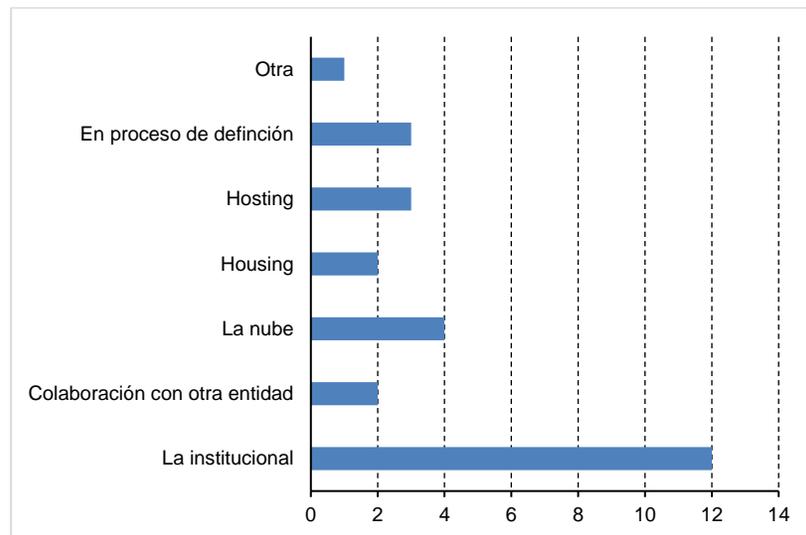


Fuente: Elaboración propia en base a la Encuesta sobre el uso de tecnología en los censos de población y vivienda de la ronda 2020, realizada por el GTC CEA-CEPAL en 2018.

Respecto a las respuestas relacionadas con la recogida de información censal en línea vía web o eCenso, se observa que la mayoría de los países (13 de 18 países) ya tiene definido el tipo de software que usarán para el ingreso de datos en los censos de la década de 2020. De los 13 países, el 16% declara interés en un uso mixto de soluciones libres y desarrolladas in-house; mientras los demás se inclinan en porcentajes similares por una u otra línea en forma exclusiva o están en proceso de definición (28% respectivamente).

Los países que han indicado el uso de un sistema de recolección de datos en línea, a través de la web (eCenso), para la década de 2020, declararon que, en general, utilizarían más de una plataforma de almacenamiento, siendo el sistema institucional la alternativa mayoritaria (12 países), seguido por la nube (cuatro países) y, posteriormente, las alternativas de *Hosting*; mientras que, en proceso de definición, hubo solo tres preferencias.

Gráfico 6
Plataformas tecnológicas para almacenamiento de datos mediante captura vía web en censos de la década de 2020
(En cantidad de países)

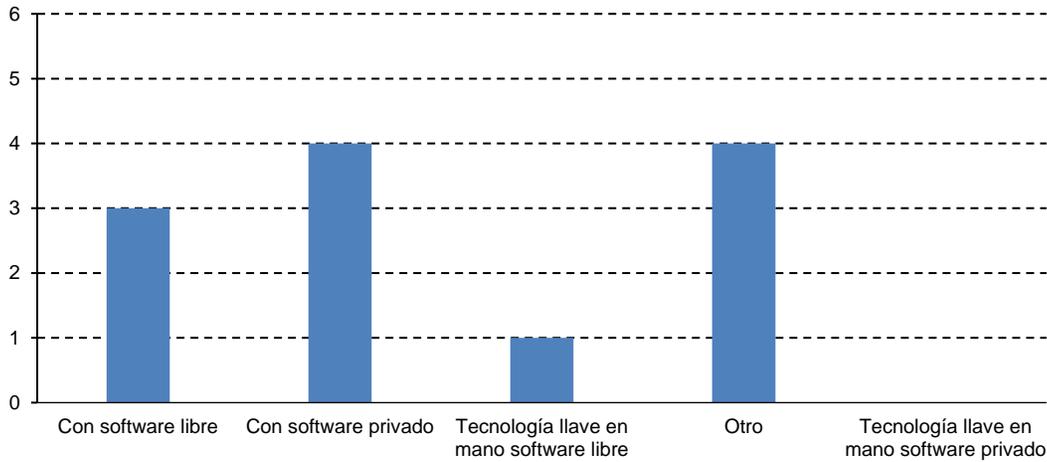


Fuente: Elaboración propia en base a la Encuesta sobre el uso de tecnología en los censos de población y vivienda de la ronda 2020, realizada por el GTC CEA-CEPAL en 2018.

Respecto a los sistemas de control, operativo y de cobertura, durante el trabajo de campo, 12 países declararon que, para los censos de la década del 2020, tenían pensado utilizar sistemas automatizados y cuatro que estaban en proceso de definición. Estos sistemas constituyen una gran ventaja y oportunidad para mejorar la calidad de los datos censales, ya sea por la vía del control de cobertura, así como por las oportunas alarmas de sesgos en la captura de los datos, como por ejemplo el olvido de personas de edades menores o muy adultas u omisión selectiva de personas en edad de trabajar, entre otros.

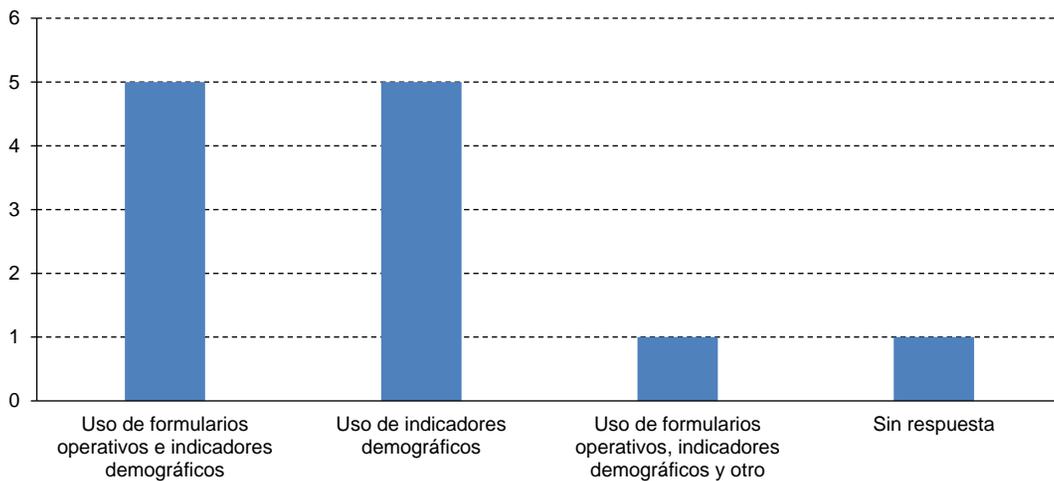
De esos 12 países, tres declararon que desarrollarían sus respectivos censos con software libre, cuatro con software privado, uno con llave en mano (tercerizado) con software libre (open source) y cuatro declararon que con otro sistema. Adicionalmente, se destaca que cinco países expresaron su interés en utilizar formularios operativos e indicadores demográficos en dichos sistemas para operacionalizar el control; mientras otros cinco solo utilizarían indicadores demográficos.

Gráfico 7
Cómo se desarrollará el sistema automatizado de monitoreo y control de cobertura en censos de la década del 2020
(En cantidad de países)



Fuente: Elaboración propia en base a la Encuesta sobre el uso de tecnología en los censos de población y vivienda de la ronda 2020, realizada por el GTC CEA-CEPAL en 2018.

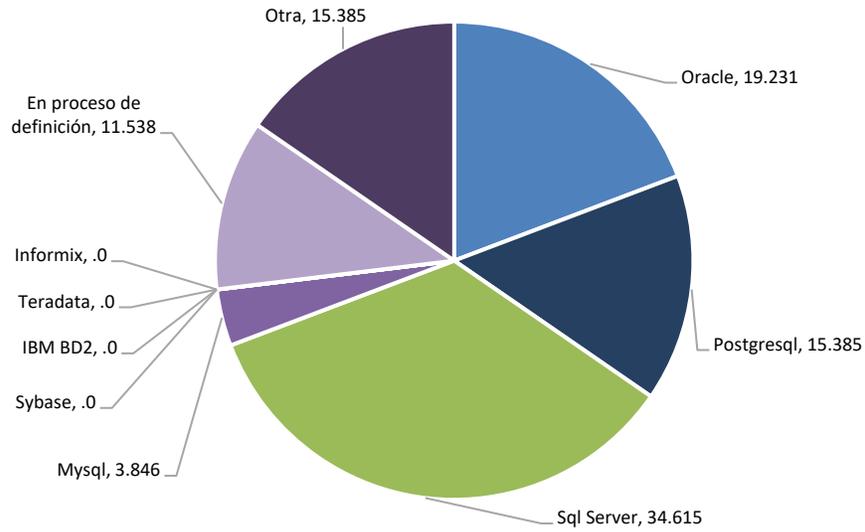
Gráfico 8
En qué se basa el sistema automatizado de monitoreo y control de cobertura en censos de la década del 2020
(En cantidad de países)



Fuente: Elaboración propia en base a la Encuesta sobre el uso de tecnología en los censos de población y vivienda de la ronda 2020, realizada por el GTC CEA-CEPAL en 2018.

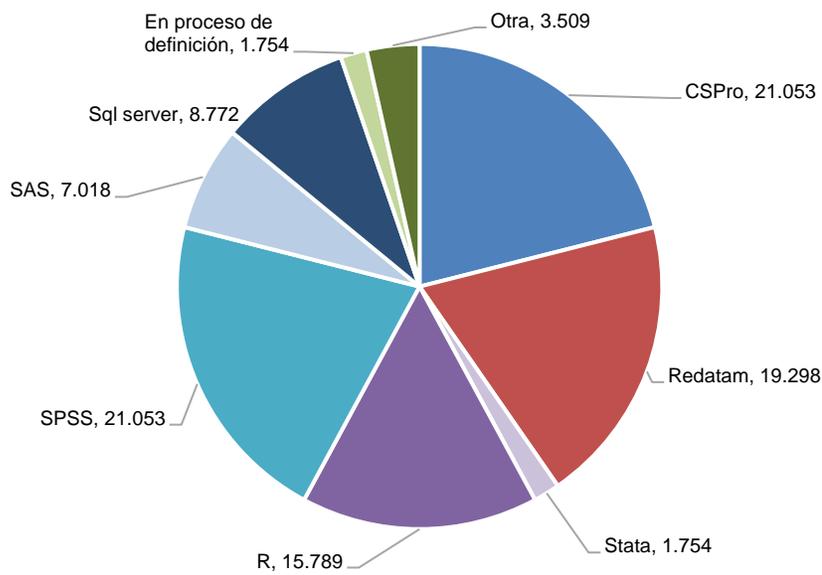
Los resultados de la encuesta reflejan la preferencia mayoritaria de los países por utilizar sistemas como SQL Server (35%), Oracle (19%) y Postgresql (15%) para el almacenamiento de la base de datos censales, persistiendo aún un alto porcentaje de países que usará otro sistema (15%) o que aún está en proceso de definición (12%). A la vez que, para el procesamiento de la base de datos censal 21% declaran su interés en utilizar CSPRO y SPSS; mientras que 19% prefieren usar REDATAM. Luego, siguen alternativas como R (16%), SQL Server (9%) y SAS (7%).

Gráfico 9
Sistemas de almacenamiento de datos en los censos de la década 2020
(En porcentajes)



Fuente: Elaboración propia en base a la Encuesta sobre el uso de tecnología en los censos de población y vivienda de la ronda 2020, realizada por el GTC CEA-CEPAL en 2018.

Gráfico 10
Software de procesamiento de datos en los censos de la década 2020
(En porcentajes)



Fuente: Elaboración propia en base a la Encuesta sobre el uso de tecnología en los censos de población y vivienda de la ronda 2020, realizada por el GTC CEA-CEPAL en 2018.

Los países, al ser consultados sobre los aspectos que deberían considerar para incorporar nuevas tecnologías de la información en su próximo ejercicio censal, declararon con mayor frecuencia el “Conocimiento para asimilar dentro de la institución” y “Adecuarse a las necesidades y dinámica de la organización”, ubicándose los recursos financieros en un tercer lugar. Según los encuestados, los altos costos de la tecnología ya no son la principal razón para la no adopción de tecnología, sino los aspectos relacionados con el capital humano y los factores organizacionales.

Toda esta información, entregada por los países en 2018, revela expectativas auspiciosas hacia la incorporación de mayor tecnología en esta ronda censal; primero, porque el hecho de evolucionar a censos de derecho facilita el uso de DMC y, luego, el levantamiento Web (eCenso) cobra mayor fuerza como complemento al método de captura principal. A la vez, se observa que las decisiones sobre herramientas tecnológicas se inclinan por utilizar, mayoritariamente, softwares como CPro, REDATAM y SPSS para el procesamiento y generación de tabulados, mientras el CPro mantiene su posición como la principal herramienta tanto para la recolección de información en terreno como para la consistencia de datos.

Por otro lado, se evidencia una brecha entre el deseo de avanzar hacia mayor tecnología y la infraestructura con que cuentan las oficinas de estadísticas en el momento de la encuesta. La mayoría de los INE de la región utilizan infraestructura tecnológica propia (interna) que incluye lo básico de un centro de datos, pero que no cumplen, necesariamente, con estándares tecnológicos ideales. En general, sólo cuentan con un sitio web principal sin respaldo y no con una plataforma de alta disponibilidad.

III. Métodos para la captura de datos censales

La recolección de datos censales incluye tanto la forma en que se contacta a la población como la forma en que se obtienen y registran sus respuestas, procesos que varían según los países. En un mismo operativo censal se puede optar por una única forma de recolección, como lo hizo el Instituto Nacional de Estadística e Informática (INEI) del Perú en el censo de 2017, empleando entrevistas cara a cara con censistas y cuestionarios en papel, los que luego se escanearon o una combinación de dos o más formas, como se implementó en el censo de Colombia de 2018, donde se utilizó entrevistas con censistas con DMC como método de captura principal en combinación con papel (digitación), e innovando con autoempadronamiento web (eCenso). Otro ejemplo de método mixto es el censo de Guatemala de 2018, donde el método principal de recolección fue la entrevista cara a cara con censistas y cuestionarios en papel, los cuales fueron posteriormente digitados en combinación con eCenso como una forma de mejorar y recuperar la cobertura.

Al seleccionar una o varias formas de recolección de datos, un INE debería considerar — preferentemente durante una etapa temprana de la planificación del censo—, cierto conjunto de factores, tanto estadísticos (cobertura, tasas de respuesta y errores de medición) como no estadísticos (oportunidad, costo, población de interés), en función de múltiples consideraciones, las cuales se tratan con mayor detalle en el capítulo VI de este mismo documento.

La estrategia de interacción con la población está vinculada con la forma en que se recopilan los datos durante el operativo censal, también llamado método de enumeración; es decir, entrevistas cara a cara o autoempadronamiento²⁴. En la mayoría de los países de la región, el primer método se ha operacionalizado a través del uso de papel o DMC y, más incipientemente en algunos países, el segundo método ha sido través del eCenso. A continuación, en el Cuadro 4, se detallan las características de cada uno.

²⁴ Para más detalle ver **Términos o conceptos utilizados** en el Capítulo I,

Cuadro 4
Método de enumeración y momento de captura de datos

Entrevista cara a cara con papel	La recolección de datos se realiza mediante boletas o cuestionarios impresos, los cuales deben pasar por todo un proceso de logística para llegar en orden a un local de almacenamiento donde, posteriormente, se realizará la captura de datos por medio de la digitación o lectura óptica.
Entrevista cara a cara con DMC	La recolección y captura de datos se hace simultáneamente, mediado por un censista que ingresa las respuestas en el dispositivo, generando inmediatamente un registro digital que, luego de un proceso de transferencia, queda disponible prácticamente de forma inmediata para resguardo, seguimiento, procesamiento y análisis.
Autoempadronamiento en línea vía Internet (eCenso)	La recolección y captura de datos se realiza, simultáneamente, durante el llenado del cuestionario en línea y sin la intervención de un censista. Los datos quedan disponibles de forma inmediata (sin necesidad de transferencia) para resguardo, seguimiento, procesamiento y análisis.

Fuente: Elaboración propia a partir de los intercambios de experiencias del GTC CEA-CEPAL realizados en 2020.

Otro elemento que se debe considerar es la utilización de formas complementarias de recopilación y/o recuperación de datos censales y que son, básicamente, otras combinaciones entre métodos de enumeración y de captura. Si bien, estas formas de recolección no se han usado tanto en la región por diferentes razones, algunos países están dispuestos a usarlas complementariamente, debido a los efectos de la pandemia, la mayor inseguridad o la falta de interés por contestar el censo, entre otras razones.

Entre los métodos de recopilación más valorados se encuentran:

- Entrevistas por medio de un centro de llamadas con censistas calificados (call-center) quienes ingresan las respuestas en papel y/o electrónico (DMC o aplicativo web);
- Autoempadronamiento por correo en papel (postal) y/o electrónico (formulario digital vía email); y,
- Transferencia directa desde registros administrativos de subpoblaciones específicas.

Esta última forma de recopilación de datos no hace referencia a un censo por registros administrativos (censo no tradicional), sino más bien a la recopilación de las características de las personas, hogares y/o viviendas cuyos datos se han consignado en un registro administrativo para otros fines, generalmente, no estadísticos, el que se utiliza de forma complementaria a un método principal de recopilación para alcanzar una mayor cobertura censal.

En la región de América Latina y el Caribe existen experiencias como la del censo de Colombia 2018 que recopiló información de subgrupos poblacionales en algunos tipos de viviendas colectivas como penitenciarías, centros de salud o cuidado de larga estadía, entre otros.

Recuadro 1**Uso de registros administrativos en el Censo de Colombia 2018**

La recolección en las viviendas colectivas (Lugares Especiales de Alojamiento, LEA), se realizó con la metodología de “focalizado”, por medio de una plataforma digital, donde las instituciones registraban su información. El proceso se apoyó en las entidades que las regulan, el Instituto Nacional Penitenciario y Carcelario (INPEC), el Instituto Colombiano de Bienestar Familiar (ICBF), el Ministerio de Defensa y la Agencia Colombiana para la Reintegración, entre otras. Esta información se capturó en un cuestionario análogo, a través de una plantilla disponible en la web.

Fuente: Texto editado a partir la Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe, del GTC CEA –CEPAL enviada a los países de la región durante el año 2020 para la elaboración del presente documento. Ver Anexo 2.4.

Por otro lado, se están desarrollando pruebas en Colombia, el Ecuador y el Uruguay tendientes a contrastar la información obtenida entre el siguiente censo y los registros administrativos con la esperanza de que los censos de esta ronda 2020 sean los últimos ejecutados de manera tradicional y que los posteriores se implementen aplicando una metodología no tradicional.

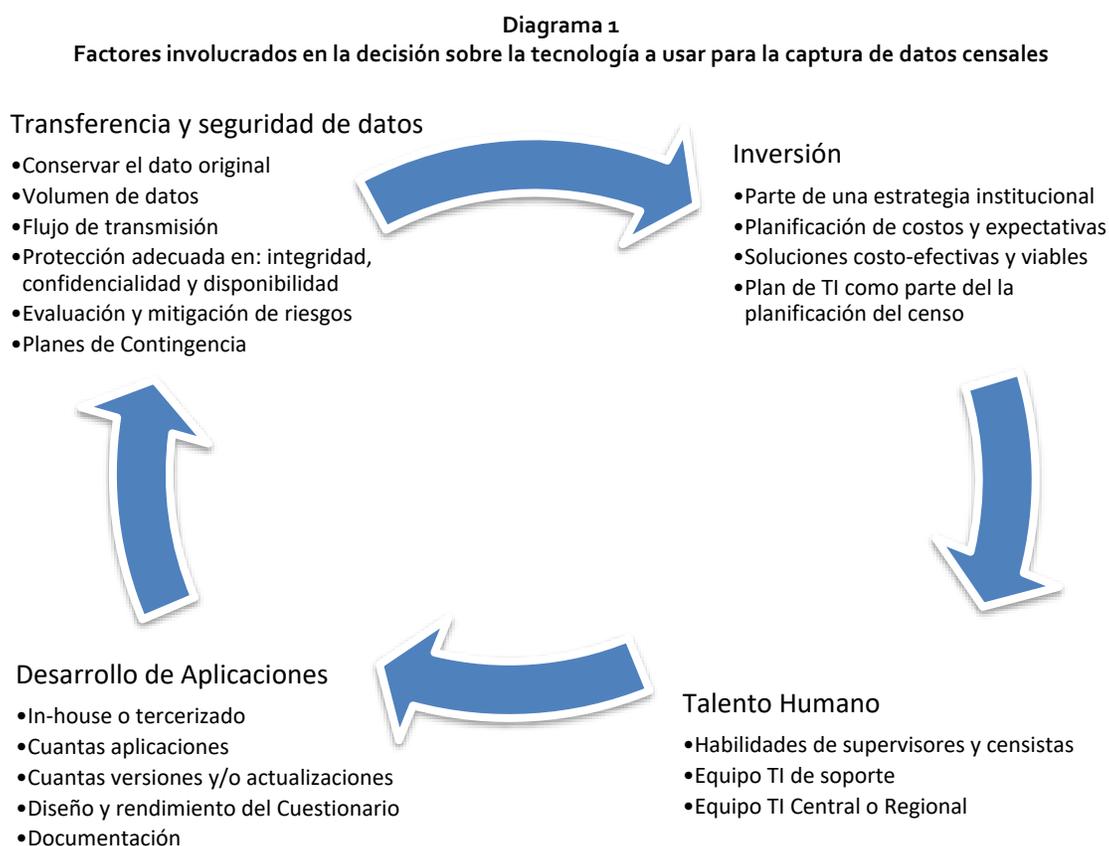
Un censo por registros administrativos corresponde a otra metodología censal (tema que no se aborda en esta publicación), en la que mediante algún identificador común se “capturan los datos” desde el registro administrativo, a nivel de registro individual de vivienda, hogar o persona, según corresponda. El proceso consiste en recopilar las características de las personas, hogares y/o viviendas desde un registro administrativo para fines, generalmente, no estadísticos. Este proceso de vinculación de registros administrativos con el censo va a depender, mayormente, de la robustez de tales registros en cada país.

Cabe enfatizar que el fortalecimiento de los registros administrativos debiera ser una política pública de mediano y largo plazo, teniendo como uno de sus objetivos centrales la reducción de la carga a los informantes y logrando que los próximos censos sean menos extensos, en volumen de preguntas y de temas a ser cubiertos. Sin embargo, hay quienes defienden la combinación de datos censales con los datos provenientes de las estadísticas vitales, incluso con la posibilidad de hacer la consistencia (y corrección) del censo; pero, toda vez que se debate esto —que en teoría podría ser perfectamente factible—, no aparecen ideas suficientemente claras sobre cómo implementar de manera pragmática dicha corrección.

Por otra parte, en la región se han implementado experiencias relevantes en el uso de registros administrativos para la construcción de directorios telefónicos y/o de viviendas, así como en la información de apoyo a los sistemas de seguimiento y control operativo, así como también en la etapa de procesamiento como fuentes complementarias. Por ejemplo, en Chile en el pre censo 2016 y en la actualización cartográfica para el censo 2023, se ha utilizado información de los registros de los certificados de recepción final de las viviendas³⁵, de las cuentas de agua y de la electricidad, como información de contraste a fin de evaluar la cobertura de viviendas para la actualización cartográfica en la etapa pre censal de 2016 y, actualmente, para priorizar zonas que se actualizarán en terreno.

En todos los métodos de captura empleados emerge, a mayor o menor escala, el uso del componente tecnológico, supeditado a la combinación que se decida entre el acceso a la población (forma de recopilación) y el método de captación de los datos (captura). Además, sujeto a múltiples factores que cada INE debe ponderar a fin de saber si el valor agregado —que aporta la innovación tecnológica para mejorar la calidad, oportunidad y cobertura—, es suficiente versus los costos y riesgos asociados. Esto se puede apreciar en el siguiente diagrama.

³⁵ Los Certificados de recepción final constituyen la autorización oficial de las direcciones de obras municipales para que una edificación (o conjunto de edificaciones) sea habitada o utilizada, los que se envían periódicamente al INE donde son georreferenciados.



Fuente: Elaboración propia a partir de los intercambios de experiencias del GTC CEA-CEPAL realizados en 2020.

La ponderación entre el valor agregado, los riesgos y los costos de incluir tecnología no indica que se le reste importancia al uso de mayor tecnología en el proceso de captura de los datos censales. Por el contrario, se trata de asegurar que su uso proporcione un mayor valor agregado al proceso censal. Debe comprenderse que la tecnología es una herramienta y no un fin en sí mismo y que, para su implementación de forma exitosa, es necesario considerar la variedad de factores que se tensionan entre sí, como se puede apreciar en el diagrama anterior.

En cuanto a los costos monetarios asociados, es muy difícil realizar comparaciones teóricas entre cada método de captura, ya que dependerán de la forma de implementación (doble digitación del 100% o del 10% de los formularios), los procesos circundantes que se incluyan (crítica y codificación), la formación y capacidades del equipo directivo, los tiempos en que se quieran presentar los resultados y del costo relativo de la mano de obra y la tecnología, entre otros elementos. Estos aspectos hacen variar los costos fijos y variables, por lo que cada país deberá analizar en su momento y según su contexto, cuáles son los pros y contras de cada método de captura en función de sus bondades, nivel de riesgo deseado y costos.

A. Captura digital durante el operativo

En las últimas rondas censales, la penetración del uso de tecnología, especialmente como método de recolección, ha trascendido al punto de reformular las etapas "clásicas" o la "secuencialidad" de las etapas censales, generando cierta confusión a la hora de definir roles y responsabilidades en las

actividades censales. De ahí que, se deba hacer un gran esfuerzo en la planificación, definición de los diferentes equipos de trabajo y su integración, como un elemento clave para la implementación exitosa al momento de usar alguno de los métodos de captura disponibles.

En la región se ha producido, principalmente, una transición desde el papel hacia dispositivos móviles de captura y —como ya se ha dicho—, se prevé que, en esta ronda, habrá más países que elijan este tipo de método. Algunos países cambiarán de metodología censal, desde el censo de hecho al de derecho, no solamente para la aplicación de este método de captura, sino que también a raíz del contexto que ha impuesto la pandemia y la posibilidad de resolver los imponderables de mejor forma.

1. Dispositivos Móviles de Captura (DMC)

La captura de datos con DMC es ejecutada, directamente, por los propios censistas durante el operativo, ya sea usando computadoras de mano (*palmtops*), tabletas, celulares o *netbooks* (modelos menores que los *notebooks* y *laptops*). Estos dispositivos, apoyados en programas instalados (*apps*) y hechos a la medida, almacenan en sus archivos las respuestas dadas por los entrevistados.

El uso de DMC brinda una serie de ventajas, entre ellas, la eliminación del uso del papel (y sus costos involucrados), el uso de dispositivos para guiar al censista en el flujo de preguntas (los “saltos” entre las preguntas) —eliminando la posibilidad de hacer preguntas que no correspondan (no se aplican) y disminuyendo la posibilidad de omitir preguntas específicas para ciertas subpoblaciones (fecundidad, educación, entre otras)—, apoyo a los censistas al mostrar textos de ayuda en la pantalla y permitir la programación de chequeos de inconsistencias entre las preguntas. Es decir, el uso de tales dispositivos ofrece la ventaja de una mayor calidad y una mejor oportunidad en la recolección del dato.

Sin embargo, los DMC también presentan algunas desventajas, entre las cuales se pueden mencionar: la necesidad de mayor capacitación a los censistas en el uso del DMC (algunos de ellos podrían, inclusive, no adaptarse totalmente al uso del dispositivo); los dispositivos no siempre se pueden usar, ya sea por razones de seguridad del censista en zonas peligrosas o por necesidad de cargar baterías de duración limitada —las cuales deben ser adquiridas previamente, en cantidad excedente debido a las pérdidas o desperfectos, al igual que los protectores y cargadores—; el costo de los dispositivos y de los accesorios podría alcanzar cifras elevadas dependiendo, además, de la definición técnica del tipo de censo, el número de empadronadores y del tiempo que se planifique para estar en terreno, entre otros inconvenientes que presentan estos aparatos.

Los aspectos positivos para utilizar este método de captura, mencionados por los países que llenaron las fichas (GTC – CEA, 2020), están relacionados con:

- Adecuación de enunciados de preguntas para distintos operativos.
- Inclusión de reglas de validación del tipo de variable, rango y saltos.
- Inclusión de reglas de validación entre preguntas.
- Inclusión de códigos y precarga de textos en ciertas preguntas.
- Mayor oportunidad en el monitoreo de datos.
- Posibilidad de georreferenciación vía GPS de la ubicación del censista y de la vivienda.
- Reutilización de dispositivos para otros productos estadísticos o estatales.

A lo señalado anteriormente, se puede agregar:

- Adecuación de preguntas a varios idiomas.

En cuanto a los desafíos que se presentan al momento de utilizar este método de captura, y que fueron mencionados por los países, se señalan los siguientes:

- Habilidades tecnológicas del personal de recolección, como censistas y supervisores.
- Infraestructura tecnológica de los INE.
- Asumir por defecto que no se requiere procesamiento de los datos.
- Asumir por defecto menores tiempos para presentar resultados.

A tales desafíos, se suman las siguientes situaciones:

- Censistas y supervisores calificados que presentan dificultades para adaptarse a las nuevas tecnologías.
- Adquisición de los dispositivos “*off the shelf*” o a medida.
- Desarrollo del aplicativo “*in-house*” o tercerizado para la captura.

En este sentido, es necesario explicitar que se deben diseñar e implementar mecanismos de supervisión sobre los DMC —por los procedimientos administrativos y normativas de cada país—, relacionados con el cuidado de los bienes estatales, considerando a la vez lo lento que puede ser el proceso de reemplazo de los DMC debido a las pérdidas por robos, daños o defectos de fábrica.

Los países que colaboraron con el llenado de fichas para la preparación de esta guía muestran un panorama diverso respecto del desarrollo del aplicativo, como se puede apreciar en el siguiente cuadro:

Cuadro 5
Características del desarrollo del aplicativo para DMC según países seleccionados

País y Censo	Brasil Censo 2010	Colombia Censo 2018	México Censo 2020 ^a	Venezuela (República Bolivariana de) Censo 2011
Tipo de Desarrollo	In-house	In-house	In-house	In-house
Software de Desarrollo	TotalCross		Delphi, Android Studio	CSPro 4.2
Sistema Operativo	Windows Mobile	Android	Android	Windows
Seguridad	Encriptación	Encriptación	Encriptación y cifrado de datos	Encriptación
Transmisión	Cable USB	Datos móviles	Cable USB y datos móviles	Cable USB
Aplicativos de Gestión			Administrador Censal Módulo Cartográfico Censal Kiosko INEGI ^b	Gestión en CRTG Gestión de Archivos Sala Situacional

Fuente: Elaboración propia a partir de las Fichas de Sistematización de Experiencias para Países de América Latina y el Caribe, del GTC CEA- CEPAL enviadas a los países para la elaboración del presente documento “Lineamientos generales para la captura de datos censales” durante el año 2020 que puede encontrar en el Anexo 2 de este documento.

^a Datos de preparación de su censo.

^b Instituto Nacional de Estadística y Geografía de México.

Nota: Las celdas vacías indican información no disponible o inexistente.

La selección de países en el cuadro anterior, para la sistematización de experiencias en el desarrollo del aplicativo para DMC, se debe a su reciente experiencia en la implementación de un método de captura a través de dispositivo móvil en un Censo de Población y Vivienda.

Así, el desarrollo de un aplicativo a través de un desarrollo propio (in-house) es lo que ha marcado estas primeras experiencias. El desarrollo propio de tecnología les permite a los países tener un control absoluto sobre las funcionalidades y usabilidad del aplicativo que se implementará. Por otro lado, actualmente, existe una diversidad de software para comercialización masiva (off the shelf) que resuelve varios de los desarrollos para la implementación del aplicativo de recolección de datos. Sin embargo, si se decide utilizar alguno de ellos, se debe considerar que la mayoría de dichos programas no ha sido implementada en ningún operativo a gran escala como lo es un Censo de Población y Vivienda; por lo cual, sigue siendo necesario llevar adelante la mayor cantidad de pruebas sucesivas previo a su implementación.

2. Transmisión de datos

La transmisión de datos debe permitir, de forma rápida y segura, el traspaso de información desde los DMC o las oficinas censales hacia los servidores centrales. Este procedimiento no solo permite una mayor oportunidad para las fases de procesamiento y análisis, sino que también para el seguimiento y monitoreo durante el operativo, permitiendo a su vez procesos de retroalimentación, recapitación y otras medidas previamente planificadas y/o soluciones a nuevos imprevistos.

La experiencia en la región indica que, usualmente, se utilizan dos métodos de transmisión de datos:

- i) A través de la red o datos móviles y,
- ii) A través de un cable desde un dispositivo de almacenamiento.

Ambas formas de transmisión pueden emplearse durante un mismo proceso de recolección de datos, considerando las características de conectividad que suelen tener las distintas áreas geográficas de un país. En ambos contextos, tanto los dispositivos como el sistema de transmisión deben garantizar la seguridad de los datos, de manera que no sean vulnerables.

De acuerdo con el Cuadro 5, los países usan métodos de encriptación, cifrado y claves de acceso durante la transmisión de datos. De todos modos, siempre es necesario realizar pruebas de integridad de datos para asegurar que la información generada en el dispositivo llegue íntegramente a los servidores centrales.

3. Dato primario

En cualquiera de las formas de transmisión de datos que se emplee se debe asegurar su encriptación para cautelar la privacidad de la información, junto con asegurar su integridad. Este proceso, si no ha sido bien probado, puede inducir a errores por la "transformación del dato" y la pérdida de información. Adicionalmente, un elemento clave será la posibilidad de "resguardar" el primer dato recolectado en el dispositivo, equivalente a lo que se hace con el cuestionario en papel. En algunos países, los datos primarios se guardan un cierto número de años por temas legales y, en la etapa de procesamiento, se tiene la posibilidad de consultar el cuestionario original. Es importante que se prevea una determinada forma de cautelar el primer dato, a manera de respaldo e inclusive de uso posterior, a la vez que sirva para el control de calidad y evaluaciones posteriores.

4. Aplicativos de Gestión

En general, al adoptar esta metodología de captura de datos, el énfasis principal está en el diseño e implementación correcta del formulario en el DMC. Sin embargo, un factor clave que se debe considerar es el o los aplicativos de gestión y monitoreo de la implementación y operativo censal elaborados dentro de una plataforma. Un ejemplo de esto es la plataforma OPERA empleada por el Instituto Nacional de Estadística y Geografía de México (INEGI), un sistema completo en el cual las operaciones estadísticas como el Censo 2020 entran a ser parte de este ecosistema, el que permite integrar funciones y tomar acciones para la correcta utilización del dispositivo, llevar a cabo otras operaciones y optimizar el uso de recursos.

Diagrama 2
Plataforma Opera del INEGI



Fuente: Diagrama presentado por el INEGI de México en las videoconferencias realizadas por la coordinación del GTC CEA-CEPAL para la revisión de las fichas de Sistematización de Experiencias para Países de América Latina y el Caribe, durante el año 2020.

B. eCenso (Autoenumeración o autoempadronamiento vía web)

El eCenso se relaciona con la posibilidad de responder las preguntas del censo de manera autónoma y sin la necesidad de un censista. En su versión de captura de datos en línea vía Internet (eCenso), el informante se conecta, directamente, al servidor de Internet que ha implementado el instituto para estos fines y responde el cuestionario.

Este método facilita la respuesta de los barrios u hogares “dormitorio”, donde la percepción de inseguridad es alta o donde no se permite el acceso a las viviendas ubicadas en zonas de estratos altos. Al ser los propios moradores quienes llenan sus cuestionarios, pueden organizar su tiempo para responderlo, incluyendo pausas. De esta manera, se evita tanto los errores de interpretación de los censistas como el contacto personal en tiempos de pandemia. Y, al igual que en los DMC, el programa dirige el flujo de las preguntas y su consistencia, permite versiones diferentes del cuestionario por idioma y diseños con enfoque diferencial para personas con capacidades diferentes.

Esta estrategia, como método complementario al método principal de recolección, es considerada altamente positiva, en especial para disminuir la omisión de personas, el número de rechazos —que se esperaría aumenten en un contexto de distanciamiento social exigido por la pandemia—, y mejorar la captación de las viviendas con moradores ausentes. Recientemente, del total de respuestas en el censo de México 2020, el 0.3% fueron por eCenso.

Otros aspectos relevantes, y a favor del uso de este método, son la reducción de los costos de equipos, bodegaje y distribución, en comparación con el uso de cuestionarios en papel o DMC. Así como también se disminuye sustancialmente la contratación y administración de recursos humanos, tanto

para el operativo como para la captura de datos¹⁶. En el recuadro a continuación, se muestran algunos resultados del eCenso de Colombia 2018.

Recuadro 2
Características de los resultados del eCenso 2018 de Colombia
<p>Porcentaje de personas y hogares censados mediante eCenso:</p> <ul style="list-style-type: none"> • Población censada: 11,73% • Hogares censados: 12,64% <p>Se obtuvo un promedio de 2,86 personas por hogar con el método de autoempadronamiento web, versus un promedio de 3,1 personas por hogar con los otros métodos de recolección (mayoritariamente entrevistas con DMC).</p> <p>Algunas características:</p> <ul style="list-style-type: none"> • Restricción de un hogar por vivienda • Estratos socioeconómicos 3 y 4 con mayor frecuencia de respuesta • Enfoque diferencial por discapacidad visual • Construcción del usuario (nombre, correo y clave), verificado con Registraduría para validar el documento. <p>Fuente: Texto editado a partir la Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe, del GTC CEA-CEPAL puede encontrar la ficha en el Anexo 2.4 de este documento.</p>

Por otro lado, este método exige desarrollar programas de captura de alta seguridad vía Internet, efectuar un control estricto de quienes responden por esta vía y seguir los casos que no responden, dejando entrevistas incompletas, lo cual puede requerir recursos humanos adicionales en el proyecto.

Igualmente, es necesario implementar formas de ayuda para los informantes, por ejemplo, definiciones dentro del sistema, así como un centro especializado de llamadas con personal capacitado que pueda guiar virtualmente a los encuestados para diligenciar el censo de manera apropiada. En esta modalidad, las personas se encuentran solas frente al cuestionario y les pueden surgir dudas durante el proceso.

A su vez, el uso de avances tecnológicos influye en la imagen de la institución, lo cual permitiría socializar con la población la posibilidad de autoenumerarse vía web o eCenso, generando confianza. Una buena estrategia, en este sentido, es usar los medios de comunicación y redes sociales para promover el método y mejorar la cobertura.

1. Diseño del formulario

La principal característica de este método es que es “autodiligenciado”. Esto significa que no existe un “censista” que oriente o guíe el llenado del cuestionario, de ahí que los requisitos funcionales del diseño del formulario deben estar enfocados hacia el informante y no hacia el censista. Esto es importante porque, en el diseño del formulario para DMC y para el eCenso, se debiera tomar en cuenta que tienen objetivos distintos y que están dirigidos a poblaciones objetivo diferentes. Afortunadamente, existen muchas herramientas y tecnologías de diseño e implementación de formularios para autoenumeración como JavaScript y CSS, entre otros, habiéndose incrementado sus usos y opciones de implementación, como consecuencia de la pandemia por COVID-19.

¹⁶ Para más información véase el capítulo III.4.d

2. Seguridad

En este método de recopilación en línea o eCenso, la seguridad de la información es controlada de forma centralizada. Para ello, es deseable que el INE y el personal de TI tengan experiencia previa y habilidades en la utilización de estas tecnologías. Además, la tecnología utilizada debe ser acorde con los niveles de seguridad, garantizando la continuidad operacional del sistema, sin poner en riesgo ni la credibilidad del censo ni de la institución. Es importante establecer un control de acceso que sea robusto, al no existir control sobre quién ingresa o desde dónde se ingresa, así como garantizar la seguridad de la infraestructura y las bases de datos. Por ejemplo, el censo de Chile 2017 tuvo un sistema aislado o encapsulado del resto de la infraestructura tecnológica del INE para evitar ciberataques, cautelando accesos indebidos a su sitio. Por último, hay que señalar que la imagen institucional y del censo se pueden dañar si se producen accesos indebidos, por eso es recomendable la asesoría de instituciones estatales especializadas en el tema.

3. Conectividad

En cuanto a la conectividad, debe contarse con garantías tecnológicas máximas para disponer de servidores de red capaces de satisfacer las demandas durante los horarios de congestión, especialmente en la primera y últimas semanas del censo cuando se produce mayor concurrencia. Hay que destacar que una buena conectividad impacta positivamente en el presupuesto del censo, por lo que algunos países optan por manejar este riesgo. Por último, tener un sitio web que no responda o sea "lento" puede tener consecuencias comunicacionales y repercusión que dañen la confianza en el censo y en el INE.

4. Estrategia de contacto a la población

Es necesario definir previamente cómo se va a contactar a la población; es decir, cuál será la forma de entregar la clave única que permita relacionar una vivienda específica con el cuestionario web, dado que en este método es un desafío vincular la vivienda al territorio. Por ejemplo, en el Censo de Chile 2017, los censistas cuando encontraban una vivienda con moradores ausentes dejaban una "citación" con un número único de identificación de la vivienda (con una identificación del portafolio y número de la vivienda) y con este código los moradores quedaban habilitados para autoenumerarse vía web (eCenso) o dirigirse a una dependencia del INE para hacerlo personalmente. En el caso del censo de México 2020, la carta de invitación para responder en la web estaba asociada a la lista o nómina de inmuebles y se entregó, solamente, si la persona lo pedía o después de la tercera visita.

C. Captura desde papel

La entrevista cara a cara con cuestionarios en papel representa el método más tradicional que los INE han usado en la región para la recolección censal. Con este tipo de entrevista, se tiene relativa experiencia en la digitación o lectura óptica (escáner) como método de captura. Lo importante de estos métodos es que están plenamente definidos y se ejecutan posteriormente al operativo, sin estresar tanto el operativo censal en terreno –dado que el manejo, la revisión, el empaquetado y la entrega de los cuestionarios impresos al centro de captura implican mayor carga de trabajo para la estructura operativa– y pudiendo hacerlo con debida calma, siempre y cuando los tiempos vinculados a la entrega de resultados lo permitan.

En estos métodos, existe un factor de riesgo asociado al retorno de las boletas censales al lugar de acopio o bodegaje. Si existe alguna pérdida de material, se debe volver a ejecutar el operativo, algo que con los métodos digitales no sucede necesariamente, pues el dato está disponible en formato digital prácticamente de forma inmediata en los servidores centrales. Uno de los mayores riesgos con los DMC es que la información no haya sido transmitida oportunamente. Un ejemplo de esto fue lo ocurrido con el Censo de Chile 2017. Una vez concluido el operativo en terreno, los cuestionarios se encontraban en

una bodega regional, para luego ser trasladados a la oficina central, lugar donde serían escaneados. Lamentablemente, producto de las lluvias en la cordillera del Desierto de Atacama, se generó un aluvión de gran magnitud, el cual arrasó con los cuestionarios de cuatro municipios, obligando a repetir un nuevo operativo censal en las semanas posteriores, con el habitual impacto que ello significa en términos de recursos humanos, financieros y en la calidad de los datos.

1. Digitación

La digitación consiste en ingresar a mano los cuestionarios censales a una base de datos mediante el teclado y se clasifica en dos tipos: numérica y alfanumérica. La numérica implica que se usa solamente el teclado numérico; por lo tanto, las preguntas abiertas (migración, ocupación, rama de actividad y otras) deben haber sido previamente codificadas (también manualmente). En la digitación alfanumérica, las preguntas abiertas son capturadas en forma textual, lo más similar a como se respondieron en terreno. En una etapa posterior, dichas respuestas son codificadas mediante un proceso semiautomático.

En general, la digitación debe entenderse como el proceso por el cual se ingresan todas las respuestas del cuestionario lo más aproximado posible a como están en el papel, por lo que en este proceso de digitación no se requiere efectuar ninguna consistencia en los datos entrantes, más allá de la verificación de los rangos de las variables.

Lo cierto es que la digitación se utiliza cada vez con menor frecuencia en la región. En la década de 2010, siete países usaron este método de captura y, si bien, en México se realizó la digitación de los cuestionarios captados en papel, este método no fue el medio principal para la captación de información, por lo que, en la presente década (2020), ningún país ha informado que realizará la captura por medio de la digitación.

Entre las principales razones para migrar hacia métodos más tecnológicos están las desventajas de la digitación en términos de calidad y oportunidad. Sin embargo, a veces los países la consideran como una alternativa válida o como plan de contingencia frente a dificultades o problemas con los otros métodos, dado que su ejecución es relativamente sencilla, a un bajo costo y, además, se tiene experiencia previa en su implementación.

Hoy, estos argumentos han perdido validez en forma diferenciada según los países. Por ejemplo, la argumentación de ser un método de captura de menor costo en relación con los otros ya no es tan real, producto de un mayor desarrollo, del mejoramiento en los salarios y en las condiciones laborales y de la disminución de los precios de la tecnología, entre otras razones.

Desde otra perspectiva, la alta rotación laboral, que afecta frecuentemente a los INE de la región, hace que no exista personal suficiente con la experiencia previa en el proceso y la logística requerida, especialmente por las crecientes exigencias administrativas y normativas que requieren de espacio físico y de la contratación de un gran volumen de colaboradores.

Por otro lado, las ventajas más destacables del método de captura por digitación son:

- requiere bajo nivel de profesionalización;
- no demanda mayor tecnología, solo terminales o computadoras de escritorio simples y de bajo costo;
- se pueden utilizar software de ingreso de datos gratuitos o de bajo costo; y,
- los procesos son secuenciales, lo que simplifica la planificación y las habilidades gerenciales requeridas para el proceso, facilitando su implementación y la realización de la captura con mayores niveles de certidumbre en aquellos INE de menor tamaño, con poca infraestructura tecnológica y bajo nivel de profesionalización.

Otro aspecto que se debe analizar con mucho cuidado a la hora de aplicar este método de captura es el mecanismo de validación del ingreso de datos, pues afectará considerablemente los costos, la productividad y la duración de la digitación. Lo recomendable es que se haga doble digitación; es decir, que un mismo cuestionario sea digitado dos veces y, luego por procesos de validación automatizada, se opte por el "correcto". Algunos países de la región han innovado sus mecanismos de validación de ingreso de datos utilizando un proceso de doble digitación por muestreo, el que se inicia con la verificación del 100% de los cuestionarios y se va descendiendo, conforme va avanzando la curva de aprendizaje y consolidándose los niveles de error, hasta llegar a un porcentaje de verificación en torno al 10 o 5% al finalizar la digitación del censo. Otra alternativa es verificar solamente las preguntas categóricas, dado que las preguntas textuales se usan solamente como insumo para la codificación asistida. En el cuadro siguiente se observa el porcentaje de validación usando doble digitación en algunos países con experiencias relativamente recientes.

Cuadro 6
Uso de doble digitación

País	Año del Censo	% Validación (doble digitación)
Cuba	2012	100
El Salvador	2007	80
Guatemala	2018	10
Honduras	2013	30
Paraguay	2002	100

Fuente: Elaboración propia, sobre la base de información entregada por los países.

Asimismo, cabe señalar que, al momento de decidir el tipo de metodología censal, se debe considerar si el programa de ingreso de datos se diseñará solamente para ser un "espejo" del dato en papel o si se implementarán validaciones (rangos, flujos), tal como se desarrolla cuando se utiliza DMC o eCenso. En este punto, es recomendable aprovechar el ingreso de información para incorporar ciertas validaciones, sean estructurales u otras, con la posibilidad de detección automática y preventiva de errores, a fin de identificar si el error es producto de la digitación o viene del operativo. La desventaja, en la incorporación de validaciones, es la gran dificultad de estandarizar la aplicación de criterios para la corrección de errores en los procesos masivos de digitación.

De esta manera, la incorporación o no de validaciones y la doble digitación están altamente vinculadas con el factor tiempo, condición que siempre muestra una fuerte tensión entre los tiempos técnicos y políticos, especialmente si este método no fue parte de la planificación inicial, sino de la contingencia.

Quizá el factor más relevante al considerar la digitación como método de captura es la intensidad del trabajo realizado manualmente, agregando una fuente de error adicional. Es justamente en la digitación donde el factor humano está más presente, no solo por la digitación en sí misma, sino que también en diferentes actividades manuales vinculadas al proceso de digitación. Esto influye en bajar la calidad del proceso, además de su lentitud si se le compara con los otros métodos, por lo que a veces se exige la contratación de más personal para tratar de alcanzar la velocidad de las otras alternativas de captura.

La historia de la región ha mostrado una propensión a iniciar con retraso la digitación, especialmente cuando el equipo de profesionales o técnicos que lidera el censo está ocupado en el cierre del operativo y la logística necesaria para recopilar todo el material censal, afectando tanto los plazos para la contratación de personal (digitadores, supervisores y codificadores) y su capacitación, como también la adecuación del espacio necesario para la digitación y el archivo censal o la compra de equipos, entre otros efectos negativos. Con frecuencia, todo esto ha resultado en intentos para recuperar el tiempo perdido, aumentando los turnos (con personal menos entrenado), presionando para

mayor productividad, reduciendo (o eliminando totalmente) la doble digitación, todas medidas que, invariablemente, sacrifican la calidad de la información censal.

En dicho contexto, resulta necesario explicitar también que los países no siempre acostumbran a publicar los resultados de sus procesos de doble digitación, lo cual serviría, entre otras cosas, para medir con un mayor asidero la calidad o los errores provenientes de la captura manual. De este modo, se podría evaluar cuánto afecta específicamente la digitación en la calidad, pues la digitación termina soportando, en muchas ocasiones, todos los errores que surgen en las etapas anteriores a la edición e imputación, como aquellos errores provenientes del llenado de los cuestionarios en terreno, los errores debidos a una capacitación deficiente o un mal diseño del formulario. En tal sentido, esto se puede considerar, equivocadamente, como una "ventaja" del método de digitación, al asumir toda la culpa de una mala calidad en el proceso de captura y no reconocer que los problemas estaban radicados en las etapas previas.

2. Reconocimiento óptico (Escaneo)

En este método de captura, todas (o al menos la mayoría) de las páginas del cuestionario en papel son transformadas en una imagen por medio de máquinas de lectura digital (escáner) y, luego, esas imágenes pasan por una etapa de interpretación automática, que las convierte en marcas, números y caracteres. Cuando la interpretación no es posible por algún problema en el llenado del cuestionario, el sistema las envía a una estación de verificación manual para dirimir dudas. Este tipo de reconocimiento óptico es, por tanto, una combinación entre Reconocimiento Óptico de Marcas (OMR, *Optical Mark Recognition*), usado en censos desde los años 80, y Reconocimiento Óptico de Caracteres (OCR, *Optical Character Recognition*) es decir, letras y números, utilizado para los censos de la ronda 2000. Esta combinación de técnicas OMR y OCR se conoce también en la literatura especializada como Reconocimiento Inteligente de Caracteres o ICR (*Intelligent Character Recognition*) y como Reconocimiento Inteligente de Marcas o IMR (*Intelligent Mark Recognition*).

Además de las variables no interpretadas por los algoritmos de escaneo, puede que una variable vaya a la estación de verificación visual, cuando se usan otros controles de consistencia, como en el caso de los rangos para la edad. Aunque se acepten valores hasta 130 años, se puede usar un límite más estricto, por ejemplo 80 años, para evitar valores altos poco probables. Lo mismo se puede hacer con el número de hijos, donde se recomienda que el límite aceptable sea 30 hijas o hijos, a pesar de que se rechazaría si se interpreta arriba de 10 hijas o hijos. Evidentemente, no es posible tener un límite muy bajo, porque aumentaría mucho la carga de trabajo de la verificación. De ahí que, una de las claves de éxito de este proceso es tener un buen balance y afinamiento en la cantidad de las alertas disparadas por el sistema. Muchas alertas con "falso positivo" aumentan, innecesariamente, el trabajo de los verificadores. Mientras que pocas alertas ("falso negativo") aumentan la posibilidad de que pasen errores que, posiblemente, podrían ser atrapados y corregidos en la verificación. Por lo tanto, es crucial haber efectuado pruebas suficientes que permitan conocer el tipo de error al que se van a enfrentar los verificadores en el proceso de escaneo.

El escaneo tiene diversas ventajas sobre la digitación, entre ellas:

- requiere de menos personal y, por ende, se puede tener personal más capacitado en menos espacio físico;
- es mucho más rápido y se generan copias digitales de los cuestionarios, lo cual puede significar un ahorro de espacio para almacenar los cuestionarios, además de una mayor preservación de los mismos; y,
- posibilita la captura de campos adicionales, que tradicionalmente no se digitan, como la lista de integrantes del hogar sin mayores costos monetarios o de tiempo.

Consecuentemente, todos estos factores redundan en una mayor calidad en la captura de datos, ya que involucra un grado mucho mayor de estandarización y menor incidencia del factor humano, el cual actúa solamente en la estación de verificación de los casos donde la interpretación del escaneo no fue capaz de “traducir” la imagen en un código, número o letra apropiado. Podría considerarse, incluso, que los campos categorizados (sexo y parentesco, entre otros) no tendrían errores en la interpretación de las marcas, porque generalmente se obtiene un porcentaje igual o superior al 99% de aceptación.

Recuadro 3

Ecuador, fases y actividades del próximo censo de población y vivienda¹⁷

En el Ecuador se han establecido las siguientes fases y actividades para el procesamiento asociado a la captura de los datos censales:

Archivo inicial

- Recepción de la información enviada luego del empadronamiento
- Ordenamiento y ubicación de las estanterías

Pre-Crítica

- Separación de cuestionarios en blanco y otros materiales
- Revisión y corrección de la identificación geográfica de cada cuestionario
- Revisión y ordenamiento de los cuestionarios por área de empadronamiento y vivienda en orden ascendente
- Revisión de la estructura vivienda-hogar-población

Guillotina

- Corte del lote censal en la línea de corte (impreso en cada cuestionario censal)

Escaneo

- Escaneo de cada lote censal (sector censal), no se deben escanear sectores unidos o parciales
- Verificación en el monitor del sistema que el escaneo sea completo
- Verificación que las imágenes de cada sector censal vayan a su correspondiente carpeta

Interpretación

- El software especializado interpreta el contenido de las imágenes y transforma en valores numéricos o alfanuméricos y almacena en una base de datos

Verificación

- Verificación de las variables interpretadas (números y caracteres alfanuméricos)
- Corrección manual (teclado) entre la imagen de la variable y el contenido en la base de datos

Control de Cobertura

- Corrección de identificaciones de cuestionarios (ubicación geográfica errada o diferente a la base pre censal)
- Matching o pareo entre la base pre censal y la base censal a niveles geográficos (zona, sector y áreas de empadronamiento)
- Re-escaneo de sectores con errores

Codificación (Porcentajes)

- Codificación automática de las variables con respuesta alfanumérica (geográficas y económicas)
- Codificación asistida de las variables que no pudo realizar la codificación automática

Fuente: Texto editado a partir la Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe, del GTC CEA-CEPAL. Véase Anexo 2.6 de este documento.

Entre las desventajas, habitualmente se menciona el costo total, aunque existen corrientes que advocan que el proceso es más barato que la digitación, dados los costos de la mano de obra local -no se requiere contratar digitadores-, instalaciones más pequeñas, se evita introducir nuevos errores producto de la digitación, y mejora la oportunidad, entre otras. Otra desventaja mencionada es la necesidad de mejorar el entrenamiento de los censistas, quienes deben tener mucho cuidado en el llenado completo de las marcas, en escribir bien los números y en la caligrafía mayúscula de las letras. Además, se argumenta que no funciona en zonas rurales, a pesar de que el escaneo ya se usó de forma

¹⁷ El INEC se encuentra en un proceso de análisis para determinar la nueva fecha del censo debido al impacto que ha tenido la pandemia por COVID-19 en las tareas pre censales.

exitosa en varios países con condiciones adversas, incluyendo zonas rurales y lluviosas, y con censistas que eran estudiantes de nivel medio.

Por otro lado, el escaneo necesita de un papel de mayor calidad, con un gramaje y opacidad adecuados que no se encontrarían en imprentas locales con la calidad y rapidez necesaria. Por último, el escaneo requiere que, en terreno, se mantenga una gestión adecuada de los cuestionarios, porque podrían aumentar los porcentajes de rechazo en la lectura, la que debe encontrarse dentro de los márgenes aceptables calculados para la productividad.

A continuación, se presenta el proceso de escaneo que ha diseñado el Instituto Nacional de Estadística y Censos (INEC), del Ecuador, para su próximo censo.

Diseño y contenido de la boleta censal

La utilización del método de captura por reconocimiento óptico o escaneo implica ciertos ajustes a la hora del diseño de la boleta censal, adecuaciones que permitan la correcta interpretación por parte de las máquinas lectoras, las que además generan discrepancias, en ocasiones, dentro del equipo censal a la hora de decidir el diseño óptimo del cuestionario. Por ejemplo, es deseable (según las necesidades de información) minimizar el uso de respuestas abiertas ("Otro, ¿cuál?"), porque impactan negativamente en el operativo y en el proceso de captura al exigir una codificación posterior, con una ganancia muy cuestionable sobre las posibles existencias de otros tipos de categorías en esas preguntas. Además, se sugiere no llenar los campos de valores con ceros (o) a la izquierda, debido a que provocan errores de confusión con otros números (6, por ejemplo), impactando desfavorablemente en la calidad final de la lectura.

En la transición del método de digitación al de escaneo, se sugiere reformular los códigos de localización geográfica con códigos jerárquicos (región/departamento, provincia, municipio/comuna, distrito, zona censal) hacia códigos secuenciales con 7 u 8 dígitos. Una de las ventajas que tiene el escaneo es que se puede substituir esa "caravana" de dígitos por una identificación única en la boleta con el número de dígitos suficiente y máximo compatible con el número de viviendas (boletas) del censo, esto además ocupa menos espacio en el cuestionario. Otra recomendación es el uso de códigos de barra o QR para la identificación; sin embargo, estas medidas no son suficientes y el proceso debe apoyarse con la localización detallada.

Identificación de las boletas de continuación y número de personas

La identificación de las boletas de continuación¹⁸ y número de personas son temas que deben ser analizados e implementados detenidamente, debido a que puede resultar complejo en la fase de procesamiento: el enganche de boletas de continuación en los casos en que la vivienda tenga más de un hogar y/o que el hogar tenga más personas que el número soportado en la boleta. En este sentido, la recomendación es tener una marca que especifique claramente si la boleta es de continuación e insistir en tener la información correcta de la identificación o número de serie del cuestionario general. Una forma eficiente para disminuir la existencia de boletas de continuación es con un número adecuado de hojas de personas por boleta. Este número dependerá de la cantidad de preguntas para el módulo de población, del promedio de personas por hogar de censos anteriores y de encuestas de hogares recientes. Además, debe lograrse el equilibrio, a través de la simulación de escenarios que den cifras óptimas, entre el número de páginas, la cantidad de cuestionarios que se deben imprimir y, obviamente, los costos asociados a ello.

¹⁸ Boletas de continuación son aquellas que necesitan usarse para las viviendas que tengan más de x personas, en donde x es el número de personas que se puede anotar en cada boleta.

Por ejemplo, en el último censo de Chile (2017), se agregaron más personas al cuestionario censal pasando de 6 (censo 2002) a 8 personas por cuestionario. Esto permitió reducir, de forma considerable, la utilización de cuestionarios de continuación del hogar, llegando a cubrir el 99.3% de las personas con un cuestionario que logra abarcar hogares de hasta 8 personas, como se puede observar en los cuadros 7 y 8:

Cuadro 7
Porcentaje de población cubierta según el número de personas por cuestionario,
en censos seleccionados de la década de 2010
(En porcentajes)

Personas por cuestionario	Argentina 2010	Bolivia 2012	Chile 2017	Costa Rica 2011	Ecuador 2010	Guatemala 2018	Perú 2017	Rep. Dominicana 2010
1	17,6	21,1	17,7	11,2	12,1	6,1	17,1	14,7
2	40,2	38,8	41,7	30,6	28,0	17,9	34,9	32,7
3	59,9	55,7	64,4	53,9	48,0	35,5	54,9	52,6
4	78,2	71,5	83,2	76,7	69,1	56,1	74,3	72,0
5	88,9	82,9	92,8	89,7	83,5	72,7	86,5	86,4
6	95,1	90,0	96,8	95,6	91,2	83,2	94,3	94,6
7	97,3	94,3	98,5	97,9	95,2	89,7	96,8	97,3
8	100,0	98,2	99,3	99,0	97,4	94,2	98,2	98,7
9		98,9	99,6	99,5	98,6	96,4	99,0	99,3
10		99,3	99,8	99,7	99,5	97,9	99,5	99,7

Fuente: elaboración propia a partir de procesamiento de los censos en línea con la plataforma Redatam que disponen los países para el procesamiento, análisis y disseminación de microdatos censales (<https://redatam.org/es/procesar-en-linea>).

Cuadro 8
Páginas por cuestionario censal según módulo de preguntas y total,
en censos seleccionados de la década de 2010

País y año del censo	Páginas por temática censal						Total contenido de temática	Información general para el censista y/o censado	Total páginas del cuestionario
	Vivienda	Hogar	Listado de personas	Personas					
				Total personas	Cantidad de páginas por persona	Cantidad de personas por cuestionario			
Chile 2017	1	-	1	8	1	8	1+1+8=10	0	10
Ecuador 2010	2	1	1	20	10	10	2+1+1+2*10=24	0	24
Guatemala 2018	1	1	1	16	2	8	1+1+1+2*8=19	1	20
Perú 2017	1,5	0,5	1	12	2	6	2+1+2*6=15	1	16

Fuente: elaboración propia partir de las boletas censales recopiladas en <https://www.cepal.org/es/temas/censos-de-poblacion-y-vivienda/enlaces-institutos-nacionales-estadistica-america-latina-caribe>.

D. Ventajas y desventajas de los diferentes métodos de captura

Con la finalidad de contribuir a la toma de decisiones respecto del método de captura que debe elegir cada país, se propone considerar las principales características relativas de cada método utilizado en la región, teniendo en cuenta que las mismas se pueden transformar en potenciales ventajas y desventajas en función de su implementación, seriedad en la evaluación y consistencia entre los procesos. Es decir, por ejemplo, no basta solo con utilizar DMC para asegurar una mejor calidad en los datos censales, sino que este debe ser correctamente implementado en cada uno de sus procesos previos, simultáneos o posteriores, tanto en términos técnicos, como de gestión y de administración. En cambio, podría ser mayor la calidad de un censo en papel que sea bien implementado. En el siguiente cuadro se pueden apreciar, comparativamente, las características de cada método de captura empleado en la región de América Latina y el Caribe.

Cuadro 9
Características relativas entre los métodos de captura

Características	Digitación	Escaneo	DMC	eCenso
Implementación técnica	Menos compleja que con escáner	Más compleja que digitación	Compleja	Más compleja que digitación, pero menos compleja que con DMC
Sofisticación en el diseño e impresión del cuestionario	Menos sofisticada	Más sofisticada	No aplica	No aplica
Perfil de los censistas	Alfabetizados con nivel académico de los últimos años de secundaria	Alfabetizados con nivel académico de los últimos años de secundaria	Requiere de censistas con un mayor nivel de alfabetización digital	No aplica
Perfil del censado	Informante calificado	Informante calificado	Informante calificado	Informante calificado con un mayor nivel de alfabetización digital y acceso a internet
Errores en la captura	Mayores al escaneo	Cercanos a 0 en marcas y menores en alfanuméricos	No hay transcripción posterior	No hay transcripción posterior
Estandarización en los criterios de captura	Criterios de digitación particulares según cada digitador	Mayor estandarización que en digitación	Mayor estandarización que en papel	Mayor estandarización que en las tres anteriores, porque no hay una tercera persona involucrada en la captura
Cuidado del cuestionario en terreno	Bajo	Mayor al de digitación	No aplica /Más cuidado con el DMC y seguridad personal	No aplica
Tiempo entre el cierre definitivo del cuestionario y el censo	Extenso, pero menor que en escáner	Extenso, mayor que en digitación por premura en cerrar el cuestionario por tareas de impresión	Menor que en papel, el cuestionario puede ajustarse muy próximo al censo (y si es necesario, a veces durante el censo mismo, solo por excepción)	Menor que en papel, el cuestionario puede ajustarse muy próximo al censo (y si es necesario, a veces durante el censo mismo, solo por excepción)
Logística de los cuestionarios y materiales de ida y vuelta	Riesgo de pérdida del material antes de digitarse	Riesgo de pérdida del material antes de escanearse	Menor en volumen, pero compleja por seguridad (daño y robo)	No aplica
Desarrollo e implementación del servicio	Interno	Externo	Preparación de los programas de captura con mucha antelación	Preparación de los programas de captura con mucha antelación

Características	Digitación	Escaneo	DMC	eCenso
Chequeo de los saltos lógicos	Inexistentes	Inexistentes	Aplicación de los "saltos" estructurales para dirigir al empadronador	Aplicación de los "saltos" estructurales para dirigir al encuestado
Organización y control de la producción (talento humano, mobiliario, logística, etc.)	Compleja por el gran volumen de personas, turnos, espacio físico, mobiliario y manejo del archivo censal	Menos compleja que la digitación, al ser un menor volumen de personas, espacio, y manejo de archivo censal más simple	Compleja al juntarse la forma de recolección con el método de captura en una misma acción	Compleja por el seguimiento de cobertura y falta de respuestas al cuestionario
Equipo de cómputo y mobiliario	En gran cantidad, pero puede ser utilizado por el instituto y el Estado después del censo	No es relevante en cantidad	En gran cantidad, pero puede ser utilizado por el instituto y el Estado después del censo	No aplica
Almacenamiento y seguridad de los cuestionarios	Costoso, de gran volumen y frágil ante riesgos humanos y naturales	Las imágenes son un medio más seguro de respaldo de la información	No aplica	No aplica
Tiempo de producción	Depende del número de digitadores, generalmente entre 6 a 12 meses	Depende del número de escáner y cantidad de personal en la verificación, generalmente 4 a 6 meses.	Más rápido en la producción de resultados, dado que se elimina la etapa de transferencia del papel a un medio digital y también se reduce la cantidad de inconsistencias	Muy rápido en la producción de resultados, dado que adelanta la etapa de captura
Exigencia de tiempos legales de resguardo	Alto, frecuentemente 5 años con un alto costo de bodegaje	Si la legislación nacional lo permite, se reduce el costo al guardar solo las imágenes. De lo contrario es igual a digitación.	A muy bajo costo	A muy bajo costo
Recursos necesarios para el manejo del archivo censal (físico), durante el proceso de digitación.	Importantes y mayores al escaneo	Bastante menores a la digitación	No aplica	No aplica
Diferentes Idiomas del cuestionario	Si, pero no estarán siempre disponibles	Si, pero no estarán siempre disponibles	Siempre disponibles en múltiples idiomas	Siempre disponibles en múltiples idiomas y con diseño especializado para personas con capacidades diferentes
Pilotaje	Requiere menos pilotajes que las otras alternativas	Requiere más pilotajes que la digitación, pero menos que los otros métodos	Requiere de más y complejas pruebas piloto, permite explotar y analizar los datos obtenidos de las pruebas.	Similar a DMC sumando el tema de seguridad
Control de la No-respuesta	No aplica, se debe ingresar en el procesamiento	No aplica, se debe ingresar en el procesamiento	Se puede tener un mayor control de la no-respuesta que en papel por la posibilidad de implementar durante el operativo	Se puede tener un mayor control de la no-respuesta que en papel, pero menor que con DMC, porque la responsabilidad final de responder recae en las personas
Transmisión de los cuestionarios a la central	No aplica	No aplica	Sofisticación en la transmisión de los cuestionarios a la central	No aplica
Localización de la Vivienda	No aplica	No aplica	Captura de la localización geográfica de las viviendas (y del censista)	Difícil, sin observación y/o verificación. Exige un control de cobertura muy sofisticado

Fuente: Elaboración propia a partir de los intercambios de experiencias del GTC CEA-CEPAL realizados en 2020.

E. Combinación de métodos de captura ¹⁹

En el intento de garantizar una buena cobertura y minimizar la tasa de no-respuesta en el censo, habitualmente, se hace necesario considerar el uso de más de un método de captura, de forma simultánea y secuencial. En la combinación simultánea de métodos, entre otros, se puede ofrecer al usuario la opción de una de las alternativas para responder al censo, a través de un método presencial o un cuestionario en línea. En la combinación de métodos secuenciales, se empieza por uno de los métodos, como la entrevista presencial y, en caso de no ser posible dicha entrevista —por la ausencia de los informantes, por ejemplo—, se entrega la invitación para el uso de la internet, como se efectuó en el reciente censo de México 2020 y como se está planificando para los censos de el Brasil (2022) y Chile (2023). O al revés, se entregan citaciones a los informantes (por medio de los mismos entrevistadores o correo regular) para que contesten el censo por internet. Pasado un tiempo, si no se recibe la entrevista completada, se avanza al siguiente método, tal como lo realizó Colombia (2018) y la República Bolivariana de Venezuela (2021).

Otra manera de explicar los modelos paralelo y secuencial es que, en el diseño secuencial, por ejemplo, se debe elegir un método principal, el que más se adecúe a la gran mayoría de la población. Este método debería usarse en su máxima potencialidad, mientras que los otros métodos pueden ser usados como opciones auxiliares o para grupos específicos de la población, como se ha diseñado para el censo del Brasil 2022, donde el método de recolección principal es con DMC y de forma complementaria se utilizará el eCenso. Por otro lado, en el diseño paralelo o concurrente, no hay un método dominante, sino que todos son tratados por igual. En general, el diseño secuencial tiene efectos menos marcados sobre la diferencia de calidad entre los métodos (*mode effect*). A continuación, se presenta un diagrama para resumir las características del modelo mixto.

Diagrama 3
Modelos de recolecciones de datos censales



Fuente: Elaboración propia a partir de los intercambios de experiencias del GTC CEA-CEPAL realizados en 2020.

¹⁹ Este acápite se basa en una adaptación libre y resumida del documento United Nations (2019) “Guidelines of the use of electronic data collection Technologies in population and housing censuses, chapter E. (2019), Department of Economic and Social Affairs, United Nations Statistics Division, (UNSD), New York, January.

Por otro lado, se puede considerar también la aplicabilidad de métodos dirigidos a grupos específicos, sean determinados por la accesibilidad (área urbana o rural, zonas de difícil acceso geográfico, viviendas ocupadas con moradores ausentes), por condiciones socioeconómicas (acceso a internet) o por limitaciones, incluso de seguridad (zonas de alta peligrosidad o condominios protegidos con un alto grado de dificultad de acceso a los censistas).

En la ronda censal de 2020, se ha observado que los países de la región están eligiendo y planificando la combinación de métodos, manteniendo un método de recolección presencial principal (papel con escaneo o DMC) combinado con autoempadronamiento vía internet (eCenso).

El uso de una combinación de métodos de recolección tiene muchas ventajas y puede ser muy efectivo en el sentido de disminuir la no-respuesta y aumentar la cobertura, pero hay que considerar también algunas dificultades y/o desventajas, tales como:

1. Desarrollo

La existencia de más de un método de recolección presupone la necesidad de desarrollo y programación de más de un sistema; es como si hubiera dos o más censos, los que, además, se interconectan. La sincronización y unificación de la información proveniente de los distintos métodos constituye un desafío desde el punto de vista del desarrollo, más aún si se pretende realizar un análisis del monitoreo de la información durante el operativo que permita generar alertas y desencadenar los planes de contingencia predefinidos.

2. Parcialidad (mode effect o bias)

Es la entrega de resultados disímiles como consecuencia de usar métodos distintos de recolección. Es decir, las diferencias observadas en los datos pueden atribuirse a cómo el dato ha sido recolectado y no a características reales existentes en la población. No obstante, esta situación se puede minimizar si se ha considerado dentro de la planeación. Existen tres factores relativos a la calidad de la información y asociados al *mode effect*, a saber:

- i) cobertura;
- ii) tasa de respuestas; y,
- iii) calidad de las respuestas en preguntas sensibles.

Estas diferencias deben ser cuidadosamente estudiadas y, dentro de lo posible, propender a que sus efectos sean minimizados. No es una tarea fácil, las diferencias pueden no ser específicas del uso de distintos métodos y sí por las diferencias entre las poblaciones que contestan el censo usando cada método; por ejemplo, las personas con mayor nivel de educación que responden por internet.

3. Operacionalización

Este procedimiento presenta el mayor desafío de un modelo mixto. Como se ha dicho anteriormente, cada método tiene que trabajarse en su contexto debido a sus particularidades específicas, por cuanto hay que tratarlos separadamente, como si fuesen censos distintos. De todos modos, a la larga todos tendrán que estar totalmente integrados.

El control cuantitativo de la recolección tiene que asegurarse cuidadosamente en cada método, a fin de producir un resultado que refleje realmente lo que está ocurriendo con cada método y también con el censo. Los sistemas de control de cada método tienen que monitorearse regularmente generando reportes oportunos en los diferentes niveles geográficos.

En este caso, la administración de un Sistema Integral de Gestión Operativa (SIGO) debe considerar, al menos, algunas funciones que permitan tener una visión clara de los grandes grupos de tareas a desarrollar: la identificación unívoca de cada cuestionario, el seguimiento de los cuestionarios

iniciados con un método de captura y finalizados con otro, el monitoreo de los rezagados, las no-respuestas, el envío de recordatorios, las acciones de seguimiento y la administración del personal de campo, según los avances y la integración de los datos capturados por diferentes métodos.

4. Costo

Existen algunas controversias sobre el tema de los costos. Por un lado, se argumenta que el esfuerzo del operativo disminuye con el uso de la internet para autoenumerarse; pero, por otro lado, hay que calcular los gastos adicionales en el desarrollo de los sistemas computacionales, involucrando pruebas, posibles contrataciones externas, implementación del sitio institucional y toda la complejidad adicional asociada al aumento de la capacidad del servidor para soportar los accesos simultáneos, seguridad, respaldos del mismo, paralelismo (o replicación) para garantizar el acceso.

El párrafo D.114 del documento “*Guidelines*” trae una lista con ejemplos de algunos países fuera de la región que implementaron el censo con un modelo mixto secuencial (*multi-phase approach*), entre ellos, presenta una descripción más detallada sobre los censos de Canadá.

En nuestra región, los ejemplos más recientes son Colombia 2018 y México 2020, mientras que actualmente se está aplicando el Autoempadronamiento Censal —eCenso— como primera fase en la República Bolivariana de Venezuela 2021.

5. Integración de datos

La integración de datos es el proceso en el que se combinan diferentes fuentes, en este caso, distintos archivos digitales provenientes del operativo censal. Generalmente, suele identificarse como una fase del procesamiento de datos; es decir, posterior a la captura de datos. Sin embargo, cuando se usa una combinación de métodos de captura durante el operativo (DMC o eCenso), pasa a formar parte de la recolección misma, dado que también alimenta los sistemas de gestión y aseguramiento de la calidad.

En este sentido, los países de la región han innovado en:

- Generación de identificadores únicos para unidades de empadronamiento, permitiendo identificar duplicados en las bases de datos, según el método de captura utilizado.
- Utilización de códigos QR, metodología implementada por México (2020) y planificada para la República Bolivariana de Venezuela (2021).
- Metodología de sincronización de información con priorización de información de un método sobre otro.

F. Métodos complementarios de recolección

Con la embestida de la pandemia por COVID-19, el contacto con la población vía entrevistas cara a cara tuvo que repensarse, evitarse o minimizarse. El autoempadronamiento por internet (eCenso) se ha valorado, pero aún representa un gran desafío para los países de la región. Una alternativa poco considerada antes de la pandemia, que se había utilizado básicamente como mecanismo de recuperación en los censos de población y vivienda es la *computer assisted telephone interview* (CATI), esto es, entrevistas usando el teléfono como medio de comunicación para la recolección de datos. Sin embargo, en el contexto del distanciamiento social este método ha ganado visibilidad, sobre todo para las encuestas de hogares.

En este método, los entrevistadores trabajan conectados a un “centro de llamadas” (call center) e ingresan las respuestas con la ayuda de un computador programado de una manera similar a la digitación, DMC o aplicativo Web. Se trata de una transferencia de información usando computadores con las mismas ventajas y desventajas asociadas al uso de un programa de captura (saltos programados,

validaciones). En términos de “costos” cambia, ya que no se va a terreno, pero se invierte en la infraestructura del centro de llamadas y del personal asociado. En este método alternativo, no es necesario que se usen computadores interconectados, es posible trabajar con papel o DMC (celulares o tabletas), sacrificando el control y administración del proceso y la evaluación de los entrevistadores.

Entre las desventajas inherentes a este método complementario está el hecho que la realización de una entrevista por teléfono también impone más restricciones a la longitud del cuestionario y la complejidad de las preguntas, en comparación con otros métodos. Algunos efectos de respuesta, como el orden de respuesta, es más probable que ocurran en las encuestas telefónicas que en las autoempadronadas por internet. Por ejemplo, es más probable que los entrevistados seleccionen la última opción de respuesta (efecto reciente) en una encuesta telefónica.

Igualmente, hay que estudiar las dificultades para obtener los contactos telefónicos de los posibles entrevistados y, además, asociar esos teléfonos a las direcciones físicas de las viviendas y la identificación geográfica completa de la base cartográfica, sumado a la problemática de diferenciar entre hogares y viviendas. Como el hogar puede tener más de una persona, cada una de ellas con su teléfono, hay que evitar la duplicidad de conteo. Otro problema para considerar es cómo definir por teléfono quién será la persona responsable por el hogar, de modo que esta persona sea la interlocutora de la entrevista.

IV. Matriz de decisión–criterios de evaluación

Con el fin de aportar a la toma de decisiones sobre el método de captura de datos más apropiado para un INE, se presenta a continuación un modelo basado en matrices cualitativas y cuantitativas de decisión, las que utilizadas en su conjunto permitirán una evaluación más completa, según las condiciones específicas de cada país. Aunque se considere el uso del modelo mixto de captura, siempre hay que elegir el método principal; por tanto, las matrices están pensadas para evaluar cada método independiente de la combinación que se decida usar como complemento al método principal.

La idea base de una matriz de decisión es enlistar todos los criterios que se deben evaluar, los cuales estarían representados en las filas, mientras los métodos de captura estarían en las columnas. Cada celda de la matriz lleva la información correspondiente a la aplicación del criterio (fila) y al método de captura (columna). Según el tipo de matriz en la que se esté trabajando, la “información” en cada celda puede ser:

- cualitativa (o subjetiva), por ejemplo, la ventaja, desventaja o un comentario de ese método en relación con los otros métodos; o,
- cuantitativa, estableciendo una jerarquía entre los métodos, para cada criterio específico.

Es relevante recordar que las comparaciones entre métodos se dan por criterio (fila) y en separado; es decir, los métodos son evaluados y comparados entre sí en relación con cada criterio.

Como puede intuirse el quid del método está en la selección de los criterios a evaluar, una selección equivocada (o condescendiente) de criterios, obviamente puede conducir a una decisión defectuosa, por lo que es necesario prestar especial atención a la inclusión de criterios que sean realmente relevantes y diferenciadores para la toma de decisión en cada matriz.

La diferencia entre ambas señala que en la matriz cualitativa se escriben comentarios o descripciones en las celdas, mientras que en la cuantitativa se dan puntajes a cada celda de acuerdo con la jerarquía de los métodos para cada criterio. Ambas matrices son necesarias, dado que para decidir el método a utilizar los factores que se deben observar o cautelar son diversos y no todos son

cuantificables. De todos modos, en la decisión final se deberán ponderar factores institucionales, tecnológicos, operativos, culturales y económicos, los cuales pueden agregarse a través de las matrices.

El uso de información cualitativa en el proceso de decisión no debe entenderse como información menos robusta, debido a que la base para un juicio sólido sobre la capacidad y riesgos de este tipo de información se debe construir con énfasis en información cuantitativa y cualitativa documentada.

En el recuadro siguiente, se puede leer una breve síntesis de los factores que se evaluaron en el proceso decisional, seguido por el Instituto Nacional de Estadística y Censos, INEC, de Costa Rica al momento de planificar el censo de la década de 2020²⁰.

Recuadro 4

Síntesis del proceso de decisión en la elección de uso de DMC como método para la recolección censal para el censo 2021, Costa Rica.

Costa Rica no ha realizado aún ningún censo nacional utilizando DMC. Sin embargo, el INEC se prepara para implementar el Censo 2021 con DMC.

En el proceso de decisión se analizaron los siguientes factores:

- La experiencia del INEC desde 2010 con el uso de DMC en otras operaciones estadísticas como encuestas y registros.
- La inversión de la compra de DMC y desarrollo tecnológico frente a la impresión de papel y contratación de lectores ópticos.
- La automatización de los flujos de las entrevistas y los controles que se realizan durante la captura de datos para mejorar la calidad de los datos.
- La oportunidad de mejorar la cobertura censal.
- La oportunidad de validar parte de la información en tiempo real y disminuir los tiempos de procesamiento de datos.
- La posibilidad de brindar los resultados censales de una forma más rápida y oportuna.
- La accesibilidad y usabilidad del sistema de captura de datos, en los diferentes flujos posibles para disminuir los posibles errores de aplicación del cuestionario.
- Las condiciones del país respecto a la cobertura de internet e infraestructura tecnológica.
- Las habilidades tecnológicas, en cuanto al uso de dispositivos móviles por parte de la población como posibles censistas.

A partir de esas valoraciones, se determinó que, el Censo 2021 de Costa Rica se lleve a cabo utilizando DMC.

Fuente: Texto editado a partir la Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe, del GTC CEA-CEPAL, véase Anexo 2.5 de este documento.

A. Criterios de evaluación

Al realizar una evaluación que permita decidir sobre el tipo de método de captura que se utilizará en un censo, la lista de criterios a considerar puede ser tan sencilla como la que se propone en el siguiente cuadro, el cual contiene cinco criterios globales o más minuciosa, según la profundidad con que se quiera evaluar cada factor.

²⁰ Originalmente este censo estaba planificado para realizarse en 2021, pero debió ser pospuesto por los efectos de la pandemia por COVID-19, por eso el texto hace referencia al Censo 2021.

Cuadro 10
Criterios de evaluación de los métodos de captura de datos

Criterio	Descripción
Precisión	Medida en que los datos capturados reflejan con la mayor (o menor) exactitud la entrevista o la "fotografía" del país en la fecha del censo; es decir, la distancia entre el valor estimado y el valor verdadero (desconocido).
Oportunidad	Rapidez con la cual los datos del censo se ponen a disposición de la sociedad.
Control de Cobertura	Capacidad para apoyar el control de la cobertura de población y viviendas efectivamente censadas.
Costo monetario	Costo total que hay que pagar por el uso de la estrategia, incluyendo todos los recursos necesarios: tecnológicos, logísticos, financieros y humanos, independiente de si son inversión o gasto.
Riesgo	Probabilidad de que existan problemas que dificulten, en menor o mayor grado, la llegada a un buen término del proyecto.

Fuente: Elaboración propia a partir de CEPAL (2011a).

Al analizar la información en el cuadro de arriba, se entiende que una buena evaluación usando la matriz radica en la definición y pertinencia de los criterios. Aunque exista una definición inicial de cada criterio en la tabla, al ser una lista de criterios genéricos, se hace difícil aplicarla sin que se levante una serie de dificultades, excepciones, casos particulares y especificidades nacionales, entre otras, como por ejemplo en el criterio de "Costo monetario". Por otro lado, las matrices más detalladas facilitan la definición y entendimiento de los criterios. Adicionalmente, se determinarán los criterios más pertinentes para utilizarse en una matriz cualitativa, cuantitativa o mixta.

Es posible elaborar listas con criterios generales o específicos del método de captura. El cuestionario o boleta censal es uno de los múltiples formularios o documentos utilizados en un censo que se requieren en formato digital, ya sea como instrumentos para el control operativo, en el apoyo a los procesos siguientes y/o como auditoría o gestión del censo. En algunos países de la región se entregan resultados preliminares en base a los formularios de control; por tanto, sería relevante evaluar también el uso de tecnología para su captura asegurando su calidad y oportunidad. Así, se podría establecer una lista de criterios que considere la captura no solo del cuestionario censal, sino de todos los formularios, dado que los países pueden establecer diferentes combinaciones: digitación o escaneo de la boleta censal y uso de DMC para los formularios de control operativo y de logística.

De este modo, las listas de criterios para las matrices cualitativas o cuantitativas deben entenderse como características a evaluar que permitan elegir un método de captura sobre otro. Complementariamente, se puede establecer otra lista de criterios que contemple aspectos como los señalados en el Cuadro 11.

Para la evaluación de los costos de cada método se debiera hacer un ejercicio similar de cálculo de "costo total de propiedad" (*Total cost of ownership* o TCO, por su sigla en inglés), que se refiere a la estimación de los costos directos e indirectos de desarrollar e implementar un sistema en particular. Hay tres componentes claves para calcular el TCO: costos de software/hardware, costos operativos y costos de personal. A continuación, se muestra el ejemplo de una plantilla para calcular el costo total de propiedad de posibles métodos alternativos para la recopilación de datos:²¹.

²¹ Ver párrafos B.50-52 de United Nations (2019).

Cuadro 11
Ejemplo de criterios a considerar

Criterios a considerar			
1.	Impresión de cuestionarios	9.	Formularios de control de cobertura
2.	Nivel educativo de los censistas	10.	Controles de consistencia
3.	Esfuerzo en la captura de datos	11.	Experiencias regionales y mundiales
4.	Esfuerzo de codificación de textos (preguntas abiertas)	12.	Casos óptimos de uso
5.	Flexibilidad en la entrevista	13.	Costos
6.	Dificultades logísticas en terreno	14.	Plazos de adquisición
7.	Dificultades logísticas en bodega	15.	Niveles de precisión
8.	Proporción de la población que tiene conocimientos informáticos básicos	16.	Cobertura de telefonía celular / móvil

Fuente: Elaboración propia a partir de la información proporcionada por CELADE-CEPAL (2011) y United Nations (2020 y 2019).

Cuadro 12
Ejemplo de planilla para cálculo de costos fijos y variable según método de captura

	Método de captura		
	Escáner	DMC	eCenso
Costos fijos			
Hardware			
Software			
Servicios de soporte técnico			
Construcción de infraestructura tecnológica			
Otros equipamientos electrónicos			
Mantenimiento			
Costos variables			
Personal temporal (reclutamiento y salarios)			
Capacitación y formación del personal			
Impresión de cuestionarios y formularios de control			
Bodegaje y traslado de materiales			
Digitación			
Transferencia de datos			

Fuente: elaboración propia sobre la base de la información proporcionada por United Nations (2019).

B. Evaluación cualitativa

Con el propósito de ejemplificar el diseño y uso de la matriz, en este apartado se mencionan algunas características, ventajas y desventajas de cada método para construir un juicio de valor que ayude en la decisión de cuál método usar para la captura de los datos del censo, teniendo en cuenta que las características de cada método son presentadas en el capítulo III. Métodos para la captura de datos censales.

La evaluación cualitativa se inicia con la elección de la lista de criterios (filas) y se arma la matriz con los métodos de captura (columnas) y en cada cuadrante se agregan los detalles característicos de

cada combinación criterio/método, indicando si el método cumple de mejor o peor forma en comparación con los otros métodos en estudio.

Aunque la evaluación cualitativa no lleve directamente a la elección del método de captura, es un insumo valioso para la construcción de la matriz en la evaluación cuantitativa, incluso para comprender mejor cada uno de los criterios y atribuir los puntajes a cada celda. Por lo que, se recomienda realizar primero la matriz cualitativa y después la cuantitativa, aunque probablemente se hagan ajustes en ambas durante el proceso de análisis y las implicancias de cada cuadrante.

Respecto de las consideraciones cualitativas que se deben tomar en cuenta para la toma de decisiones sobre el método de captura, se encuentran los antecedentes sobre si la legislación actual permite el cambio de método o incluso si es factible cambiar la ley. Igualmente, se debe valorar la capacidad institucional para adoptar una nueva tecnología, respecto de la capacidad organizacional, recursos humanos, técnicos y financieros, incluyendo la percepción de la población sobre el uso de esta tecnología, entre otras consideraciones.

En el ejemplo siguiente, se ilustra la construcción de una matriz cualitativa. La lista de criterios se basa en algunas fases o subfases del Modelo Genérico del Proceso Estadístico o Generic Statistical Business Process Model (GSBPM)²² de la Comisión Económica de las Naciones Unidas para Europa (UNECE) la cual, a criterio de los autores, es relevante al momento de pensar exclusivamente en el método de captura, pero que en ningún caso se considera como única ni excluyente de otros criterios inspirados en el GSBPM o en sus adaptaciones nacionales a los censos de población y vivienda. Luego, desde el ámbito vertical, se identificaron los principales métodos de captura indicados como posibles de utilizar por los países del grupo de trabajo CEA-CEPAL para la ronda 2020.

Tal como se ilustra en el siguiente cuadro, se recomienda realizar el llenado por cuadrícula respecto a la fase o subfase y hacer el análisis de cada una agregando las descripciones o afirmaciones cualitativas, tomando en cuenta las necesidades, posibilidades y realidades de cada país.

En el Anexo 1 se presenta otro ejemplo de una matriz cualitativa que permite ilustrar la utilidad de la herramienta para la toma de decisiones. En este caso, no es una matriz teórica, sino un ejercicio hecho en 2014, por el Instituto Nacional de Estadística e Informática, INEI, del Perú, en el cual se exponen criterios cualitativos y sus respectivas valoraciones para evaluar la utilización de digitación, escaneo o DMC. En ese entonces, la alternativa vía Internet (eCenso) no se consideró como un posible método principal, por lo que no se incluyó en la evaluación.

Cuadro 13
Ejemplo de matriz cualitativa para elegir métodos de captura de datos censales

	Método de captura			
	Digitación	Escaneo	DMC	eCenso
Fase de diseño del proceso de recolección				
Diseño de los sistemas de producción y de los flujos de trabajo	Más sencillos que con DMC o eCenso y con menores posibilidades de monitoreo, control de cobertura y alertas tempranas.	Más sencillos que con DMC o eCenso y con menores posibilidades de monitoreo, control de cobertura y alertas tempranas.	Más complejos que con papel y eCenso, con mayores posibilidades de monitoreo, control de cobertura y alertas oportunas.	Más sencillos que con DMC, pero con mayor complejidad en los sistemas de seguimiento de cobertura y alertas.

²² Disponible en https://ec.europa.eu/eurostat/cros/content/gsbpm-generic-statistical-business-process-model-theme_en

	Método de captura			
	Digitación	Escaneo	DMC	eCenso
Fase de construcción del instrumento de captura				
Impresión	Sí	Sí, pero de mayor complejidad.	No	No
Programación del software de captura	Sí, sistema de digitación y pruebas simples.	Sí, sistema de identificación de marcas y caracteres, pruebas simples.	Sí, sistema con mayor precisión en los validadores dada la oportunidad y consistencia, mejora la calidad de los datos. Pruebas complejas.	Sí, sistema con mayor precisión en los validadores dada la oportunidad y consistencia, mejora la calidad de los datos. Pruebas complejas.
Logística	Mayores esfuerzos en el diseño de la logística, debido a las dificultades de traslado, acopio y manipulación de grandes volúmenes de papel, que, además, implican cantidades importantes de personal solo para disponer los cuestionarios para la digitación diariamente.	Los mismos esfuerzos en el diseño de la logística, debido a las dificultades de traslado, acopio y manipulación de grandes volúmenes de papel que en la digitación y bastante menos cantidad de personal para disponer los cuestionarios para la línea de guillotinado y escáner.	Logística más sencilla que en el caso de papel, debido a que los volúmenes son mucho menores y no hay tanta necesidad de acopio; sin embargo, es necesario tener más consideraciones por seguridad ante actos humanos y fragilidad a la temperatura en algunas zonas climáticas.	Menores que en los otros métodos, pero se requiere de soluciones para entregar la clave de acceso que requerirá de logística.
Construcción o mejora de componentes del sistema de captura	No	Sí, cuando se usan paquetes cerrados.	Sí, cuando se usan paquetes cerrados.	Sí, cuando se usan paquetes cerrados.
Pruebas al sistema de georreferenciación de unidades	No permite georreferenciar	No permite georreferenciar	Sí	Sí
Prueba piloto del proceso estadístico	Sí, son pruebas más simples, pero se deben realizar con mayor anticipación a la fecha del censo.	Sí, similares al caso de papel para digitación.	Sí, se requiere realizar más pruebas y más complejas, pero permiten mayores cambios y hasta muy cerca de la fecha del censo.	Sí, se requiere realizar pruebas más complejas ante la ausencia de censistas, pero permiten mayores cambios y hasta muy cerca de la fecha del censo.
Fase de recolección				
Preparación de la recolección: Reclutamiento del personal que captura la información	Contratación masiva de censistas, editores y digitadores.	Contratación masiva de censistas, editores y manipuladores de la línea de producción del escáner.	Solo censistas.	No
Preparación de la recolección: Capacitación de censistas y supervisores	Sí, sobre forma de contestar (marcas, y tipografía), conceptos y secuencia del cuestionario.	Sí, sobre forma de contestar (marcas, y tipografía), conceptos y secuencia del cuestionario.	Sí, sobre el uso del DMC y conceptos.	No se capacita, pero se debe generar un sistema amigable y sencillo para habilitar a las personas para contestar, entregando información sobre definiciones y aspectos conceptuales para contestar adecuadamente el censo.
Ejecución de la recolección: Funcionalidad en la captura	Puede volver fácilmente a las preguntas en el orden que desee.	Puede volver fácilmente a las preguntas en el orden que desee.	Tiene mayor dificultad en devolverse y generar cambios durante el empadronamiento.	Tiene mayor dificultad en devolverse y generar cambios durante el empadronamiento. Pero, permite la autoadministración del o de los tiempos para responder.

	Método de captura			
	Digitación	Escaneo	DMC	eCenso
Fase de recolección (continuación)				
Ejecución de la recolección: Sistema de control y gestión de la recolección	Sistema lento, de poca oportunidad para controlar el avance de terreno y no favorece un oportuno control de calidad masivo.	Sistema lento, de poca oportunidad para controlar el avance de terreno y no favorece un oportuno control de calidad masivo.	Sistema en línea, oportuno para el control, de gestión y avance en terreno, así como para advertir oportunamente problemas de calidad o sesgos.	Sistema en línea complejo para conocer el avance en cobertura de viviendas y personas.
Ejecución de la recolección: Cierre de la recolección	Guardado en papel.	Guardado de imágenes.	Guardada imagen del dispositivo.	Guardado en un registro digital.

Fuente: Elaboración propia a partir de los intercambios de experiencias del GTC CEA-CEPAL realizados en 2020.

C. Evaluación cuantitativa

En base a los cinco criterios generales descritos en el Cuadro 10: Criterios de evaluación de los métodos de captura de datos, se presenta un ejemplo sencillo de matriz que ejemplifica dichos "conceptos" en una escala de valoración de 1 a 4, en donde 1 es el método que menos contribuye a potenciar cada criterio y 4 sería el que más aporta. Así, se puede armar una aproximación para el diseño de una matriz de decisión, recordando que los dos últimos criterios, costo monetario y riesgo, se ordenan de manera inversa. Por lo que, la alternativa 1 representa el de mayor costo y la 4 la de menor costo. Al igual que con el riesgo, a mayor probabilidad de ocurrencia e impacto del riesgo, la clasificación es 1 y a menor riesgo de ocurrencia es 4.

Para ello, se puede adoptar cualquier sistema de puntaje, aunque en el documento se elige un puntaje de 1 a n, en donde n es el número de métodos (en nuestro caso, 4). El número 1 significa el menor puntaje y el 4 el mayor. Al final, se suman los puntajes de cada método donde el de mayor puntaje debería ser el método elegido.

A continuación, se presenta un ejemplo de la matriz descrita; sin embargo, es necesario decir que los resultados expresados en esta tabla para el criterio "costo monetario", en un ejercicio real debiera atender al cálculo de "costo total de propiedad" de cada alternativa, ya que estará supeditado a los precios de los mercados nacionales, referentes a tecnología, recursos humanos y capacidad instalada de cada INE.

Cuadro 14
Ejemplo de matriz de decisión basada en 5 criterios globales

Criterios	Métodos			
	Digitación	Escaneo	DMC	eCenso
Precisión	1	2	4	3
Oportunidad	1	2	4	4
Control de Cobertura	1	2	3	1
Costo monetario	4	3	1	3
Riesgo	1	2	3	1
Total	8	11	15	12

Fuente: elaboración propia, sobre la base de la experiencia de los autores.

Esta matriz sirve única y exclusivamente para ejemplificar de una manera sencilla la adopción de puntajes para valorar los criterios en cada método. Por ejemplo, siguiendo la matriz:

- En el criterio precisión, la digitación presenta menor valoración (1) en relación con el método DMC (4); es decir, obtiene una menor valoración en cuanto a la medida en que los datos capturados reflejan la distancia entre el valor estimado y el valor verdadero (desconocido).
- En el criterio oportunidad, la digitación presenta menor valoración (1) en relación con los métodos eCenso y DMC (4); es decir, obtiene una menor valoración en cuanto a la rapidez con la cual se ponen a disposición los datos del censo a la sociedad.
- En el criterio cobertura, los métodos eCenso y digitación presentan menor valoración (1) en relación con el método DMC (3); es decir, obtiene menor valoración como una herramienta que ayuda a mejorar la cobertura de viviendas y personas.
- En el criterio costo monetario, la digitación presenta una mayor valoración (4), es menos costoso en comparación con el método DMC (1); es decir, obtiene una menor valoración en cuanto al costo total que hay que pagar por el uso de la estrategia, incluyendo todos los recursos necesarios, tecnológicos, logísticos, financieros y humanos, independiente de si son inversión o gasto.
- Es necesario insistir en que este es un ejercicio hipotético y dependerá de los precios locales e internacionales que enfrenta cada país la determinación de si un servicio, intensivo en mano de obra como lo es la digitación, es más o menos costoso que uno intensivo en tecnología (DMC).
- En el criterio riesgo, la digitación presenta menor valoración (1), es más riesgoso en relación con el método DMC (3); es decir, se considera que existe mayor probabilidad de problemas que dificulten la llegada a un buen término del proyecto.

En resumen, y en base a las experiencias de censos pasados, se puede concluir que la digitación es el método de captura menos preciso, menos oportuno, de costo relativamente bajo y de mayor riesgo. En el otro extremo, la captura por eCenso presenta una precisión alta, mayor oportunidad, costo no tan alto y más riesgoso.

La fila del Total en el Cuadro 14 representa la evaluación general de cada método, según el cual eberíamos usar el DMC que alcanzó el mayor valor al sumar los puntajes de cada criterio. Resultando que la captura por DMC es la de mayor precisión, oportunidad, cobertura y costo, pero también de moderado riesgo.

Adicionalmente, se pueden asociar ponderaciones por relevancia, importancia o prioridad de los criterios. Por ejemplo, criterios de menor importancia llevarían una ponderación de 0 o un valor de ponderación muy bajo; los criterios de importancia media llevarían peso de 1; mientras los criterios más relevantes podrían llevar pesos de 1,2; 1,5; 2 o cualquier sistema de ponderación que permita asignar una jerarquía a los criterios para discriminar entre un método y otro.

Considerando la misma matriz anterior, se pueden asociar pesos a cada uno de los cinco criterios. Por ejemplo, se definen ponderaciones considerando que el criterio de Riesgo es el más importante (peso tres), luego la Precisión (peso 1,8), el Control de la Cobertura (peso 1,5), la Oportunidad (peso 1,2) y el Costo (peso 1). Con ese conjunto de ponderadores, el método a seleccionar es el DMC (como se puede observar en el ejemplo del Cuadro 15: Matriz de decisión con ponderadores, dado que arroja una mayor valoración final al ponderar, según las preferencias específicas de cada país.

Cuadro 15
Matriz de decisión con ponderadores

Criterios	Métodos				
	Ponderación	Digitación	Escaneo	DMC	eCenso
Precisión	1,8	1,8	3,6	7,2	5,4
Oportunidad	1,2	1,2	2,4	4,8	4,8
Control de Cobertura	1,5	1,5	3	4,5	1,5
Costo monetario	1	4	3	1	3
Riesgo	3	3	6	9	3
Total		11,5	18,0	26,5	17,7

Fuente: Elaboración propia sobre la base de la experiencia de los autores.

El uso de ponderadores permite establecer la relevancia que cada país le atribuye a los criterios según su contexto y objetivos estratégicos. En ciertos casos, por ejemplo, será más relevante el costo; mientras que, en otros, el riesgo o la cobertura. Este proceso facilita, además, la diferenciación entre cada método y reduce la posibilidad de empates, aunque esto es difícil, pues la matriz que será utilizada por los países de la región, probablemente, tendrá más de cinco criterios.

V. Monitoreo del proceso de recolección censal²³

El presente capítulo ilustra los componentes que debieran tener los sistemas de control y seguimiento del operativo censal, con énfasis en el proceso de recolección de un censo de derecho, aunque pudiera aplicarse parcialmente en censos de hecho. Adicionalmente, busca relevar las potencialidades del uso de DMC para mejorar el control operativo y de cobertura del trabajo de terreno. Se suma a ello el aporte en el control y aseguramiento de la calidad de los datos, por medio de la transformación y estandarización de la supervisión del trabajo de los censistas, al poder observar indicadores “nuevos” que sería costoso hacerlo sin tecnología.

La coordinación y el seguimiento de la recolección censal presenta siempre un desafío. Esta fase ha evolucionado con las innovaciones tecnológicas, transformando los procesos manuales en sistemas automatizados, integrados y dinámicos. Como consecuencia, la supervisión del avance en terreno ha mejorado, también, la gestión del personal y el monitoreo de indicadores de desempeño y demográficos, permitiendo un mayor control de la calidad de los datos durante el operativo.

En este sentido, el monitoreo de la recolección censal se sustentará sobre un Sistema Integral de Gestión Operativa (SIGO), el cual será más o menos complejo, según el método de captura utilizado y los requerimientos establecidos por la “gerencia” del censo, entre otros aspectos. Encontrar un equilibrio entre el tipo, la cantidad, la periodicidad y la desagregación geográfica, de la información requerida para la gestión de la recolección no es una tarea sencilla, menos cuando la cruzamos, además, por la estructura organizacional, la que obviamente requiere diferentes tipos y formas de información. En este contexto, una de las recomendaciones más claras y presentes desde hace ya 20 años en el Manual de Censos (Naciones Unidas, 2001) es que se debe asegurar que la información que se va a recolectar sea necesaria, oportuna y que se utilice, constructivamente, en la gestión de la recolección censal.

²³ El presente capítulo se fundamenta en la información disponible en United Nations (2001) y (2019); en CELADE, CEPAL (2011); además de las consideraciones expuestas por Colombia, la República Bolivariana de Venezuela, el Brasil, México, el Ecuador, Costa Rica y Chile en el marco del GT Censos 20-21.

Los métodos de captura, como los DMC o el eCenso, mejoran la oportunidad de la información requerida para la gestión del operativo, también aumentan la cantidad y diversidad de información que se puede requerir y procesar para estos fines, a diferencia de los censos en papel. Cuando la recopilación censal se hace por medio de entrevistadores con DMC y en censos de derecho, relativamente extendidos en el tiempo (3 semanas o más, por ejemplo), las posibilidades de control del avance del operativo se ven, además, potenciadas por las oportunidades que ofrece para supervisar y asegurar la calidad del proceso.

Se deben diseñar diferentes módulos, pantallas y resúmenes de información para los diferentes niveles jerárquicos y decisionales, ya sean nacionales, regionales, provinciales, comunales y/o locales, atendiendo a las responsabilidades de cada puesto, así como también información específica para los supervisores y los propios censistas. La consistencia y jerarquización de las responsabilidades y ámbitos decisionales de cada puesto son vitales para el correcto diseño del sistema y el aseguramiento de la calidad y la cobertura del censo.

Por su parte, la naturaleza de la recolección censal implica que la información de gestión puede tardar algún tiempo en ser adquirida y recibida por los distintos usuarios del sistema, ya sea porque los DMC no tienen transmisión en línea y se debe esperar uno o dos días para que transmitan y, luego, para que esta información se sincronice, integre y procese. Por lo tanto, si se considera que la información no puede recolectarse en un plazo razonable, para que pueda utilizarse eficazmente, entonces se recomienda que no sea recopilada. Aunque en algunos casos, el tipo de información o el nivel de detalle puede reconsiderarse y reorganizarse.

En términos generales, la experiencia reciente en la región muestra que los SIGO centralizan información asociada a los procesos de preparación, ejecución y calidad de la recolección censal para los distintos perfiles gerenciales y operativos del censo, incluyendo:

- *Gestión de personas*, en sus aspectos de reclutamiento, selección, contratación, rotación y pago.
- Capacitación, por ejemplo, cantidad de personal capacitado y fechas de capacitación.
- Logística, incluyendo sistema de tracking para distribución y mantención de materiales y equipos en los locales censales.
- Monitoreo y control de la transferencia de datos del operativo: que se transfieran todos los datos, que estos sean recibidos y leídos correctamente y que se integren sin omisiones ni duplicaciones (en censos con DMC).
- Indicadores del proceso de recolección (cobertura poblacional y de viviendas), considerando asignación de cargas de trabajo, estado de la unidad de empadronamiento, avance de la recolección respecto de lo planificado, entre otros indicadores.
- Indicadores de calidad, que implican indicadores demográficos para evaluación del proceso durante la recolección, como estructura etaria de la población, relación de dependencia e índice de masculinidad, entre otros (en censos con DMC).
- Indicadores de consistencia de la información del cuestionario censal, que implica indicadores de no-respuesta, consistencia entre preguntas y completitud de la información.

Es necesario precisar que, la experiencia en la región —en el uso de DMC, como método de captura principal—, es bastante reciente en términos de calendarios censales, a saber: Brasil 2010, Uruguay y Venezuela (República Bolivariana de) en 2011, Paraguay en 2012, Colombia en 2018 y México en 2020; es decir, los países recién acumulan un único evento censal con uso de DMC, por lo que el mayor desarrollo de indicadores de calidad —asociados al desempeño y supervisión de los censistas—, es un desafío para la región, aunque el Brasil ya ha mostrado un avance en este sentido en los preparativos

para el censo que estaba pronto a realizar y que debió ser suspendido por causa de la pandemia por COVID-19. Estos indicadores debieran contener información como la siguiente: tiempo de entrevista, promedio de personas por hogar, selección de la ruta más corta de la entrevista²⁴, entre otros, los cuales se deberían integrar con los indicadores sociodemográficos y mejorar sustancialmente la capacidad de supervisión y, por tanto, la calidad de los datos, mediante alertas tempranas que permitan rectificar errores de la capacitación o malas prácticas del personal de campo, reduciendo los costos y acciones de supervisión en terreno.

Recuadro 5

Forma de monitoreo de la recolección con dispositivos móviles de captura, el Brasil, Censo de 2010

Los sistemas de control eran descentralizados. Toda la vinculación del equipo de campo con los DMC y la recepción de los datos ocurrían en los "Puestos de Recolección", con uso de Notebooks que contenían un sistema llamado Sistema de Gerenciamiento del Puesto de Recolección (SIGPC). Los DMC no se comunicaban con el ambiente central, solamente lo hacía el SIGPC. Cuando se hacía la comunicación del SIGPC con el ambiente central, toda la estructura del Instituto Brasileño de Geografía y Estadística (IBGE) tenía acceso a los datos para hacer el acompañamiento (generalmente, se hacía esa comunicación una vez al día). Todo el control del trabajo dependía de la conexión de los DMC con el SIGPC. Los DMC contaban también con una tarjeta de memoria para respaldo, que podía descargarse directamente en el SIGPC en caso de problemas con el equipo. Toda la validación de los datos se realizaba en el ambiente central, donde había criterios preestablecidos para cada caso. Las pérdidas de datos o datos erróneos fueron muy insignificantes.

Fuente: Texto editado a partir la Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe, del GTC CEA-CEPAL, véase Anexo 2.1 de este documento.

En el caso que un INE decida hacer un censo con más de un método de captura (método mixto), el diseño y administración de un SIGO debe considerar las siguientes funciones que permitan tener una visión clara de los grandes grupos de tareas a desarrollar:

Cuadro 16

Funciones mínimas que debiera contener un SIGO para hacer un censo con método de captura mixto

Función	Descripción
Gerenciamiento de cuestionarios	Identificar unívocamente cada cuestionario, asignando un número único para cada vivienda. Se debe seguir el estado de cada cuestionario durante todo el período de enumeración. En un censo tradicional, los informes se refieren a sectores censales como el nivel de desagregación más bajo, pero en un modelo mixto, dado que los cuestionarios podrían cambiar de método de captura durante el proceso, ese <i>tracking</i> tiene que ser ejecutado a nivel de cuestionario. Ese número único de cuestionario también debe usarse para evitar duplicación de respuestas.
Administrar la transferencia de cuestionarios	Entre los métodos de captura, los cuestionarios iniciados con un método pueden terminarse usando otro método, particularmente cuando la vivienda tiene varios miembros o no-respuesta en las primeras visitas.
Administración de no-respuestas	Tener un sistema de monitoreo para enviar recordatorios a los "rezagados" y determinar acciones de seguimiento.
Administración de personal de campo	Asignar o reasignar entrevistadores durante el operativo, según el aumento o disminución de la carga de trabajo (en algunas áreas la tasa de no-respuesta puede ser alta, exigiendo acciones de seguimiento, en otras puede ser más baja).
Integración de los datos	Definición de sistemas para combinar los varios formatos de datos provenientes de los métodos de captura.

Fuente: Elaboración propia a partir de las Fichas de Sistematización de Experiencias para Países de América Latina y el Caribe, del GTC CEA-CEPAL durante el año 2020.

²⁴ Se refiere, por ejemplo, a cuando los censistas omiten la población en edad de trabajar o a las mujeres en edad fértil para concluir la entrevista más rápido.

A. Monitoreo de la transmisión e integración de los datos

Para el adecuado monitoreo del proceso de recolección de datos censales capturados por DMC se requiere que la transmisión y sincronización de los datos, desde el terreno hacia la oficina central del censo, se realice de forma correcta, completa, expedita y segura, dado que esto marcará, en parte, la oportunidad con la que se pueda analizar y gestionar la información sobre la evolución del trabajo en terreno y favorecer la oportuna toma de decisiones respecto de la cobertura y calidad de la fase de recolección.

El contexto nacional será un determinante importante a la hora de establecer el sistema de transferencia de datos, dependiendo de las condiciones y medios disponibles de transmisión. En aquellos países que posean una infraestructura de comunicaciones robusta y de amplia cobertura, de conectividad de datos celulares, deberían preferir esta alternativa, dado que proporciona la mejor y más poderosa herramienta para transferir datos directamente desde el DMC a los servidores centrales, por lo que aumenta la eficiencia y confiabilidad, e incluso podría ahorrar en costos al reducir la necesidad de volver a los locales censales o puntos de carga por parte de los censistas. El impacto en el costo y el rendimiento del uso de redes celulares dependerá de la frecuencia y la cantidad de datos transmitidos.

Por el contrario, cuando la infraestructura de redes para celulares es deficiente, será necesario desarrollar mecanismos alternativos para la transferencia de datos desde el campo, como estaciones de recolección-recepción de datos y/o suscribir acuerdos asociativos con entidades públicas regionales o locales, el ejército o el servicio de telecomunicaciones para acceder a sus líneas telefónicas, computadoras y conexiones de Intranet o instalar antenas de radio para la conexión a Internet para proporcionar servicio de banda ancha en áreas remotas. De esta forma, los datos pueden transmitirse, periódicamente, sin la necesidad de esperar hasta que se complete un área de enumeración, algo que permite el monitoreo centralizado de la recopilación de datos y hace que la revisión y el análisis comiencen antes. Adicionalmente, el acto de guardar datos en una tarjeta de memoria SD disponible en el DMC puede ayudar a evitar la pérdida de datos en caso de que el dispositivo se bloquee, se congele o hasta que los datos recopilados se descarguen en las estaciones de recolección de datos.

La sincronización es la transferencia de datos entre los DMC y los servidores centrales o la nube. En general, envía automáticamente cualquier tarea completada al supervisor para su revisión. En segundo lugar, introduce automáticamente todas las nuevas asignaciones e instrumentos de recopilación de datos en el dispositivo del enumerador (censista). Una vez completada la sincronización, se puede mostrar el número de nuevas asignaciones recibidas y el número de entrevistas completadas que han sido enviadas. Es común utilizar la sincronización para actualizar las aplicaciones de entrada de datos en terreno, descargando los últimos datos del servidor. De esta manera, las modificaciones de la aplicación en la oficina central se pueden distribuir fácilmente a los entrevistadores en el campo.

Para asegurar una correcta sincronización de datos, los países de la región han implementado los siguientes procesos de control:

- Control de Integridad de la Información, lo que implica asegurar por una parte la integridad física de los datos, asociada al almacenamiento y extracción de información; como también a la integridad lógica de los datos, a partir de los requerimientos de las áreas del censo.
- Gestión de Respaldos de Información, en particular los países declaran la integración de entrevistas completas para análisis de datos demográficos en los sistemas y copias de respaldo, tanto a nivel local (tarjetas de memoria) como central (servidores espejo); además del respaldo de las entrevistas parciales, en caso de interrupción de servicios durante la recolección.

Recuadro 6

Control y monitoreo de transmisión de datos, México, censo 2020

Para la integración de la información a la base de datos central se implementaron dos mecanismos:

- i) Envío de entrevistas completas desde el DMC mediante el uso de datos móviles.
- ii) Envío de la información de todos los inmuebles por medio de memorias USB. El entrevistador le enviaba la información diariamente a su supervisor y éste, mediante el módulo de integración del OPERA, transfería la información vía Internet de todos sus entrevistadores al menos una vez al día.

Para el monitoreo y control de la transferencia de datos del operativo de campo, se integró un reporte de envíos de las figuras responsables de esta tarea, que informaba lo siguiente:

- Paquetes esperados a nivel nacional, estatal y por estructura operativa hasta supervisor de entrevistadores. Los paquetes esperados fueron determinados por el personal contratado y por una clave operativa única que se le asignó a cada entrevistador y supervisor de entrevistadores.
- Paquetes recibidos hasta el momento de la generación del reporte, porcentaje con respecto al número de paquetes esperados y paquetes faltantes.
- Estatus de los paquetes recibidos, que podían ser:
 - Paquete con integración correcta.
 - Paquete con integración parcial. Se revisaba manualmente para su integración total.
 - Paquete recibido con error. Era necesario volver a generar un nuevo envío.
- Nomenclatura errónea. Era necesario volver a generar un nuevo envío.
- Error en clave operativa. Era necesario volver a generar un nuevo envío desde el dispositivo de la figura operativa correspondiente.
- Fecha de integración fuera de periodo/Fecha del archivo anterior al último envío recibido. Era necesario volver a generar un nuevo envío, en algunos casos implicaba corregir la fecha del DMC.

Aunado a este reporte, diariamente se generaba una lista para cada supervisor o responsable de los envíos, en la que se indicaba, por censista, la última fecha en la que se había recibido información o si hasta el momento no se había recibido ninguna vivienda de su carga de trabajo. Este informe se enviaba a la estructura de coordinación y dirección del Censo, para que determinaran las causas y le dieran solución.

Para la eliminación de duplicados, se definió una estrategia de consolidación de la base de datos para tener datos homogéneos y únicos, respecto a la información censal que se captó en los diferentes operativos de campo del Censo 2020. Entre sus objetivos estuvieron asegurar la conciliación histórica de mnemónicos e implementar reglas de integridad de la información, así como definir los campos requeridos para cada instrumento y operativo, para que cada equipo de desarrollo se ajustara a las reglas definidas. En la definición de campos, se determinó generar un identificador único para cada unidad de observación, así se tiene uno para cada inmueble, vivienda particular habitada y vivienda colectiva, residente de la vivienda y personas migrantes, así como para cada localidad y manzana. Lo anterior permitió que en la etapa de integración de la base operativa de datos se pudieran identificar duplicados.

Durante la entrevista cara a cara, conforme se iban cubriendo secciones del cuestionario, los datos captados se iban guardando en una tabla temporal y, una vez concluida la entrevista, se copiaban a la base de datos productiva, evitando su pérdida.

Para los casos de robo o daño total del DMC se elaboró una herramienta de restauración de información del último respaldo o paquetes de información generada. De esta manera, la pérdida de información se redujo considerablemente, ya que se complementó con una serie de comunicados a la estructura operativa, solicitando el respaldo constante de información, creando así una cultura de protección de datos.

Fuente: Texto editado a partir la Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe, GTC CEA-CEPAL véase Anexo 2.7.

B. Monitoreo de indicadores del operativo

Durante la ejecución de la recolección es necesario incorporar acciones de control y alertas tempranas antes que la omisión, duplicación u otros problemas de calidad se reflejen al final del proceso. Es mejor evitar el error en el proceso de recolección que efectuar correcciones por métodos directos o indirectos en la siguiente fase de procesamiento. Las distintas fuentes de error que se desean controlar deben estar oportunamente identificadas y sistematizadas en un set de indicadores que permite accionar protocolos durante la labor de recolección (CELADE, 2011).

Algunos elementos destacados por la experiencia en América Latina y el Caribe se refieren a construcción de indicadores de gestión operativa, visualización geoespacial durante la recolección y el seguimiento de indicadores demográficos.

1. Indicadores de gestión operativa

Los indicadores de gestión operativa corresponden a una serie de indicadores de desempeño de la recolección censal, que deberán establecerse previamente, con el objetivo de que sea posible evaluar el avance y la calidad de la enumeración. Aun cuando estos indicadores no sean del todo precisos, son una fuente invaluable para comprender los resultados del censo y entregar información para la toma de decisiones.

Recuadro 7

Monitoreo de la recolección con dispositivos móviles de captura, México, Censo 2020

El sistema OPERA es una plataforma institucional que permite atender cualquier proyecto de generación de estadísticas. Este sistema apoya tareas de seguimiento y control de las diferentes etapas del proyecto. Para su acceso se requiere que el usuario sea parte del dominio institucional; es decir, que se haya creado una cuenta y una contraseña, y en base a sus credenciales, el sistema determina el perfil de usuario y presenta los módulos a los que tiene acceso.

Los principales módulos definidos en el OPERA para el Censo 2020 fueron:

- Integración de datos y distribución de materiales
- Reclutamiento y selección de personal
- Contingencias e incidencias operativas
- Seguimiento logístico
- Soporte Informático y Gestión de DMC
- Seguimiento operativo

El módulo de *Seguimiento operativo* es la herramienta que permite la detección oportuna de problemas de cobertura de los operativos censales, tomando como base las principales unidades de observación: Viviendas, Personas, Manzanas y Localidades que integran el Marco Geoestadístico Nacional. Particularmente, las viviendas y su condición operativa: visitada, censada, pendiente, deshabitada y uso temporal, tienen la mayor relevancia, ya que indican el avance del operativo de campo y proporcionan un panorama nacional que facilita la toma de decisiones que conlleve a una mayor cobertura.

También, se presentan informes sobre el avance en el proceso de actualización cartográfica detectada por el personal de campo, así como sobre la validación que de dichas actualizaciones realiza el personal de cartografía, quienes determinan aquellas que proceden y las que no.

Fuente: Texto editado a partir la Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe, GTC CEA-CEPAL véase Anexo 2.7.

Estos indicadores permiten obtener un panorama factual del avance de la recolección, día tras día (o según la periodicidad que se establezca) para las distintas áreas geográficas que componen el territorio de un país y establecer acciones correctivas, en caso de ser necesario. Por otra parte, estos indicadores pueden observarse desde la perspectiva de los resultados, ya que señalan los principales

atributos de las unidades censadas y concluidas. La contrastación de estos indicadores con los parámetros estimados orienta las actividades de supervisión durante la recolección, pero a su vez, adelantarán los trabajos de análisis de cobertura y calidad de los datos censales (CELADE, 2011).²⁵

2. Visualización geoespacial

La geografía sustenta la mayoría de las actividades asociadas a una enumeración censal, tanto en la planificación como en la gestión, al igual que para reportar los resultados. Las nuevas capacidades geoespaciales, como resultado de los avances tecnológicos en los sistemas de posicionamiento global (GPS) y los sistemas de información geográfica (SIG), junto a la disponibilidad de imágenes aéreas y satelitales de bajo costo, han permitido que los institutos de estadísticas puedan recopilar información más precisa y oportuna sobre sus respectivas poblaciones. En la actualidad, se están produciendo avances similares en las esferas de la difusión de datos geográficos con todas las principales herramientas SIG que ahora hacen accesibles las bases de datos geoespaciales o geográficas a través de Internet. Los institutos que están adoptando esta tecnología son capaces de brindar acceso a grandes cantidades de información espacial a los usuarios de forma barata, universal y rápida. Esta tecnología ha permitido a los productores de datos pasar de una base de mapas en papel a una base digital.

De esta manera, los SIG y otras herramientas geoespaciales han permitido una producción más eficiente, tanto de mapas para los enumeradores como de mapas temáticos de los resultados del censo. En este sentido, la visualización geoespacial debe considerarse como herramienta imprescindible para el monitoreo de la recolección censal y como parte del Sistema Integral de Gestión Operativa (SIGO).

Recuadro 8

Control de cobertura, Colombia, censo 2018

El control de la cobertura en campo del Censo Nacional de Población y Vivienda 2018, mediante el método de barrido, fue complementario a través del visor geográfico, en el cual los coordinadores departamentales podían observar a la Unidad de Cobertura Urbana o Unidad de Cobertura Rural que habían trabajado y cuáles les faltaban. Posteriormente, desde el Departamento Administrativo Nacional de Estadística Central (DANE), encabezada por la Dirección de Geoestadística, se le enviaba un reporte a los coordinadores departamentales que permitía tener un control efectivo de la cobertura.

Se dispuso este visor geográfico con fines de seguimiento geográfico a los reportes de la herramienta SPOT (dispositivo que incluye tecnología GPS, con la funcionalidad de enviar reportes sobre la posición geográfica del personal asignado a los grupos de trabajo del método de rutas y que dentro de sus funciones estuvo el manejo de dicho dispositivo) en los grupos de trabajo de rutas. Esta herramienta permitió observar en tiempo real el avance de los grupos de trabajo asignados en las áreas operativas de rutas, como una alternativa para que, desde el nivel central, se evidenciara que el personal estuviera recorriendo las áreas de trabajo.

El marco de referencia utilizado para el seguimiento geográfico a las rutas mediante esta herramienta corresponde al marco de las áreas operativas diseñadas para el método de rutas, junto con la información oficial, como los límites político-administrativos departamentales y municipales, el cual se complementó con la ubicación referencial de comunidades-veredas-rancherías y demás formas de concentración de la población residente en estos territorios.

Mediante la revisión del visor geográfico dispuesto con los reportes de la herramienta SPOT, la cual fue adelantada por personal del nivel central, se evidenció el avance de los grupos de trabajo para los cuales se observaron reportes en el visor, lo cual complementó los informes provenientes por parte del personal responsable del operativo en estos territorios (coordinadores departamentales, jefes municipales, profesionales operativos de rutas).

Fuente: Texto editado a partir la Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe, GTC CEA-CEPAL véase el Anexo 2.4 de este documento.

²⁵ Una lista exhaustiva de estos indicadores se encuentra en el anexo 4 de la Guía para asegurar la calidad de los datos censales, CEPAL (2011a), pág. 84-92, disponible en https://repositorio.cepal.org/bitstream/handle/11362/5515/1/S1100929_es.pdf

C. Indicadores de seguimiento demográfico

En términos generales, las fuentes de información estadística sociodemográfica siempre están sujetas a errores, dado que se basan en la comunicación entre un informante, que en los censos suele hablar por todos los demás integrantes del hogar y un entrevistador, que en operaciones masivas no siempre cuenta con suficiente capacitación (Chackiel, 2010).

Dado esto, la evaluación demográfica de los datos debe acompañar las actividades que están ligadas al trabajo en el terreno, estableciéndose controles de calidad sobre los que se tenga información cuando se trate de medir la cobertura de la población (Chackiel, 2010).

En este sentido, el uso de métodos de captura electrónicos aumenta la oportunidad de los datos no solo para realizar controles asociados al avance de la recolección, sino que también a la evaluación de datos desde la perspectiva demográfica. En esta línea, la experiencia de la sala situacional de la República Bolivariana de Venezuela en 2011 es algo a destacar dentro de la región.

Recuadro 9

República Bolivariana de Venezuela 2011

El monitoreo y control en el Censo 2011 se realizó en dos niveles, en el nivel local en los Centros de Recolección y Transmisión Censal (CRTC), y en el nivel central en la sala situacional o de seguimiento censal o en la sala de procesamiento, revisión y recuperación de los datos censales. A nivel local se realizó un seguimiento por semáforos de un conjunto de indicadores predefinidos que se adjuntan en la siguiente imagen, donde se le hacía el seguimiento cobertura y contenido del levantamiento de datos censales a partir de consultas o informes que procesaban los datos consolidados y mostraban las distribuciones o relaciones propuestas para el seguimiento, donde los valores resultantes eran comparados con valores esperados rangos inferiores o superiores. El semáforo se colocaba en rojo si dos o más criterios estaban por debajo de lo esperado, en amarillo si sólo uno estaba por debajo y en verde si todos estaban en el rango esperado.

Imagen 1
Hay que poner título a la imagen

INSTITUTO NACIONAL DE ESTADÍSTICA REPÚBLICA BOLIVARIANA DE VENEZUELA		XIV CENSO GENERAL DE POBLACION Y VIVIENDA INDICADORES: CUARTILES (percentil 5 - percentil 95) POR CENTROS DE RECOLECCION					IQ1 XIV Censo Nacional de Población y Vivienda 2011		
Fecha de Impresión: 15/09/2011 11:57:58		Total de CRTC : 78					Page 1 of 30		
Entidad: DISTRITO CAPITAL									
CRTC	Pers/Hogar	Hogar/vivienda	Niños/Mef	Mujhij/Mef	Ind.masc	Actividad	Cant. Viviendas Acumuladas	Empad. EN CRTC	
01001	3,76	1,21	31,51	58,04	96,23	62,96	826	14	Red
01002	3,61	1,08	22,75	49,10	85,59	64,34	905	16	Red
01003	3,91	1,02	37,81	62,75	100,90	52,83	1464	19	Verde
01004	3,63	1,03	37,05	65,10	101,58	56,27	929	18	Amarillo
01005	3,79	1,06	31,95	58,03	97,42	53,48	944	16	Verde

La información llega desde los CRTC al Sistema Central en dos formatos (SQL y *.dat) de forma encriptada. La primera, SQL era presentada en la página web para ver la calidad del contenido de los datos. Luego con CSPro se verificaba y corregían los datos erróneos u omitidos y se generaba en REDATAM para actualizar una aplicación R+SP xPlan de Redatam para el análisis complementario. Se verificaban los distintos niveles de desagregación geográfica (entidad, municipio y parroquia) con informes por cortes semanales, donde se compara con el estimado en base a las proyecciones de población oficiales de 2001 y con la tendencia histórica censal y se determina si se encuentran dentro de los intervalos de confianza. Las validaciones de saltos y consistencias estuvieron programadas en el DataEntry de CSPro por registro, luego en la etapa post-censal se realizó entre registros y la limpieza final.

Fuente: Texto editado a partir la Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe, GTC CEA-CEPAL. Véase el anexo 2.9 de este documento.

VI. Consideraciones de gestión y planificación vinculadas con la captura de datos censales

El presente capítulo representa la integración, interpretación y priorización de los lineamientos comprendidos en los capítulos C y F del “*Guidelines on the use of electronic data collection technologies in population and housing censuses*” (United Nations, 2019), de la “*Guía para elaborar un proyecto censal*” (CELADE-CEPAL, 2011), de la experiencia transmitida por los países en los talleres realizados por el GTC CEA-CEPAL y de la experiencia acumulada por el CELADE, de modo que, en algunos puntos del texto, se toman fraseos o párrafos con traducciones libres de los documentos y, en otros, se fusionan o reinterpretan en función de la experiencia regional.

En este contexto, el objetivo del capítulo es ofrecer lineamientos concretos para los profesionales y técnicos de los institutos de estadística de la región de América Latina y el Caribe en la planificación y gestión de un censo al momento de decidir incorporar mayor tecnología en los métodos de captura, principalmente eCenso y DMC.

El interés en incorporar este tema a los lineamientos se debe a que, si bien las fortalezas y aportes para mejorar la calidad y eficiencia son indiscutibles, su incorporación compromete importantes requerimientos en cuanto a capacidad y gestión institucional, diseño estratégico, planificación, gestión del cambio y presupuesto, entre otros elementos, requerimientos que, en la experiencia reciente de la región, no siempre se han valorado con oportunidad.

Por otro lado, este capítulo no pretende ser una guía para el desarrollo de la planificación de un censo ni de un proyecto de inversión en tecnología para un censo o para un INE. El motivo que anima a sus autores es alertar sobre aquellos aspectos puntuales que se deben considerar en la planificación general del proyecto censal debido a que serán impactados al avanzar de un censo en papel (sea con

escáner o digitación) a un censo con DMC o eCenso²⁶. En este capítulo, tampoco, se abordarán aspectos específicos del plan del subproyecto de innovación tecnológica en DMC o eCenso dentro del plan general del censo.

Como ya se ha mencionado anteriormente, y en la literatura con respecto a los censos tradicionales de población y vivienda, estos son procesos de alta complejidad y de gran magnitud. La incorporación de mayores niveles de tecnología, en todas sus etapas, abre ventanas de oportunidad para mejorar la eficiencia, eficacia y calidad, a la vez que impone nuevos retos, los que deben ser correctamente ponderados y administrados. La adopción, especialmente, de los DMC como el método de captura principal, pero también el eCenso, implican la readecuación en tiempo, cronología y forma de una serie de actividades, agregación de nuevas pruebas, de cambios en la manera de gestionar, la identificación de nuevas habilidades y puestos de trabajo, mayores niveles de inversión en tecnología, identificación y administración de riesgos y desafíos emergentes, las que deben sumarse a la planificación y gestionarse de manera exhaustiva, oportuna y rigurosa.

Como se expresa en el documento “Guía para la elaboración de un proyecto censal” (CELADE--CEPAL, 2011), para poder desarrollar con éxito las actividades de planificación censal se precisa haber tomado una serie de decisiones metodológicas previas con respecto a los objetivos, al tipo de censo, al tiempo de trabajo en terreno y a la posibilidad de introducir innovaciones tecnológicas en el método de captura, entre otros aspectos, además de haber observado cómo estos componentes influyen positiva o negativamente sobre el desarrollo de las diversas actividades del ciclo de vida del proyecto censal, por lo que una planificación cuidadosa y completa es vital para asegurar una toma de decisiones consistente, informada y que tienda a un resultado satisfactorio.

A. Consideraciones para la gestión institucional y del censo al incorporar un método de captura electrónico

La incorporación de un nuevo método de captura, junto a todo lo que ello implica —en cuanto a requerimientos de sistemas de tecnologías de la información en el proceso de producción de datos censales—, debe partir por reconocer un hecho esencial: pasar del papel a DMC (o eCenso) cambia no solo la captura y parte del procesamiento de los datos, sino que cambia toda la forma de gestionar la mayor parte del proyecto censal y todo el trabajo en terreno. Esto, generalmente, implica el rediseño de las estructuras organizacionales, de flujos y procesos de trabajo, de roles y con ello incide en la cantidad, tipo y habilidades requeridas por el equipo humano involucrado, influyendo desde la alta dirección hasta los mismos entrevistadores, en el caso de los DMC y, por supuesto, repercutiendo en el calendario general del censo.

En este sentido, cuando una oficina de estadísticas decide evaluar la posibilidad de evolucionar de un censo en papel a uno con DMC o eCenso debe tener presente el impacto que esta herramienta tendrá en toda la organización y en la planificación del censo. En consecuencia, se deben asegurar los resultados deseados y mitigar los riesgos que implica un cambio de esta magnitud.

El tránsito del uso de papel a un dispositivo móvil de captura obliga a preguntarse si cada INE tiene la capacidad para adoptar esta nueva tecnología y lograr incorporarla en el proceso de gestión institucional, más allá del censo, de forma sostenible durante y después. Este es un punto relevante en la medida en que la institución estadística carezca de experiencia suficiente en el uso masivo de DMC o levantamientos a través de la web, dificultándose la incorporación de mayor tecnología. Por lo que, es

²⁶ Estas y otras cuestiones generales relacionadas con la planificación y gestión de las colecciones de datos del censo se elaboran y discuten en detalle en United Nations (2017) y (2001).

importante hacer un análisis de las condiciones y características técnicas, humanas, organizacionales, presupuestarias y legales existentes en la institución, de sus brechas y de la posibilidad real de adoptar la tecnología e implementarla de manera estratégica en el proceso de producción estadística, asegurando la continuidad operacional del censo, de otros productos estadísticos y, obviamente, de la calidad y oportunidad.

1. Estructura organizacional

Como ya se ha dicho, la incorporación de un método electrónico principal de captura impactará en los procesos a lo largo de todo el ciclo de vida del proyecto censal, por lo que esta nueva forma de “hacer” debe ser asumida, corporativamente, por la institución y sumarse de manera sistemática a los procesos de trabajo ya existentes, lo que implica además que debe existir un estamento de gestores o personal directivo con las habilidades indispensables para gestionar un proyecto cuyo énfasis está en el manejo de grandes cantidades y flujos de información, por lo que tales habilidades pueden, o no, estar presentes en el instituto, haciendo que el proyecto censal requiera equipos especializados en gestión de compras de bienes y servicios más complejos e involucren equipos de contrapartes permanentes con los proveedores.

En este contexto, el análisis de procesos debe ayudar a establecer, por ejemplo, las relaciones entre la plana directiva del proyecto censal y los profesionales de TI, para permitir maximizar los beneficios de los nuevos procesos de gestión. También, es altamente recomendable que proporcione un plan de implementación que ayude a los institutos a coordinar y orientar los esfuerzos de mejora, al incluir los nuevos procesos de gobernanza organizacional. De igual modo, es deseable que exista el compromiso de la alta dirección (superior al INE, del INE y del censo) y la aceptación de los actores para superar los posibles retos asociados en la implementación de un nuevo proceso de negocio en una actividad tan crítica como lo es el censo.

Dada las complejidades propias de un censo, sumadas a las de un proceso de innovación tecnológica, el equipo de implementación, para que sea eficaz, debe incluir tanto a los actores experimentados de las áreas técnicas del censo (temáticas, operativas, capacitación, logísticas, etc.) como a los de TI. Si bien, los cambios tecnológicos en los censos se realizan más rápidamente que los cambios en los métodos y principios estadísticos, tales innovaciones tecnológicas nunca deben poner en riesgo la continuidad del proceso estadístico y, siempre que sea posible, deben reforzarlos.

Recuadro 10

Recomendación a otros países que deseen incorporar el DMC como método de captura, el Brasil, censo 2021

El primer punto es, sin duda, tener una fuerte integración entre el área técnica y la de informática. Cuando se usa DMC, todas las posibilidades que estos dispositivos proporcionan dependen del desarrollo de algoritmos y presentan una gran demanda en términos de recursos humanos.

Fuente: Texto editado a partir la Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe, del GTC CEA – CEPAL véase Anexo 2.1.

En definitiva, el acto de pasar del uso de papel al DMC o eCenso como método de captura impacta, prácticamente, todo el proceso censal, cambia el orden de los procesos y, demanda nuevas habilidades y cantidades de equipo humano (no solo del personal de TI), mayor inversión en bienes y servicios tecnológicos adyacentes al DMC o eCenso, solo por nombrar algunos aspectos, de tal forma que, al momento de decidir el método que se va a usar, es necesario levantar la mirada hacia la vida completa del proyecto para incluir en la planificación y organización todas las aristas que permitan hacer una innovación tecnológica con éxito y, a la vez, mitigar los riesgos.

2. Gestión del cambio

La implementación de un nuevo método de captura, a pesar de mejorar, evidentemente, el proceso de producción estadística se enfrenta a múltiples desafíos, incluyendo la natural resistencia al cambio. A menudo, los mayores desafíos surgen no de la gestión de los aspectos técnicos u operativos de la innovación, sino de la gestión de las dimensiones humanas del cambio.

Para superar esta resistencia, los institutos necesitan comenzar a gestionar el cambio desde el mismo principio de rediseño del proceso censal y durante la fase de implementación. El plan de gestión del cambio debe vincular el nuevo proceso a cuestiones específicas, preguntas y desafíos involucrados en la implementación, como roles de trabajo, relaciones, expectativas de desempeño, métodos de supervisión y trayectoria profesional. A través de su implementación, se recomienda re-educar, alentando al personal para aceptar las nuevas ideas y para adoptar el nuevo proceso. También, es recomendable entrenar al personal y a los gerentes, para gestionar los cambios en sus respectivas funciones y expectativas profesionales.

El manejo de la gestión del cambio urge no solo por la aversión a lo novedoso o por falta de conocimiento, convencimiento o interacción con la tecnología. También se debe a que, a pesar de estar convencidos y decididos a nivel intelectual del cambio, el tránsito del papel a los formularios electrónicos implica actualizar la forma de pensar y de hacer un censo, modificar el acento y la valoración del trabajo pre censal y de las pruebas tecnológicas. Para ello, se debe dar énfasis a la seguridad informática y no solo de las personas, algo que cambia, radicalmente, la manera cómo se gestiona y controla el trabajo de terreno. Lo anterior implica pensar en las herramientas, métodos, errores e indicadores desde otra perspectiva, siendo natural que los equipos técnicos y profesionales se demoren un tiempo en descubrir e internalizar esta nueva forma de trabajo, por lo que es un requisito primordial gestionar, adecuadamente, todo el proceso censal.

B. Impacto del método de captura de datos en el diseño de los procesos censales.

El objetivo de la planificación es garantizar no sólo que cada fase esté organizada y tenga los recursos (humanos, físicos y presupuestarios) adecuados, sino también asegurar que los productos de cada fase sean de calidad suficiente para todas las etapas posteriores y dependientes. Hay que mencionar además que la planificación debe ser lo suficientemente flexible y que permita contemplar las holguras necesarias para incorporar los cambios que se requieran producto de las diferentes pruebas, así como por los hallazgos y aprendizajes que siempre ocurren en las innovaciones. Este último aspecto es de vital importancia cuando se está evaluando la inclusión del uso de DMC o eCenso por un instituto que aún no tiene una experiencia robusta en el uso de tecnología para realizar un proceso tan masivo e importante como el método electrónico de captura de datos, ya que impacta justo en el corazón del censo, tanto en la forma como en los tiempos y subproductos previos y posteriores a la etapa de recolección de datos.

El desarrollo de toda planificación y de los procesos de decisión, sobre en qué áreas innovar para un futuro censo, debe comenzar, necesariamente, por la evaluación del censo precedente, con el objeto de establecer un diagnóstico de sus problemas y dimensionar sus impactos, magnitudes y pertinencia en el nuevo contexto y así determinar las prioridades, objetivos y estrategias para la siguiente década. En este marco, y luego de la evaluación, se debe realizar un detallado análisis de los procesos, con la finalidad de determinar exactamente qué actividades, métodos, hitos, sistemas o herramientas son afectadas (positiva o negativamente) al innovar con un método electrónico de captura. En función de los resultados de dicho análisis, se deberá readecuar la planificación, cronogramas, estructuras, inversiones y personal, entre otros.

1. Calendario censal

En lo que concierne al cronograma censal, se debe realizar un detallado análisis de procesos que permita ajustar el calendario del censo a los requerimientos de implementación del nuevo método de captura. Este punto es especialmente importante en aquellos casos en que los INE tienen poca experiencia con el uso de DMC, dado que algunos tiempos se desplazan y otros se traslapan en el calendario.

Los tiempos de preparación para un censo con dispositivos electrónicos y digitales son más largos que con papel, por lo que se debe comenzar tempranamente con el desarrollo del software del cuestionario, aun cuando no se haya terminado con la definición temática del cuestionario, pues se requiere un tiempo considerable para desarrollar y probar la aplicación, configurar los sistemas de transferencia de datos y adquirir, programar y probar cada uno de los DMC.

Realizar las pruebas de cuestionario y depurar el software puede tomar un tiempo considerable hasta lograr que, por un lado, el flujo de preguntas y los patrones de omisión funcionen bien y que todos las modificaciones y cambios en las versiones del cuestionario y, por otro, que sus programas no tengan impactos negativos y el programa siga funcionando correctamente. El desarrollo de un cuestionario electrónico es iterativo; no obstante, presenta la dificultad que existen alertas y validaciones que hacen que las pruebas demoren más tiempo, o que se cometan errores que afecten su correcto funcionamiento. Es importante recordar que, a diferencia del papel, cuando en un formulario electrónico se comete un error en la captura de la información, ya no hay vuelta atrás, en la que se pueda volver a escanear o digitar el dato censal, por lo que su correcto desarrollo lo hace aún más crítico.

Por su parte, el diseño y desarrollo del sistema de seguimiento operativo y técnico también puede consumir tiempos importantes del proyecto censal y su desarrollo está directamente conectado e integrado con el sistema del cuestionario, transmisión, seguridad y procesamiento por nombrar algunos. Quizás, la principal ventaja de los formularios electrónicos, especialmente en el caso del uso de DMC es la entrega en línea de información operativa y técnica, permitiendo gestionarla en terreno y mejorar la calidad de los datos, en la medida en que el sistema esté bien definido y acorde a la estructura jerárquica a la que debe alertar y tomar decisiones. Este desarrollo también tiene sus dificultades propias y debe ser sometido a diversas pruebas, puesto que debe conversar, adecuadamente, con la compleja estructura jerárquica asociada tanto al trabajo de campo, desde supervisores hasta la jefatura máxima, como por los analistas de calidad y seguimiento. Dicha situación repercute, además, en las habilidades de la estructura jerárquica para manejar, comprender e interactuar con sistemas de información más complejos y dinámicos que los usados en los censos con papel.

2. Habilidades informáticas y desarrollo de capital humano

Al momento de implementar un método electrónico de captura de datos, con DMC o eCenso, los institutos nacionales de estadística tendrán que actualizar la evaluación de habilidades para identificar las brechas de competencias, destrezas tecnológicas, de administración y manejo de información, especialmente en aquellas áreas que aseguran la continuidad operacional y desarrollar, al mismo tiempo, planes y estrategias para cerrar cualquier brecha que sea crítica para enfrentar con éxito la recolección y el proceso censal completo.

Dependiendo de su estructura organizacional, algunos institutos tendrán que desarrollar capacidades sostenibles de gestión corporativa de la información y de habilidades de TI en toda la institución, incluidas sus oficinas regionales y locales para que consideren las necesidades de organización, implementación y control del censo.

Las mejoras sostenibles en la gestión de TI son imposibles de implementar sin optimizar los conocimientos y habilidades de gerentes, directivos y jefaturas intermedias, sin el establecimiento de una sólida relación de trabajo con los profesionales de TI de la institución. La actualización de destrezas de las jefaturas intermedias en las diferentes áreas censales permitirá desarrollar capacidades de

administración de TI a nivel organizacional. Los diferentes niveles jerárquicos de la organización necesitan disponer de las competencias primordiales para identificar problemas de administración, oportunidades y decisiones de gestión de la información. Esto garantizaría que las jefaturas de las áreas censales (temática, capacitación, operaciones, etc.) tengan una mayor comprensión de TI, mejorando su gestión e impacto en la misión, los objetivos y los problemas de las unidades.

Las gerencias o jefaturas involucradas en la adquisición, contratación y gestión de los sistemas y servicios de TI necesitan mantenerse al tanto de los desarrollos en TI. Esto es, especialmente, relevante cuando los institutos dependen, en un grado significativo, de contratistas y proveedores externos. El riesgo de un quiebre entre la institución y los contratistas aumenta cuando los institutos no disponen de profesionales de TI competentes que puedan ayudar a las jefaturas de área para evaluar y supervisar el desempeño de los contratistas.

C. Gestión de la inversión en Tecnologías de Información

Hoy es un hecho y existe evidencia suficiente de que la inversión en TI mejora el rendimiento institucional. Sin embargo, cuando dicha inversión no se gestiona eficazmente, pueden deslizarse en el calendario, llegando a ser riesgosa, costosa e improductiva. La adquisición de dispositivos, bienes y servicios tecnológicos asociados, directa e indirectamente, se debe realizar con la participación de la industria adecuada y con la concurrencia de profesionales especializados en adquisiciones (del INE o subcontratados), en el marco de una suficiente comunicación entre el personal de TI y del censo, quienes efectivamente utilizarán tanto el hardware como el software y los servicios.

Para tener éxito en la incorporación de métodos de captura, como los DMC o eCenso, la mayoría de las oficinas de estadística necesitarán identificar la demanda respecto de las innovaciones tecnológicas que se requerirá incorporar, así como los métodos, herramientas o sistemas que resultarán impactados, junto con poner en marcha un proceso de gestión sistemático y disciplinado mediante la creación de una estructura de gestión, incluyendo un grupo de gestión (o junta directiva), integrado por altos directivos del instituto; entre ellos, el director(a), el jefe(a) del censo, los jefes del departamento de TI, los de las unidades operativas y de apoyo, quienes permitirán gestionar debidamente tanto las compras de equipos o servicios, los insumos de hardware y software como el diseño conceptual que involucra su implementación y sostenibilidad, la selección y capacitación del personal técnico.

Uno de los principales objetivos de incluir las inversiones y desarrollos de TI en la planificación y gestión del censo es garantizar que los proyectos y sistemas estén alineados con las necesidades de producción del censo y no se transformen en un fin en sí mismo. La participación de los diversos equipos involucrados en la etapa de concepción del proyecto TI (operativos, capacitación, gestión, control operativo y de procesamiento), debe estar marcada por su compromiso en el desarrollo del caso de negocio y en la definición de cómo el sistema ayudará a satisfacer las necesidades del censo, en términos de usabilidad, productividad, desempeño, seguridad y manejo de la información. También deben estar muy involucrados durante las pruebas de aceptación del usuario. Durante la fase operativa y de mantenimiento del ciclo de vida del sistema, los usuarios deben desempeñar un papel activo en ayudar a identificar y documentar los beneficios que se logran de la implementación del sistema.

Recuadro 11**Recomendaciones a países que deseen incorporar el DMC como método de captura, México censo 2020.****Primera recomendación:**

Considerar que, desde el momento en que la captación de datos se lleva a cabo con el uso de dispositivos de cómputo móvil, se vuelve indispensable definir una estrategia informática para sistematizar el resto de los procesos del proyecto censal, desde la planeación operativa, el mapeo de riesgos operativos y la captación de información, hasta los procesos primarios, codificación, validación y generación de resultados. Una estrategia que responda a las necesidades funcionales y contemple un esquema de seguridad robusto, tal que contribuya a la calidad de los datos captados y a la cobertura total de áreas y unidades de observación. Para ello, es importante contar con un requerimiento de desarrollo de software por cada sistema a elaborar.

La estrategia debe incluir, además, una estructura de soporte informático lo más cercano posible a la estructura de levantamiento, con objeto de agilizar la atención de cualquier incidencia en materia informática que se presente, así como estructurar una estrategia de capacitación tal que garantice la transmisión de conocimientos y la adquisición de habilidades para el correcto desempeño de sus funciones, sin omitir que, desde su reclutamiento, se debe solicitar un perfil informático.

Segunda recomendación:

En términos de tecnología, para el uso de dispositivos de cómputo móvil y la autoenumeración para la captación de datos, los aspectos que se deben considerar son:

- Características mínimas de los dispositivos de cómputo móvil. Deben cuidarse en las especificaciones que se hacen en el proceso de licitación, ya que ello determina las características funcionales de las aplicaciones informáticas a implementar.
- Desarrollo de aplicaciones de software. Es importante determinar la plataforma de desarrollo para dispositivos de cómputo móvil, ya que puede enfocarse para un determinado sistema operativo (Android, IOs o Windows). Esto aplica también para las plataformas de desarrollo de aplicaciones en ambiente web.
- Infraestructura de TI. Considera la infraestructura de cómputo requerida para el hospedaje de aplicaciones web (servidores de aplicaciones), integración de la información en una base de datos centralizada (servidores de bases de datos); almacenamiento (dispositivo de almacenamiento conectado a una red) y respaldo de archivos (dispositivos de respaldo como cintas magnéticas), así como el equipo de escritorio y de cómputo móvil. Con respecto a las redes y comunicaciones, se deben considerar servicios de voz y datos en oficinas del Instituto y de campo, el servicio de correo electrónico institucional, datos móviles para la telecomunicación en los DMC, entre otros.
- Cobertura de los servicios de internet. Es importante tener en cuenta la capacidad de la infraestructura instalada, en términos de tecnologías de la información y comunicación (TIC), así como la cobertura nacional, ya que será determinante para la estrategia de autoenumeración.

Tercera recomendación:

Con el objetivo de verificar el correcto funcionamiento de toda la infraestructura de TIC; es decir, el hardware, el software, los elementos de red, los diferentes sistemas operativos utilizados y el almacenamiento de datos, se debe definir un plan de pruebas que considere:

- Pruebas de sistemas. Su objetivo principal es asegurar que cada aplicación informática satisfaga las necesidades funcionales incluidas en el(los) requerimiento(s) de desarrollo de software realizado(s) por las áreas responsables de cada uno de los procesos a sistematizar en el proyecto censal.
- Pruebas de volumen. Hacen referencia a grandes cantidades de datos para determinar los límites en que causan que un sistema falle. Sirven para identificar la carga máxima o volumen que el sistema puede manejar en un periodo determinado, principalmente el más crítico, así, se realizan para verificar que la aplicación informática funcione adecuadamente bajo los siguientes escenarios de volumen: máximo número de clientes conectados y tamaño máximo de base de datos.
- Hacking ético (pruebas de penetración). Su objetivo es identificar y reparar posibles vulnerabilidades en la capa de servicios implementada para el proyecto, a fin de fortalecer y mejorar la estrategia de seguridad implementada.

Fuente: Texto editado a partir la Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe, del GTC CEA –CEPAL véase Anexo 2.7.

D. Consideraciones relacionadas con adquisiciones, desarrollos y externalizaciones

Hasta ahora, se ha visto que la velocidad de los cambios tecnológicos y la decisión de implementar mejoras o cambios tecnológicos en los censos conduce a la adquisición de licencias de sistemas disponibles en el mercado, ajustes a los existentes en la institución o al desarrollo e integración de sistemas, ya sea con equipos internos o externalizados (*outsourcing*). Si bien, la externalización puede permitir lograr ciertos ahorros y resolver algunos cuellos de botella, al mismo tiempo plantea diversos problemas, incluidos los relativos a la confidencialidad. Una vez más, la situación del país y del Instituto (incluida su capacidad de gestión) serán factores determinantes de la viabilidad de recurrir a la externalización.

Se recomienda, en consecuencia, que esta definición se base en rigurosas investigaciones de mercado, pruebas, revisiones técnicas, análisis y consideración de factores tales como importancia estratégica, costo y disponibilidad de tiempo y recursos de programación interna. La decisión final de construcción o desarrollo frente a la compra es crítica, pues elegir incorrectamente podría originar una solución inadecuada o un rendimiento deficiente de la inversión de los institutos. Una vez que se ha identificado un requisito funcional claro, se debe realizar una investigación de mercado para garantizar que se identifiquen tantas soluciones alternativas viables como sea posible para su consideración. Cuando se disponga de soluciones comerciales aceptables, se debe priorizar el uso de software basado en estándares y soluciones comerciales listas para usarse de inmediato, mientras estén disponibles con poco riesgo en costos, cronogramas, rendimiento y obsolescencia técnica por sobre el desarrollo personalizado.

En muchos casos, un enfoque híbrido, que consiste en integrar las mejores plataformas “*off the shelf*” con soluciones específicas internas desarrolladas con la participación de expertos en el censo, es el mejor enfoque recomendado para construir la infraestructura necesaria que permita apoyar la recopilación de datos del censo.

Dependiendo del enfoque de cada instituto, para enfrentar sus necesidades tecnológicas, de infraestructura y de la capacidad instalada que ya disponga, es probable que algunos de ellos deban realizar adquisiciones de TI a gran escala, siendo un factor de éxito en su implantación la correcta y oportuna definición de los requerimientos y su adquisición o desarrollo temprano, proporcionando la holgura suficiente de tiempo para realizar un conjunto de pruebas que aseguren el cumplimiento de los requisitos definidos y la funcionalidad esperada. Se recomienda que cada INE disponga de un proceso de adquisiciones disciplinado y, en la medida de lo posible, simplificado para planificar, adquirir y utilizar, eficazmente, los activos de TI para lograr el máximo retorno de la inversión y, en caso de no poseerla, considerar un seguimiento exhaustivo al proceso administrativo que se requiera²⁷.

E. Plan de pruebas

Una de las características de transitar de un método de captura en papel a uno con DMC o eCenso es que el plan de pruebas sufre cambios importantes, reenfoándose a los sistemas de información y funcionamiento de las herramientas y servicios tecnológicos y adelantándose en el cronograma respecto al papel. Este es un aspecto muy relevante al realizar la planificación del proyecto, la que se ve influenciada por la capacidad de los equipos de internalizar la nueva forma de pensar el proyecto y la gestión del cambio. En un censo con DMC o eCenso, los sistemas y la infraestructura de TI desempeñan un rol fundamental y de mucho mayor impacto que en un censo con papel. Para asegurar el éxito, los sistemas y la infraestructura de TI, tanto individualmente como cuando están integrados, deben funcionar según lo previsto.

²⁷ En United Nations (2017) y (2001) y en CEPAL (2011b) se puede encontrar información más detallada acerca de la subcontratación en los censos de población y vivienda.

Las pruebas completas y exhaustivas son esenciales para proporcionar una seguridad razonable de que los sistemas de TI funcionarán como se espera, por lo que se debe establecer un plan de pruebas con estándares de calidad para cada elemento que se va a probar; de modo que, todas las características puedan evaluarse de manera coherente, con un calendario claro que establezca un ordenamiento lógico de prueba, desde los sistemas individuales hacia la integración incremental de esos sistemas, , abordando así todos los sistemas interrelacionados. Estas pruebas deben realizarse en entornos operativos los más cercanos y parecidos a la realidad que se enfrentará en el censo y, obviamente, con el tiempo suficiente para incorporar los cambios en los sistemas cuando sea necesario e incluso recurriendo a implementar planes alternativos. En estas pruebas, también es recomendable estresar a los sistemas sometiénolos a contingencias posibles, pero muy esporádicas artificialmente, con el fin de asegurar que los planes de contingencia y de continuidad operacional estén bien definidos e internalizados. Entendiendo por contingencias posibles, pero esporádicas, los terremotos, cortes de energía eléctrica de plazos no tan cortos y/o la falta de acceso a internet.

Recuadro 12

Pruebas realizadas durante la preparación de la recolección con DMC, Departamento Administrativo Nacional de Estadística (DANE), censo 2018

La mayor prueba de los dispositivos móviles de captura se realizó en Colombia, durante el Censo experimental de Jamundí, Valle del Cauca (2016), donde se evidenciaron las ventajas de su uso y las fallas (recalentamiento de máquinas y bloqueos del cuestionario) y limitaciones en el software, situaciones que fueron corregidas por el equipo de sistemas.

Se realizaron las siguientes pruebas:

- Pruebas a los cuestionarios análogos
- Pruebas temáticas de las preguntas incluidas y excluidas
- Pruebas de escritorio para identificar la comprensión de las preguntas
- Pruebas de campo con el fin de captar comprensión y tiempos de diligenciamiento

Para el DMC, se elaboraron las normas de validación de flujos, rangos y consistencia del cuestionario análogo y se gestiona con la Oficina de Sistemas. Una vez que se desarrolla el programa de captura del cuestionario análogo, se procede a verificar y validar que todas las normas se hayan incluido mediante pruebas de escritorio, igualmente se prueba la capacidad del dispositivo y el tiempo de carga de las baterías para evitar inconvenientes en el desarrollo de la encuesta.

Las pruebas se describen en una matriz que identifica los errores que se presentan, para que sean tenidos en cuenta por la Oficina de Sistemas y así, una vez se superen todos los inconvenientes, se da vía libre al programa de captura, diseñado para tal fin.

Por otra parte, una vez realizadas las pruebas de campo, se evalúa y analiza la captura con datos reales, se valida tanto el programa como los tiempos en el diligenciamiento y manejo del dispositivo.

Fuente: Texto editado a partir la Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe, del GTC CEA-CEPAL, véase el Anexo 2.4.

Recuadro 13

Prueba de funcionalidad de dispositivos móviles, México, censo 2010

Del 18 al 28 de noviembre de 2008 se realizó, en algunas áreas seleccionadas del municipio de Salamanca, Guanajuato, la denominada Prueba de Funcionalidad con el objetivo de evaluar las ventajas y desventajas en el uso de unidades móviles de captura, durante el proceso de recolección de información, como parte de los trabajos preparatorios para el Censo de Población y Vivienda 2010 en México. Para la captación de la información, se combinó el uso de PDA y mini laptop, en alrededor de 3.500 viviendas.

El resultado no fue tan positivo como se había calculado y se definió que se hará como antes, a papel. Además de los problemas técnicos —la transmisión de la información demoraría demasiado sin una tarjeta USB para conexión inalámbrica—, se imponen las restricciones presupuestarias.

Fuente: Texto editado a partir la Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe, del GTC CEA-CEPAL, véase Anexo 2.7 de este documento.

VII. Conclusiones

El uso de la tecnología es uno de los más grandes desafíos que enfrenta nuestra región para la ronda de censos 2020, junto a las dificultades que ha significado la pandemia por COVID 19 a los INE de cada país. Si bien, cada método de captura tiene sus ventajas y desventajas e impacta de manera diferente en la calidad y oportunidad, no existe un método único que pueda recomendarse. Se deben considerar y evaluar los múltiples factores y situaciones en que están insertos los INE de la región, especialmente, en lo relativo a las grandes brechas de capacidad tecnológica instalada en los institutos; algo que, ocasionalmente, se puede transformar en un factor de éxito del proyecto. No obstante, es importante mencionar que, a pesar de lo anterior, el tipo de censo que se realice (de hecho o de derecho) sigue siendo una gran condicionante en la decisión del método de captura que utilizan los países de la región.

A lo largo del presente documento, se ha podido abordar en detalle la forma en que la elección del tipo de tecnología que se usa como método de captura para la recolección de los datos censales, impacta directa o indirectamente en casi todo el proceso censal: cambia el orden de los procesos, demanda distintas habilidades, competencias y cantidades de los equipos humanos (no solo del personal de TI); determina el nivel de inversión en bienes y servicios tecnológicos, adyacentes al DMC o eCenso; modifica la necesidad de logística y cambia el rol del supervisor, solo por nombrar algunos de los aspectos. Es por ello que, al momento de decidir el método de captura que se utilizará, se hace necesario orientar la mirada hacia el lapso completo de vida del proyecto censal; de manera tal que, dicha decisión se pueda incluir, tempranamente, en la planificación y organización del proyecto, estimando todas las aristas que permitan la innovación tecnológica y mitigando los riesgos para desarrollar un censo exitoso.

En el caso de los países de la región que, actualmente, están evaluando el cambio de método de captura —por ejemplo, de escaneo a DMC—, es importante recordar que deben poner énfasis, durante la evaluación, en la factibilidad que las ganancias del cambio de tecnología radiquen en la potencialidad que brinda el acceso inmediato al dato censal, mejorando calidad, cobertura y oportunidad, junto con considerarlo como un gran proceso de inversión en infraestructura estadística, la que perdurará por, al menos, 10 años, en lugar de estimar que pudiera ser un ahorro directo en costos económicos o del

tiempo que implica la realización del censo. Por consiguiente, se hace necesario destacar que la incorporación de tecnología al proceso censal, especialmente en la captura de datos, es un medio para asegurar el éxito del censo y no un fin en sí mismo.

Por otra parte, en el caso de aquellos países de la región que ya han utilizado o que planean utilizar métodos mixtos de captura, ya sea simultáneos o de manera secuencial —lo cual conlleva a una complejización del proceso de recopilación de información—, se debe controlar el avance del censo (cobertura) aplicando diversos métodos y convivir con fuentes de error que tienen diferente comportamiento, según el método de captura de origen, entre otros factores. De manera que, el uso de un método mixto debe estar, fuertemente, condicionado por la capacidad de control, seguimiento y procesamiento de cada INE. Por ello es que, en estos casos, la implementación secuencial de captura sería la técnica más recomendable para los países de la región.

Si bien, cuando se habla de métodos mixtos, generalmente el papel continúa estando presente como método de captura, se observa una mayor intención por parte de los INE de la región para implementar sistemas de eCenso (autoempadronamiento web) con el objeto de facilitar las respuestas de los diferentes grupos poblacionales, en especial para aquellos con más difícil acceso para los encuestadores, sea por razones de seguridad, apatía o ausencia de moradores, a lo cual se suma, además, la necesidad del distanciamiento social impuesto por la pandemia del COVID 19.

Por otra parte, y en relación con el uso del eCenso en la región, se observa su uso como método subsidiario de captura, con muy bajas tasas de respuesta en los países que lo han utilizado, por lo que no debiese ser considerado aún como un método único de captura plausible de implementar para futuros censos en los países de la región, al menos en la ronda 2020. Esto se debe a diferentes razones, las que van desde la baja cobertura de internet —que existe en algunos de los países, método que no ha sido totalmente probado—, y su incorrecta aplicación por parte del informante que podría comprometer fuertemente los resultados de la operación censal; por ejemplo, al contestar solo por una persona generando un aumento erróneo de los hogares unipersonales, además de las dificultades que implica lograr conectar la ubicación geográfica de la vivienda con el cuestionario y la posibilidad de invisibilizar los segundos hogares.

Los países de la región, que han innovado e incorporado recientemente el uso de DMC como método de captura en sus procesos censales, centran sus principales recomendaciones en tres aspectos. En primer lugar, lo que dice relación con la necesidad de integrar todos los sistemas que se utilizan a lo largo del proceso censal, desde la visualización geográfica, el seguimiento de la recolección, el monitoreo, hasta la información para la fase de procesamiento y análisis. El segundo aspecto relevante dice relación con los planes de pruebas, los que deben incluir pruebas de sistemas, dispositivos, volumen y transferencia de datos, elementos de seguridad de datos como el hacking ético y, posteriormente, la integración de distintos equipos técnicos que forman parte del proceso censal. Finalmente, acerca del DMC, los países recomiendan considerar soluciones y herramientas de cuarta generación y planificar, investigar y consolidar un presupuesto para su adquisición.

En definitiva, el éxito de la implementación de mayor tecnología en la captura de datos debe pasar, necesariamente, por un proceso integral de análisis y decisiones, donde no solo se consideren los aspectos tecnológicos y su potencial de mejora en calidad y oportunidad de datos. Es necesario, además, garantizar los apoyos institucionales necesarios para implementar esta innovación, los recursos financieros y de tiempo necesarios, así como evaluar, ponderadamente, la capacidad institucional del INE para la integración de esta nueva forma de hacer y pensar en el proceso censal en todos los niveles de la estructura, potenciando así las nuevas habilidades y capacidades requeridas en los recursos humanos y generando la integración que se requiere de los equipos de TI en todas las áreas del censo: temática, operaciones, logística, capacitación y administración. Sin esta integración no se podrá obtener el potencial de ventajas que ofrece el uso de tecnología ni asegurar el éxito del proceso censal.

Bibliografía

- CEPAL (Comisión Económica para América Latina y el Caribe) (2019), *Aspectos conceptuales de los censos de población y vivienda: desafíos para la definición de contenidos incluyentes en la ronda 2020*, serie Seminarios y Conferencias, N.º 94 (LC/TS.2019/67), Santiago, CEPAL, 2019.
- _____ (2011a) *Guía para asegurar la calidad de los datos censales*, Serie Manuales N.º74, (LC/L.3431), Santiago, CEPAL, 2011.
- _____ (2011b). *Guía para la elaboración de un proyecto censal*, Serie Manuales N.º70, (LC/L.3324-P), Santiago, CEPAL, 2011.
- Chackiel, Juan (2010), "Evaluación post-empadronamiento de la cobertura en los censos de población". Vol.37 Notas de Población N.º91. (LC/G.2484-P), Santiago, Comisión Económica para América Latina y el Caribe (CEPAL), 2010.
- FLACSO-Chile, CELADE-CEPAL e INE-Chile (2018). *Experiencias, buenas prácticas y desafíos para los sistemas estadísticos nacionales de los países de América Latina frente a la ronda de censos de población y vivienda de 2020*. Santiago de Chile, Ediciones FLACSO-Chile.
- International Organization for Standardization (ISO) (2021), "ISO 9000:2015(es), (2015) Sistemas de gestión de la calidad. Fundamentos y vocabulario [en línea] <https://www.iso.org/obp/ui/es/#iso:std:iso:9000:ed-4:v1:es>.
- United Nations (2020), *Handbook on Population and Housing Census Editing Revision 2*, New York Statistics Division (UNSD).
- _____ (2019), *Guidelines on the use of electronic data collection technologies in population and housing censuses*, Nueva York, Statistics Division (UNSD).
- _____ (2017), *Principles and Recommendations for Population and Housing Censuses, Revision 3*, (ST/ESA/STAT/SER.M/67/Rev.3). New York, Department of Economic and Social Affairs, Statistics Division (UNSD). United Nations Publications.
- _____ (2001), *Handbook on Census Management for Population and Housing Censuses* (ST/ESA/STAT/SER.F/83/Rev.1), Nueva York, Department of Economic and Social Affairs, Statistics Division (UNSD). United Nations Publications.

Anexos

Anexo 1

Matriz cualitativa para censo de Perú 2017

Ejercicio hecho por el INEI de Perú en 2014, por Martha Rodríguez, en el cual se exponen criterios cualitativos y sus valoraciones, para evaluar la utilización de digitación, escaneo o DMC. En ese entonces la alternativa vía Internet (eCenso) no se consideró como un método principal posible, por lo que no se incluyó en la evaluación.

Cuadro A.1
Ejemplo de Matriz cualitativa para Censo de Perú 2017

Criterio	Digitación	Escaneo	DMC
Impresión cuestionarios	Se puede imprimir y encuadernar en la imprenta del INEI, no requiere tamaños fijos y se puede sacar copia	Se requiere un tamaño determinado de papel. Se requiere tamaño fijo de cada una de las cuadrículas donde se escribirán las respuestas en texto. Se requiere control de calidad de la impresión del cuestionario cada cierta cantidad de formularios (100,000)	No se requiere hacer gasto
Nivel de los censistas	Deben tener secundaria y ser bien capacitados por más de 15 horas	Deben tener secundaria y ser bien capacitados por más de 25 horas, de las cuales se dediquen a caligrafía por lo menos 10 horas	Deben tener secundaria completa o superior, ser bien capacitados en el uso y manejo de DMC por más de 45 horas, se debe tomar un examen para ser seleccionados
Esfuerzo de la captura de la información	Se requiere de mucha mano de obra en tres turnos, de un lugar determinado, de muchas PC, según se calcule cuánto tiempo se demorará	Se requiere tercerizar esta actividad, de tal manera que nosotros ejecutemos el control de calidad de la interpretación de los escáneres que usan	Todo depende de la capacitación y la buena vista que tenga la persona para usar el dispositivo móvil
Esfuerzo de codificación de texto abierto	Es preferible que se ejecute antes de enviarlo a Digitación	Si el texto está bien interpretado se puede pasar por el sistema de codificación automático, solo se requiere alimentar las tablas maestras	Se puede usar tablas maestras y también un campo otros para ser codificados posteriormente por el sistema de codificación
Flexibilidad en la entrevista	Se tiene a la vista el cuestionario completo, según su capacitación recogerá la información	Se tiene a la vista el cuestionario completo, según su capacitación recogerá la información, además debe tener una buena caligrafía	No tiene el cuestionario completo como para revisar alguna pregunta que ya respondió
Dificultades logísticas en terreno	Las dificultades logísticas son el traslado de ida y vuelta de los cuestionarios	Las dificultades logísticas son el traslado de ida y vuelta de los cuestionarios, además se requiere una adecuada manipulación para no ser deteriorado	Las dificultades logísticas, geográficas y relativas a la seguridad son enormes, porque son artefactos que si se golpean o se caen se puede malograr y están expuestos al robo
Controles de consistencia	No se aplica ninguna consistencia, pero se aplica el criterio de cómo entendió que debe hacer las preguntas del cuestionario	No se aplica ninguna consistencia, pero se aplica el criterio de cómo entendió que debe hacer las preguntas del cuestionario	Según la respuesta de la pregunta se ejecutará los saltos automáticamente y el dato será semi-consistente
Experiencias mundiales	No evaluado	No evaluado	No evaluado
Pilotaje (Caso óptimo de uso)	Se monitorea con un sistema la entrega de los formularios al área de digitación, para que todos los formularios se digiten y luego se devuelvan	Aquí el lote de escaneo depende de lo que soporte la bandeja del scanner, luego el reconocimiento de marcas es óptimo (90%), el problema está en el reconocimiento de valores escritos de letras y números	Solo se debe tener buena vista para dar marcar la respuesta que nos informan

Criterio	Digitación	Escaneo	DMC
Costo	Se genera labor para muchas personas y el costo es barato	Es muy costoso y se refleja en la impresión del formulario, la capacitación de las personas, en el traslado de ida y vuelta y en el escaneo de los formularios que es tercerizado	Es muy caro, depende del tipo de censo que se va a ejecutar y de la cantidad de artefactos móviles que se van a comprar, si es de hecho sería uno por cada censista, sería menor el costo si es un censo de derecho
Plazos	Demoraría por lo menos unos 3 a 4 meses, en 3 o 4 turnos	Esto depende del control de calidad en aceptar la interpretación de los datos, pero se demoraría menos que la digitación	Con artefacto móvil se tendría el dato al finalizar el día del censo
Nivel de calidad	Por más que se coloque la verificación de lotes es calificada como mala	Aquí se salvan los resultados por marcas que son reconocidos en un margen de 90 por ciento, es calificado como intermedio	La calidad depende de la capacitación y de los buenos ojos del empadronador

Fuente: Elaborado por Martha Rodríguez, profesional del INEI en el proceso de evaluación del método de captura llevado a cabo por Perú en 2014.

Anexo 2

Fichas de Sistematización de Experiencias para Países de América Latina y el Caribe, del GTC CEA–CEPAL para el documento Lineamientos generales para la captura de datos censales

A continuación, se presentan las fichas de sistematización que los siete países convocados autorizaron para su incorporación en este documento con el fin de compartir sus experiencias recientes en censos con escáner, eCenso y DMC. Cabe mencionar que estas fichas fueron respondidas por los países en agosto de 2020, por lo que algunos aspectos pueden haber cambiado a la luz del desarrollo de sus proyectos censales.

- A.1 Brasil, Censo de 2010, uso de DMC
- A.2 Brasil, censo 2010 eCenso
- A.3 Chile, censo 2017, escáner
- A.4 Colombia Censo 2018, DMC
- A.5 Costa Rica, censo 11, DMC
- A.6 Ecuador, censo 2010, escáner
- A.7 México, censo 2020, DMC
- A.8 México, censo 2020, eCenso
- A.9 República Bolivariana de Venezuela, censo 2011, DMC

Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe
Grupo de Trabajo Censos de la Conferencia Estadística de las Américas – CEPAL
Documento Lineamientos generales para la captura de datos censales

En el marco del plan de trabajo del **Grupo de censos de la CEA CEPAL** para el **bienio 2020-2021** se comprometió la realización de un documento que abordara lineamientos generales para la captura de datos censales para los países de América Latina y el Caribe. Uno de los objetivos específicos es la recopilación de experiencias por parte de los países respecto de la utilización de algunos de los métodos de recolección y captura de datos.

La presente ficha tiene por objetivo sistematizar la experiencia nacional respecto del método de captura de datos censales que ha sido solicitada a usted como representante de su país por parte del equipo de coordinación del Grupo de Trabajo de Censos de CEA Cepal.

La información aquí recopilada será utilizada como insumo para la redacción del documento, durante febrero del 2021 le enviaremos la versión preliminar del documento que será generado con su aporte para que sea previamente aprobada por la dirección de la oficina nacional de estadísticas de su país.

La ficha que se presenta consta de **5 partes**: Información de contacto; Características generales del último censo realizado en su país; Información del Método de Captura; Síntesis de aprendizajes; y Planes para el próximo censo.

Le pedimos enviar la ficha completa a más tardar el día **16 de octubre del 2020**. Su contenido será revisado por el equipo de coordinación del Grupo de Trabajo de Censos y de ser necesario le solicitará una reunión para profundizar en alguno de sus aspectos.

Agradecemos desde ya su colaboración en este proceso y ante cualquier duda o consulta por favor contactarse con Olga Barquero del INE-Chile: olga.barquero@ine.cl y Alejandra Silva de CELADE: alejandra.silva@cepal.org

1. Información de Contacto.

Por favor llenar con su información de contacto como responsable del llenado de la ficha por parte de su país.

Nombre	LUCIANO TAVARES DUARTE
Correo electrónico	luciano.duarte@ibge.gov.br
Cargo	COORDINADOR TECNICO DEL CENSO
Institución	IBGE
País	BRASIL

2. Características generales del **último censo** realizado en el país.

2.1. Marque con una X el Tipo de Censo para el último censo realizado en su país:

De Hecho De Derecho

2.2. Indique la fecha de inicio de la recolección del último censo en su país:

Día Mes Año

2.3. Indique la duración de la recolección del último censo en su país en semanas, se debe incluir el período oficial además de las semanas que se extendió la recolección más allá de ese período hasta el cierre del operativo censal.

a. Duración del período oficial: semanas. Si su censo fue de hecho anotar 0.

En caso de que quiera especificar condiciones especiales para el período oficial de recolección, como, por ejemplo, diferencia de duración entre urbano y rural, especifique en el cuadro siguiente:

--

b. ¿Cuántas semanas se extendió la recolección después del período oficial?: semanas.

2.4. Indique los métodos de recolección o captura utilizados en el último censo en su país con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas cubiertas mediante ese método.

	% de Viviendas	% de Población
<input type="checkbox"/> Papel/Digitación		
<input type="checkbox"/> Papel/Reconocimiento Óptico		
<input checked="" type="checkbox"/> DMC / CAPI	99,9%	99,9%
<input checked="" type="checkbox"/> Web / CAWI	0,1%	0,1%

2.5. Describa de forma general la estrategia operativa del censo. Incluya si los métodos de recolección o captura se aplicaron de forma secuencial o en paralelo; qué métodos de recolección o captura fueron utilizados para operativos especiales (viviendas colectivas, situación de calle, población en tránsito u otro); indicar si la captura de datos fue centralizada o descentralizada, o cualquier otra información relevante para el contexto de su experiencia.

La estrategia de recolección de datos fue por entrevistas presenciales con uso del DMC de forma descentralizada. La internet fue utilizada paralelamente como una alternativa solamente en los casos en que los informantes hacían la solicitud. No se censó toda la población en situación de calle, solamente la parte que vivía en habitaciones improvisadas o viviendas colectivas. En las viviendas colectivas, la entrevista fue presencial siempre que posible, pero en algunos casos especiales (como prisiones) se hizo por medio de los registros administrativos.

3. Características del método: **Dispositivo Móvil de Captura.**

3.2. Si usted tuviera que defender el uso de dispositivos móviles de captura frente a sus pares de otros países, ¿podría sintetizar el proceso de decisión para la elección de este método como parte del proceso de recolección censal?

Sin duda el uso de los DMC tiene muchas ventajas. Es posible, por ejemplo, diseñar preguntas diferentes para obtener la “misma información” de personas con perfiles diferentes; se puede hacer crítica de entrada de datos y consistencia en el momento de la entrevista; es posible crear reglas de salto y flujo; se puede hacer codificación en campo de muchas variables (variables de migración, educación, religión y otras); se puede poner una función ayuda en cada pregunta para sacar posibles dudas durante la realización de la entrevista; y muchos otros. La gran cuestión es el costo financiero y de desarrollo de los algoritmos. A pesar de los DMC permitirán una gran variedad de posibilidades, hay un gran costo humano y tecnológico que precisa ser llevado en cuenta antes de la decisión.

3.3. Indique las pruebas realizadas durante la preparación de la recolección con dispositivos móviles de captura, incluir todas las pruebas tecnológicas, de cuestionario, operativas y censo experimental, donde se haya evaluado este método. Identificar los principales hallazgos en esas pruebas.

Para el censo de 2010 tuvimos tres pruebas piloto y un censo experimental. Después del censo experimental, aún tuvimos que hacer otro operativo de campo para probar nuevos equipos, pues los "Netbooks" que intentamos usar como DMC para los empadronadores no se mostraron una buena opción en el censo experimental. Tuvimos también muchos tests internos de homologación de los sistemas (tanto DMC como web).

3.4. Sintetice el desarrollo TI de la recolección con dispositivos móviles de captura. Indique si el software de recolección fue desarrollado in house o se utilizó un software externo. Se solicita describir procesos de respaldo, transferencia de datos, elementos de seguridad y protección de datos personales (criptaciones, manejo de claves y accesos, etc.); Incluir consideraciones para la adquisición (licitaciones de compra, arriendo o provisión de servicios).

El software utilizado para la recolección electrónica de datos en dispositivos móviles fue desarrollado internamente por IBGE, utilizando el lenguaje Java y una plataforma llamada TotalCross, compuesta por bibliotecas y una máquina virtual para el sistema operativo Windows Mobile, que era el sistema utilizado en los DMC. Los datos se almacenaron internamente en el dispositivo de forma encriptada y en archivos independientes para cada cuestionario. Cada dispositivo tenía una tarjeta SD de backup. La información se transmitió desde el dispositivo a la computadora portátil local en el “puesto de recolección”, a través de un cable USB. Desde este portátil, los datos se transmitían a través de un canal seguro al servidor central de IBGE. Más de 200.000 dispositivos móviles de recolección fueron adquiridos para uso de empadronadores y supervisores en 2010. Además del equipo interno del IBGE que desarrolló la aplicación de recolección, IBGE contrató un soporte de la empresa responsable de TotalCross durante el período de desarrollo y recolección para ayudar al desarrollo y mantenimiento de la aplicación.

3.5. Identifique elementos de logística operativa asociados a los dispositivos móviles de captura, por ejemplo, distribución de material para la recolección, retorno de material de recolección, manejo de contingencias en caso de pérdidas, destrucción o fallas en los materiales y cualquier otro elemento relevante en este sentido.

De hecho, hay una gran carga de trabajo en los preparativos y distribución de los equipos, pero este tiempo de preparación sin duda es compensado por no tener que hacer transcripción de los datos (mismo que sea automática) y también por se ganar mucho tiempo en los procesos de crítica y codificación. Actualmente la estructura de TI nos permite que los datos colectados sean transmitidos directamente de los DMC para el ambiente central vía internet, o que nos permite aun un acompañamiento y control muy preciso de la operación.

3.6. Describa la forma de monitoreo de la recolección con dispositivos móviles de captura. Indique como se realizó el monitoreo y control de la transferencia de datos; la forma de control y monitoreo del avance de la recolección en términos de cobertura territorial y poblacional. Qué validaciones se realizaron para evitar duplicados (incluir algunos criterios de identificación), pérdida de datos durante la captura, o datos erróneos por mal funcionamiento del método.

Em 2010 los sistemas de controle eran descentralizados. Toda la vinculación del equipo de campo con los DMC y recepción de los datos ocurrían en los "Puestos de Recolección", con uso de Notebooks que contenían un sistema llamado SIGPC (Sistema de Gerenciamiento del puesto de recolección). Los DMC no se comunicaban con el ambiente central, solamente el SIGPC lo hacía. Cuando se hacía la comunicación del SIGPC con el ambiente central, toda la estructura del IBGE tenía acceso a los datos para hacer el acompañamiento (generalmente se hacía esa comunicación una vez al día). Así, todo el control del trabajo dependía de la conexión de los DMC con el SIGPC. Los DMC contaban también con un cartón de memoria para backup, que podría ser descargado directamente en el SIGPC en caso de problemas con el equipo. Toda la validación de los datos se daba en el ambiente central, donde había criterios preestablecidos para cada caso. Las pérdidas de datos o datos erróneos fueron muy insignificantes.

4. Síntesis del aprendizaje sobre el método de captura.

4.1. Si usted tuviera la oportunidad de entregar tres recomendaciones sobre este método de captura a otro país, ¿qué elementos a partir de su experiencia deberían ser críticos para que el método funcione de forma adecuada?

Recomendación 1:

El primero punto es sin duda haber una fuerte integración entre el área técnica y la de informática. Cuando se usa DMC, todas las posibilidades que estos dispositivos proporcionan dependen del desarrollo de algoritmos y presentan una gran demanda en términos de recursos humanos.

Recomendación 2:

Asegurar los presupuestos necesarios para adquirir los equipos.

Recomendación 3:

Tener una estrategia de uso posterior, incluso con la posibilidad de usarlos en otros órganos del gobierno.

5. Planes para el próximo censo de población y vivienda.

5.1. Marque con una X el Tipo de Censo que espera realizar para la próxima ronda censal en su país:

De Hecho De Derecho

5.2. Indique la duración de la recolección censal que espera para su próximo censo en el país, en semanas.

semanas. Si su censo será de hecho anotar 0.

5.3. Indique los métodos de recolección o captura que espera utilizar en su país para el próximo censo con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas que espera cubrir con cada método.

		% de Viviendas	% de Población
<input type="checkbox"/>	Papel/Digitación		
<input type="checkbox"/>	Papel/Reconocimiento Óptico		
<input checked="" type="checkbox"/>	DMC / CAPI	97%	97%
<input checked="" type="checkbox"/>	Web / CAWI	3%	3%

5.4. Identificar con una X todos los tipos de cuestionario censal que va a considerar en el próximo censo en su país.

- Viviendas particulares
- Viviendas Colectivas
- Personas en Situación de Calle
- Personas en Tránsito o Viajeros
- Otros

Si marca "otros", especificar en este recuadro:

5.5. Describa de forma breve los métodos a utilizar en el próximo censo, explicité de forma general las razones de mantener o cambiar los métodos utilizados. Además, incorpore los principales desafíos y avances en la preparación del censo en esta materia.

La estrategia de recolección de datos va a ser mantenida: entrevistas presenciales con uso del DMC de forma descentralizada. La internet va a ser utilizada nuevamente de forma auxiliar y paralelamente con la operación presencial. Intentamos proponer mudanzas especialmente en la estrategia de internet, pero preferimos mantener la estrategia adoptada en 2010 por seguridad. Además como forma alternativa para el llenado de los cuestionarios, los censistas podrán utilizar sus DMCs para hacer las entrevistas por teléfono. Esta modalidad será utilizada cuando el censista no encontrar ningún morador presente en la vivienda. En estos casos podrá dejar un billete para que el morador haga una llamada para el censista e se realice el llenado por llamada telefónica.

5.6. Si tuviera la oportunidad de realizar tres preguntas a otros países de América Latina y el Caribe que hayan implementado el método captura, o los métodos de captura, que usted proyecta para el próximo censo, ¿cuáles serían?

Pregunta 1:

? Cómo se hizo la formación de los empadronadores? Por clases presenciales o remotas?

Pregunta 2:

Pregunta 3:

Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe
Grupo de Trabajo Censos de la Conferencia Estadística de las Américas – CEPAL
Documento Lineamientos generales para la captura de datos censales

En el marco del plan de trabajo del **Grupo de censos de la CEA CEPAL** para el **bienio 2020-2021** se comprometió la realización de un documento que abordara lineamientos generales para la captura de datos censales para los países de América Latina y el Caribe. Uno de los objetivos específicos es la recopilación de experiencias por parte de los países respecto de la utilización de algunos de los métodos de recolección y captura de datos.

La presente ficha tiene por objetivo sistematizar la experiencia nacional respecto del método de captura de datos censales que ha sido solicitada a usted como representante de su país por parte del equipo de coordinación del Grupo de Trabajo de Censos de CEA Cepal.

La información aquí recopilada será utilizada como insumo para la redacción del documento, durante febrero del 2021 le enviaremos la versión preliminar del documento que será generado con su aporte para que sea previamente aprobada por la dirección de la oficina nacional de estadísticas de su país.

La ficha que se presenta consta de **5 partes**: Información de contacto; Características generales del último censo realizado en su país; Información del Método de Captura; Síntesis de aprendizajes; y Planes para el próximo censo.

Le pedimos enviar la ficha completa a más tardar el día **16 de octubre del 2020**. Su contenido será revisado por el equipo de coordinación del Grupo de Trabajo de Censos y de ser necesario le solicitará una reunión para profundizar en alguno de sus aspectos.

Agradecemos desde ya su colaboración en este proceso y ante cualquier duda o consulta por favor contactarse con Olga Barquero del INE-Chile: olga.barquero@ine.cl y Alejandra Silva de CELADE: alejandra.silva@cepal.org

1. Información de Contacto.

Por favor llenar con su información de contacto como responsable del llenado de la ficha por parte de su país.

Nombre	LUCIANO TAVARES DUARTE
Correo electrónico	luciano.duarte@ibge.gov.br
Cargo	COORDINADOR TECNICO DEL CENSO
Institución	IBGE
País	BRASIL

2. Características generales del **último censo** realizado en el país.

2.1. Marque con una X el Tipo de Censo para el último censo realizado en su país:

De Hecho De Derecho

2.2. Indique la fecha de inicio de la recolección del último censo en su país:

Día Mes Año

2.3. Indique la duración de la recolección del último censo en su país en semanas, se debe incluir el período oficial además de las semanas que se extendió la recolección más allá de ese período hasta el cierre del operativo censal.

a. Duración del período oficial: semanas. Si su censo fue de hecho anotar 0.

En caso de que quiera especificar condiciones especiales para el período oficial de recolección, como, por ejemplo, diferencia de duración entre urbano y rural, especifique en el cuadro siguiente:

--

b. ¿Cuántas semanas se extendió la recolección después del período oficial?: semanas.

2.4. Indique los métodos de recolección o captura utilizados en el último censo en su país con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas cubiertas mediante ese método.

	% de Viviendas	% de Población
<input type="checkbox"/> Papel/Digitación		
<input type="checkbox"/> Papel/Reconocimiento Óptico		
<input checked="" type="checkbox"/> DMC / CAPI	99,9%	99,9%
<input checked="" type="checkbox"/> Web / CAWI	0,1%	0,1%

2.5. Describa de forma general la estrategia operativa del censo. Incluya si los métodos de recolección o captura se aplicaron de forma secuencial o en paralelo; qué métodos de recolección o captura fueron utilizados para operativos especiales (viviendas colectivas, situación de calle, población en tránsito u otro); indicar si la captura de datos fue centralizada o descentralizada, o cualquier otra información relevante para el contexto de su experiencia.

La estrategia de recolección de datos fue por entrevistas presenciales con uso del DMC de forma descentralizada. La internet fue utilizada paralelamente como una alternativa solamente en los casos en que los informantes hacían la solicitud. No se censó toda la población en situación de calle, solamente la parte que vivía en habitaciones improvisadas o viviendas colectivas. En las viviendas colectivas, la entrevista fue presencial siempre que posible, pero en algunos casos especiales (como prisiones) se hizo por medio de los registros administrativos.

3. Características del método: **Autoempadronamiento Web.**

3.2. Si usted tuviera que defender el uso de autoempadronamiento web frente a sus pares de otros países, ¿podría sintetizar el proceso de decisión para la elección de este método como parte del proceso de recolección censal?

Dependiendo de la estrategia, lo más importante a llevarse en cuenta es la capacidad de vincular los cuestionarios a las direcciones correctas y garantizar la cobertura. Este es uno de los grandes desafíos que tenemos en Brasil para ampliar el uso de la internet, sin la necesidad del empadronador hacer este vínculo en campo. Como un método auxiliar, todavía, es una estrategia muy bien venida, especialmente por cuenta de los altos porcentuales de habitaciones con moradores ausentes que hemos encontrado en los testes de campo (mucho más grandes de que en 2010). En estos casos, o en otros también (como rechaza, o cuando la persona no tiene tiempo en aquello momento para hacer la entrevista), la internet puede ayudar mucho.

3.3. *Indique las pruebas realizadas durante la preparación de la recolección mediante autoempadronamiento web, incluir todas las pruebas tecnológicas, de cuestionario, operativas y censo experimental, donde se haya evaluado este método. Identificar los principales hallazgos en esas pruebas.*

La evaluación del cuestionario web se dio siguiendo la misma lógica del DMC. Además de las tres pruebas piloto y del censo experimental, se hizo también dos testes internos con la participación de todos los funcionarios del IBGE (alrededor de 5000).

3.4. Sintetice el desarrollo TI de la recolección con autoempadronamiento web. Indique si el software de recolección fue desarrollado in house o se utilizó un software externo. Se solicita describir procesos de respaldo, transferencia de datos, elementos de seguridad y protección de datos personales (encriptaciones, manejo de claves y accesos, etc.); Incluir consideraciones para la adquisición (licitaciones de compra, arriendo o provisión de servicios).

El sistema web para el cuestionario del censo de 2010 fue desarrollado internamente por el IBGE. Se utilizó una plataforma basada en el servidor Apache / HTTPS y PHP. Para la base de datos, se utilizó Oracle en una arquitectura de clúster (Oracle RAC). También se empleó una arquitectura de red con firewall F5 y equilibrio de carga. Los servidores web de la aplicación tenían su propia rutina de copia de seguridad interna y la base de datos de Oracle también se configuró para realizar copias de seguridad periódicas utilizando las propias herramientas de Oracle. Se firmó un contrato con una empresa especializada en soluciones Oracle para soporte y mantenimiento durante el período de desarrollo y recolección de datos.

3.5. Identifique elementos de logística operativa asociados al autoempadronamiento web, por ejemplo, distribución de claves de acceso a la población, manejo de sesiones y soporte de ayuda a usuarios y resolución de problemas y cualquier otro elemento relevante en este sentido.

Lo más importante a llevarse en cuenta, como mencionado arriba, es la capacidad de vincular los cuestionarios a las direcciones correctas y garantizar la cobertura. Así, la estrategia adoptada en Brasil en 2010 fue de entregar al informante una carta con un código de acceso único y exclusivo, vinculado en el momento de la entrega por el empadronador en el DMC. Para 2021 vamos a repetir este mismo proceso, pero la clave de acceso va a llegar por correo electrónico. Solamente en los casos en que no se encuentren los moradores después de muchas tentativas (viviendas ocupadas con moradores ausentes), el supervisor va a dejar una carta con un código de acceso para aquella vivienda.

3.6. Describa la forma de monitoreo de la recolección con autoempadronamiento web. Indique como se realizó el monitoreo y control de la transferencia de datos; la forma de control y monitoreo del avance de la recolección en términos de cobertura territorial y poblacional. Qué validaciones se realizaron para evitar duplicados (incluir algunos criterios de identificación), pérdida de datos durante la captura, o datos erróneos por mal funcionamiento del método.

El control de la recolección con autoempadronamiento web siguió la misma lógica del DMC. Cuando un informante llenaba el cuestionario, sus informaciones iban para un banco de datos intermediario hasta que la clave entre el código del cuestionario y la dirección fueran transmitidos del DMC para el SIGPC, y después del SIGPC para el sistema central. Así, cuando se hacía la comunicación del SIGPC con el ambiente central, toda la estructura del IBGE tenía acceso a los datos de colecta vía DMC y web para hacer el acompañamiento. Toda la evaluación y control de integridad en el caso de los cuestionarios web se daba en el ambiente central.

4. Síntesis del aprendizaje sobre el método de captura.

4.1. Si usted tuviera la oportunidad de entregar tres recomendaciones sobre este método de captura a otro país, ¿qué elementos a partir de su experiencia deberían ser críticos para que el método funcione de forma adecuada?

Recomendación 1:

Si la intención es tener una participación masiva de la población, hay que invertir mucho en publicidad.

Recomendación 2:

Las personas que van a llenar el cuestionario web no van a tener un entrenamiento para entender todas las preguntas como los empadronadores, entonces es fundamental que se cree un conjunto de herramientas para auxiliar los informantes en caso de duda, así como invertir en una central de apoyo telefónico.

Recomendación 3:

No se debe cambiar el método de recolección únicamente por cuestiones presupuestales. La mayoría de los países que cambiaron de papel o DMC para web no lo hicieron por eso. Debemos garantizar la calidad y la cobertura en primer lugar.

5. Planes para el próximo censo de población y vivienda.

5.1. Marque con una X el Tipo de Censo que espera realizar para la próxima ronda censal en su país:

De Hecho De Derecho

5.2. Indique la duración de la recolección censal que espera para su próximo censo en el país, en semanas.

semanas. Si su censo será de hecho anotar 0.

5.3. Indique los métodos de recolección o captura que espera utilizar en su país para el próximo censo con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas que espera cubrir con cada método.

		% de Viviendas	% de Población
<input type="checkbox"/>	Papel/Digitación		
<input type="checkbox"/>	Papel/Reconocimiento Óptico		
<input checked="" type="checkbox"/>	DMC / CAPI	97%	97%
<input checked="" type="checkbox"/>	Web / CAWI	3%	3%

5.4. Identificar con una X todos los tipos de cuestionario censal que va a considerar en el próximo censo en su país.

- Viviendas particulares
- Viviendas Colectivas
- Personas en Situación de Calle
- Personas en Tránsito o Viajeros
- Otros

Si marca "otros", especificar en este recuadro:

5.5. Describa de forma breve los métodos a utilizar en el próximo censo, explicita de forma general las razones de mantener o cambiar los métodos utilizados. Además, incorpore los principales desafíos y avances en la preparación del censo en esta materia.

La estrategia de recolección de datos va a ser mantenida: entrevistas presenciales con uso del DMC de forma descentralizada. La internet va a ser utilizada nuevamente de forma auxiliar y paralelamente con la operación presencial. Intentamos proponer mudanzas especialmente en la estrategia de internet, pero preferimos mantener la estrategia adoptada en 2010 por seguridad.

5.6. Si tuviera la oportunidad de realizar tres preguntas a otros países de América Latina y el Caribe que hayan implementado el método captura, o los métodos de captura, que usted proyecta para el próximo censo, ¿cuáles serían?

Pregunta 1:

? Cómo se hizo la formación de los empadronadores? Por clases presenciales o remotas?

Pregunta 2:

Pregunta 3:

Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe
Grupo de Trabajo Censos de la Conferencia Estadística de las Américas – CEPAL
Documento Lineamientos generales para la captura de datos censales

En el marco del plan de trabajo del **Grupo de censos de la CEA CEPAL** para el **bienio 2020-2021** se comprometió la realización de un documento que abordara lineamientos generales para la captura de datos censales para los países de América Latina y el Caribe. Uno de los objetivos específicos es la recopilación de experiencias por parte de los países respecto de la utilización de algunos de los métodos de recolección y captura de datos.

La presente ficha tiene por objetivo sistematizar la experiencia nacional respecto del método de captura de datos censales que ha sido solicitada a usted como representante de su país por parte del equipo de coordinación del Grupo de Trabajo de Censos de CEA Cepal.

La información aquí recopilada será utilizada como insumo para la redacción del documento, durante febrero del 2021 le enviaremos la versión preliminar del documento que será generado con su aporte para que sea previamente aprobada por la dirección de la oficina nacional de estadísticas de su país.

La ficha que se presenta consta de **5 partes**: Información de contacto; Características generales del último censo realizado en su país; Información del Método de Captura; Síntesis de aprendizajes; y Planes para el próximo censo.

Le pedimos enviar la ficha completa a más tardar el día **16 de octubre del 2020**. Su contenido será revisado por el equipo de coordinación del Grupo de Trabajo de Censos y de ser necesario le solicitará una reunión para profundizar en alguno de sus aspectos.

Agradecemos desde ya su colaboración en este proceso y ante cualquier duda o consulta por favor contactarse con Olga Barquero del INE-Chile: olga.barquero@ine.cl y Alejandra Silva de CELADE: alejandra.silva@cepal.org

1. Información de Contacto.

Por favor llenar con su información de contacto como responsable del llenado de la ficha por parte de su país.

Nombre	Daniel Allende
Correo electrónico	daniel.allende@ine.cl
Cargo	Jefe Área Técnica – Censo de Población y Vivienda
Institución	Instituto Nacional de Estadísticas
País	Chile

2. Características generales del **último censo** realizado en el país.

2.1. Marque con una X el Tipo de Censo para el último censo realizado en su país:

De Hecho De Derecho

2.2. Indique la fecha de inicio de la recolección del último censo en su país:

Día Mes Año

2.3. Indique la duración de la recolección del último censo en su país en semanas, se debe incluir el período oficial además de las semanas que se extendió la recolección más allá de ese período hasta el cierre del operativo censal.

a. Duración del período oficial: semanas. Si su censo fue de hecho anotar 0.

En caso de que quiera especificar condiciones especiales para el período oficial de recolección, como, por ejemplo, diferencia de duración entre urbano y rural, especifique en el cuadro siguiente:

Sin comentarios

b. ¿Cuántas semanas se extendió la recolección después del período oficial?: semanas.

2.4. Indique los métodos de recolección o captura utilizados en el último censo en su país con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas cubiertas mediante ese método.

		% de Viviendas	% de Población
<input type="checkbox"/>	Papel/Digitación	99,995	99,9955
<input checked="" type="checkbox"/>	Papel/Reconocimiento Óptico		
<input type="checkbox"/>	DMC / CAPI		
<input checked="" type="checkbox"/>	Web / CAWI	0,005	0,0045

2.5. Describa de forma general la estrategia operativa del censo. Incluya si los métodos de recolección o captura se aplicaron de forma secuencial o en paralelo; qué métodos de recolección o captura fueron utilizados para operativos especiales (viviendas colectivas, situación de calle, población en tránsito u otro); indicar si la captura de datos fue centralizada o descentralizada, o cualquier otra información relevante para el contexto de su experiencia.

El Censo 2017 fue un operativo de emergencia que se realizó con una metodología de hecho, para dicho efecto se utilizó personal voluntario que fue convocado con funcionarios públicos, estudiantes secundarios y universitarios además de organizaciones de la sociedad civil y voluntarios reclutados vía web.

El operativo se dividió en cuatro: viviendas particulares; personas en viviendas colectivas; personas en tránsito y; personas en situación de calle. Para la recolección se utilizaron tres cuestionarios distintos, el cuestionario de viviendas particulares que correspondía a un cuadernillo de 10 páginas, la primera correspondía a los datos de identificación geográfica y los datos de vivienda; la segunda página correspondía a una sección de registro de personas en el hogar; las páginas 3 a 10 eran idénticas y correspondían al llenado de las preguntas para cada persona del hogar con un máximo de 8 personas por cuestionario. Esta decisión fue tomada dado que con 8 personas era posible cubrir el 99,8% de los hogares sin aplicar un cuestionario de continuación que es más complejo de manejar durante la lectura óptica y el enlace de viviendas-hogares y personas.

Se trabajó el proceso de impresión, distribución y digitalización de datos a través de un contrato que en este caso se llevó a cabo con tres empresas especializadas en su área, pero bajo el mismo contrato público. De esta manera quedaban vinculados los procesos que impactan en toda la cadena del reconocimiento, desde el diseño gráfico del cuestionario, el tipo de hoja, los tamaños de los cuadros para letras, números y marcas además de los códigos asociados al tracking y los controles de completitud de cuestionario durante el procesamiento.

Los cuestionarios fueron impresos y se mecanizó un paquete completo que correspondía al material de un sector de empadronamiento, de esta manera venía el cuestionario, el tipo de lápiz y goma adecuada, stickers para el código geográfico (aumentaba la calidad de la lectura óptica para vinculación geográfica); y cuestionarios extra en caso de encontrar más personas de las esperadas en el sector. Además, se mecanizaron carpetas de contingencia y material granel que se distribuyó a lo largo de los locales censales para abordar situaciones de contingencia en terreno.

Los cuestionarios eran revisados en los locales y organizados para ser enviados en las mismas cajas que fueron enviados para su retorno a nivel central y ser digitalizados de forma central en una bodega de la capital.

3. Características del método: Reconocimiento Óptico de Caracteres.

3.2. Si usted tuviera que defender el uso de Reconocimiento Óptico de Caracteres frente a sus pares de otros países, ¿podría sintetizar el proceso de decisión para la elección de este método como parte del proceso de recolección censal?

El reconocimiento óptico de caracteres es un método utilizado por los Censos en Chile desde el año 1982, en este sentido, se consideraba que había una experiencia acumulada en la institución para poder realizar de forma adecuada este tipo de método.

Luego que el Proyecto Censal del año 2012 fuera declarado inválido, la principal recomendación de los expertos nacionales e internacionales fue realizar un Censo de emergencia con carácter abreviado utilizando la forma tradicional de recolección de los censos anteriores en el país, esto se tradujo en programar un censo de hecho con cuestionarios en papel y reconocimiento óptico.

Si bien esta decisión consideró esto, es importante destacar que el proceso del año 2012 también produjo un cambio importante en los equipos que históricamente habían trabajado los censos de población por lo que fue necesario realizar una mayor cantidad de pruebas asociadas al proceso para construir ese aprendizaje por parte de los equipos técnicos del INE.

Una vez que el INE ha definido esto se abordan algunas ventajas que ofrecen los métodos de reconocimiento actual:

- a. Escaneo de imágenes a color: es necesario considerar que el reconocimiento opera sobre una imagen de contraste en blanco y negro que elimina la información, esto hacía necesario mantener en bodega por mucho tiempo los cuestionarios físicos a su vez que impedía revisar los posibles problemas del contraste con el cuestionario original de forma rápida, hoy los métodos de almacenamiento permiten guardar la imagen a color de manera que sea un fiel reflejo del cuestionario en papel y mantener un archivo digital de consulta más eficiente ante posibles errores.
- b. Desarrollo de validaciones: la organización de la carga de trabajo del escaner es clave para mantener la calidad el proceso en ese sentido los paquetes de escaneo se mantuvieron organizados por sector de empadronamiento, lo que facilitaba la consecución de validaciones geográficas y de consistencia asociadas a perfeccionar la construcción de una base relacional. Los softwares actualmente permiten este desarrollo lo que aumenta la calidad del proceso.
- c. El escáner **NO** soluciona la construcción de una base: El proceso de reconocimiento si bien es un avance en cuanto a estandarizar un proceso de digitalización del papel, no soluciona el problema de la construcción de una base relacional validada, en este sentido es necesario establecer criterios de control de calidad que aseguren los supuestos básicos (contabilización de persona y reconstrucción de viviendas y hogares) con los que funcionaran los algoritmos una vez recibida la base desde la empresa.

3.3. Indique las pruebas realizadas durante la preparación de la recolección con dispositivos móviles de captura, incluir todas las pruebas tecnológicas, de cuestionario, operativas y censo experimental, donde se haya evaluado este método. Identificar los principales hallazgos en esas pruebas.

Prueba Integral Censo: Esta prueba se realizó en abril del 2016 exactamente un año antes del censo de población y vivienda. El principal objetivo era probar la carga de trabajo de supervisores y censistas bajo los supuestos que se habían establecido la estrategia operativa del censo.

Como prueba extra para el equipo técnico se consideró el contrato con una empresa de reconocimiento óptico que no era la misma que se utilizó finalmente, pero que permitía sostener un aprendizaje sobre el reconocimiento óptico, a dicha empresa se le solicitó la base bruta sin validaciones, ni corrección del reconocimiento, de esta forma se observaron las fuentes de error asociadas al proceso de digitalización y establecer los criterios para abordarlos durante el proceso en conjunto con la empresa que ganó finalmente la licitación.

Ensayo Operativo Censal: Correspondiente al Censo Experimental realizado en noviembre del año 2016, se consideró relevante que la empresa que ganó la licitación implementara el proceso de reconocimiento, esto originalmente no estaba en el contrato y fue negociado a posteriori, esta experiencia permitió trabajar en conjunto con los equipos de la empresa en establecer el flujo de producción de datos y probar el mecanismo de validaciones que estaba propuesto desde el INE. Acá el principal aprendizaje fue eliminar la codificación de campos abiertos de este proceso debido a que no permitía manejar criterios estandarizados para los casos difíciles, por lo que el proceso se llevó adelante en las oficinas centrales del INE.

El segundo aprendizaje es que el proceso de reconocimiento tiene características fabriles que son ajenas muchas veces a los equipos técnicos, por lo que la aproximación al establecimiento de una cadena de producción con medición diaria de productividad y la presión del tiempo de contrato fue clave para proyectar el trabajo posterior en el Censo.

3.4. Sintetice el desarrollo del método de reconocimiento óptico. Indique si el proceso se realizó in house o a través de un tercero; se solicita describir proceso de respaldos de imágenes, transferencia de datos, elementos de seguridad y protección de datos personales (encriptaciones, manejo de claves y accesos, etc.); Incluir consideraciones para la adquisición (licitaciones de compra, arriendo o provisión de servicios). Debe indicar qué softwares se utilizaron para el reconocimiento óptico y gestión del proceso de digitalización.

El Instituto Nacional de Estadísticas, en el marco de la realización del Censo de Población y Vivienda 2017, licita la adquisición de material censal y los servicios de impresión, mecanizado, distribución, procesamiento de datos y bodegaje. Dicha licitación fue adjudicada por un proveedor externo formado por una unión temporal de las empresas Xerox de Chile SA e Imprenta Alfredo Molina Flores S.A. (AMF/XEROX).

Para el procesamiento de datos fue necesario contar con un sistema que fuese capaz de capturar datos de los cuestionarios de censo utilizando la digitalización y el reconocimiento inteligente de caracteres escritos a mano (ICR), marcas (OMR) y caracteres impresos (OCR), así como otras técnicas de procesamiento de formularios, codificación y transferencia de los datos a la Base de Datos del Censo, llevó a utilizar TIS (Top Imaging Systems) como la plataforma de procesamiento del proyecto y Eflow como el software de captura de datos contratado por Xerox.

Respecto al proceso de seguridad y protección de los datos, se siguieron estrictos protocolos de seguridad física y lógica, exigiendo al proveedor, por ejemplo, el envío encriptado de los datos, los que solo personal autorizado del INE podía descifrar.

Para el almacenamiento digital se definió un protocolo de traspaso de información desde la empresa hasta el INE. En este caso, el proveedor dispuso un servidor en el centro de datos principal de la institución. Lo anterior facilitó el proceso de traspaso de información, que se hacía con un disco duro externo proveniente del servidor interno de la empresa que se instalaba en el servidor del centro de datos, llamado también “servidor de paso”. Esta actividad fue monitoreada por personal técnico de la Departamento de Tecnologías de la Información (TI) INE. Una vez copiados los archivos en el “servidor de paso”, funcionarios de los equipos Censo revisaban la entrega de acuerdo con la documentación previamente definida, y así se verificaba que estuviera completa y que lo informado correspondiera con lo efectivamente copiado en el servidor.

Una vez verificada la completitud de las cantidades de datos e imágenes, personal del INE aprobaba la entrega y se generaba la copia desde el “servidor de paso” hacia servidores internos del INE. Como medida de precaución, tanto el INE como la empresa realizaron copias de respaldo de cada entrega bajo estrictas medidas de seguridad y confidencialidad.

3.5. Identifique elementos de logística operativa asociados al reconocimiento óptico, por ejemplo, distribución de material para la recolección, retorno de material de recolección, manejo de contingencias en caso de pérdidas, destrucción o fallas en los materiales asociados al método de captura y cualquier otro elemento relevante en este sentido.

La logística operativa se realizó entre mayo y principios de septiembre de 2017 y contó con la participación de alrededor de 300 funcionarios de la empresa externa y 60 funcionarios INE.

El proceso se llevó a cabo en 8 etapas:

1. Recepción y registro de material censal: Recepción y clasificación de las cajas con material censal en las instalaciones de la empresa externa.
2. Preparación del material censal: se realizaron actividades para preparar el material censal en miras del ingreso de los cuestionarios censales al sistema de captura de datos.
3. Digitalización: Una vez revisado el material, las hojas de los cuestionarios eran separadas y guillotizadas, adaptándolas para su paso por el escáner.
4. Lectura óptica: Una vez digitalizado el material, las imágenes eran comparadas con una plantilla del cuestionario a procesar para identificar el número de página al que correspondía cada una, e ingresaban a un sistema de Reconocimiento Óptico de marcas y caracteres. Es en este proceso de Reconocimiento Óptico que se capturan los datos presentes en el cuestionario.
5. Verificación interactiva: El personal de la empresa externa hacía un reconocimiento visual de posibles errores de lectura del escáner, especialmente en los campos de texto.
6. Verificación INE: Los funcionarios del INE revisaban las variables claves de los cuestionarios, comparando las imágenes en pantalla con la información que se recogía en la base de datos, en aquellos casos en que el reconocimiento automático no lograba el nivel de confianza deseado.
7. Controles y seguimiento del flujo de trabajo a través de un sistema de trazabilidad.
8. Salida: Una vez concluidos los procesos anteriormente mencionados, se generaban las bases de datos e imágenes finales que son entregadas al INE.

3.6. Describa la forma de monitoreo del funcionamiento del reconocimiento óptico. Indique como se realizó el monitoreo y control del scanner; la forma de control y monitoreo de los cuestionarios en papel para asegurar que todos los cuestionarios fueron escaneados. Qué validaciones se realizaron para evitar duplicados (incluir algunos criterios de identificación), pérdida de datos durante la captura, o datos erróneos por mal funcionamiento del método.

Durante la digitalización, cada página escaneada generaba dos imágenes: una a color (como respaldo del original en papel) y otra en blanco y negro (que era la que se utilizaba para la lectura óptica). Si durante el proceso se detectaba un error o problema en la calidad de la imagen obtenida, entonces se detenía el procesamiento del portafolio y se volvía a digitalizar completamente hasta obtener imágenes con la calidad esperada. Junto a esto se escaneaba el código de barra de cada página de las cédulas censales, con el objetivo de verificar los formatos, completitud y orden de las imágenes contenidas en cada uno de los cuestionarios. Por último, las imágenes escaneadas eran sometidas a un control de calidad, de modo que se revisaban un 25% del total.

El INE hizo un control y seguimiento a la calidad del trabajo desarrollado en la empresa. En primer lugar, revisó las imágenes de los cuestionarios censales y las bases de datos, usando criterios como la coincidencia de portafolios y cuestionarios con la base de datos. También revisó la posición y correspondencia de cada hoja al interior del cuestionario, mediante la verificación del código de barras, el que hacía único a cada registro. Se cotejaron imágenes de color de los cuestionarios censales contra la base de datos de cada una de las entregas recibidas por el INE, al igual que con las imágenes en blanco y negro. Junto con ello, se comprobó que las bases de datos entregadas por la empresa tuviesen la identificación geográfica correcta de cada portafolio, la que fue entregada al inicio del proceso.

En segundo lugar, se efectuó un seguimiento de cada uno de los portafolios en las distintas etapas del proceso realizado dentro de la empresa. A través del sistema de trazabilidad, fue posible observar los estados de avance del proceso y verificar que ningún portafolio quedase estancado en alguna de las fases. Con este sistema de seguimiento se pudo conocer los avances de las etapas de preparación física de los cuestionarios y las posteriores a la digitalización.

En tercer lugar, se llevó a cabo un control sobre la consistencia estructural de la base de datos. Esto significó verificar que cada una de las variables definidas para el proceso de captura de datos de los cuestionarios censales se encontrara dentro de sus dominios. En este paso también se verificó la consistencia con las categorías de las respuestas. El objetivo de esta fase era verificar tres elementos estructurales de los datos:

1. Control sobre ubicación geográfica: En esta fase se revisaban las cargas de trabajo que tenían inconsistencia entre el sector geográfico declarado en el sistema versus el leído por lectura óptica, esto era revisado por personal del INE que verificaba y corregía la vinculación geográfica a partir de un análisis de las imágenes y direcciones.
2. Control de enlace vivienda-hogar: En caso de inconsistencia de los números vivienda y hogar era revisado por personal del INE que revisaba las imágenes y reconstruía los hogares y viviendas en la empresa para poder viabilizar el algoritmo posterior que se aplicaría una vez terminada la fase de digitalización.
3. Control poblacional: Revisión de marcas en hojas de persona y consistencia entre registro de personas del hogar contra marcas en hojas de persona.
4. Control de calidad de lectura: Se revisaban los casos que presentaban dobles marcas, verificación de datos como portafolios que arrojaban personas solo de un mismo sexo, por ejemplo, puros hombres. En esos casos siempre se revisaba con la imagen o se reprocesaban imágenes una vez traspasados los cuestionarios a uno nuevo para facilitar la lectura.

4. Síntesis del aprendizaje sobre el método de captura.

4.1. Si usted tuviera la oportunidad de entregar tres recomendaciones sobre este método de captura a otro país, ¿qué elementos a partir de su experiencia deberían ser críticos para que el método funcione de forma adecuada?

Recomendación 1:

La primera recomendación es ejecutar el censo experimental con la misma empresa o método de lectura óptica a utilizar en Censo, esto fue de gran utilidad para el INE de Chile al identificar las principales fuentes de error del método y establecer los protocolos para que se cumplieran los tiempos de ejecución de la fase de reconocimiento óptico acordadas en el contrato, pero manteniendo un control de calidad en cada parte del proceso.

Recomendación 2:

Existen diferencias claras en la calidad de los software de reconocimiento cuando estos se enfrentan a una marca (99,7%), a un número o una letra, y en particular cuando los números y letras son escritos a mano (80%) o impresos (99%), en este sentido, un avance central para la ubicación geográfica de los datos fue la incorporación de un sticker impreso para el código de ubicación el cual se pegaba en un cuadro del cuestionario lo que aumento su calidad. En ese sentido el diseño del cuestionario debe considerar esas diferencias para evitar la mayor cantidad de escritura a mano posible.

Recomendación 3:

El principal problema durante la fase de procesamiento es la construcción de una base relacional que incorpore vivienda-hogar y contabilice a todas las personas, en este sentido, se recomienda establecer un estricto control de los campos del cuestionario que son utilizados como llave para vincular cuestionarios y construir la vivienda y hogar, además de establecer control sobre los cuestionarios con pocas marcas para identificar si son datos de personas reales o errores de lectura asociados a dobleces o borrones.

5. Planes para el próximo censo de población y vivienda.

5.1. Marque con una X el Tipo de Censo que espera realizar para la próxima ronda censal en su país:

De Hecho De Derecho

5.2. Indique la duración de la recolección censal que espera para su próximo censo en el país, en semanas.

semanas. Si su censo será de hecho anotar 0.

5.3. Indique los métodos de recolección o captura que espera utilizar en su país para el próximo censo con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas que espera cubrir con cada método.

	% de Viviendas	% de Población
<input type="checkbox"/> Papel/Digitación		
<input type="checkbox"/> Papel/Reconocimiento Óptico		
<input type="checkbox"/> DMC / CAPI	90	90
<input type="checkbox"/> Web / CAWI	10	10

5.4. Identificar con una X todos los tipos de cuestionario censal que va a considerar en el próximo censo en su país.

<input checked="" type="checkbox"/>	Viviendas particulares
<input checked="" type="checkbox"/>	Viviendas Colectivas
<input checked="" type="checkbox"/>	Personas en Situación de Calle
<input type="checkbox"/>	Personas en Tránsito o Viajeros
<input type="checkbox"/>	Otros

Si marca "otros", especificar en este recuadro:

5.5. Describa de forma breve los métodos a utilizar en el próximo censo, explicita de forma general las razones de mantener o cambiar los métodos utilizados. Además, incorpore los principales desafíos y avances en la preparación del censo en esta materia.

El Censo 2023 considera para su ejecución una serie de innovaciones operativas, entre las que destacamos la incorporación de distintas herramientas tecnológicas y estrategias operativas para recoger los datos de las personas y la extensión del período de recolección, al tratarse de un censo con metodología de derecho. Estas innovaciones impactan de distinta forma en el diseño operativo, ya que muchas de las decisiones que se deben tomar no cuentan con antecedentes de los resultados de su implementación y, por ende, se vuelve vital la necesidad de realizar pruebas censales, que nos permitan obtener información sobre la pertinencia del diseño, los ajustes necesarios y, al mismo tiempo, que nos aporten datos cruciales para la planificación de las actividades de la recolección.

Uno de los grandes desafíos de este proyecto, consiste en realizar un levantamiento censal en el que la estrategia presencial y no presencial se implementen de manera integrada, permitiendo su funcionamiento alternado al mismo tiempo que se amplían las posibilidades de respuesta de los informantes, todo esto con el objetivo de facilitar y potenciar la participación de la población.

La recolección de la información considera un recorrido inicial del territorio, en donde se deben ubicar geográficamente las unidades, sean edificaciones o sitios, lo que permite asociarlas a un código único y, con este, identificarlas y distinguirlas entre ellas. Una vez realizada su identificación en terreno, se procede con el levantamiento de la información a través de la realización de una entrevista presencial con un informante en la unidad en el caso que corresponda, usando para ello un dispositivo móvil de captura de información (DMC).

Por otro lado, la estrategia no presencial considera la recolección de la información censal de las personas residentes habituales de la unidad previamente identificada, por medio de 1) una entrevista no presencial con un operador telefónico o de 2) el auto empadronamiento en una plataforma web.

Debido a la relevancia de la georreferenciación, la ubicación de las viviendas debe realizarse a través de una visita inicial que le permita al censista identificar en terreno el tipo de unidad y si en ella existen residentes habituales. Esto último es lo que posibilitaría iniciar el contacto con algún informante. Esta información es importante, debido a que la presencia o disposición de los informantes determinará el tipo de estrategia (presencial o no presencial) con la que se recogerá la información.

Uno de los principales desafíos que se observan son los problemas de duplicación de información y optimizar el uso de los recursos disponibles, junto con el implementar de forma masiva por primera vez estas estrategias combinadas en el marco de un censo.

5.6. Si tuviera la oportunidad de realizar tres preguntas a otros países de América Latina y el Caribe que hayan implementado el método captura, o los métodos de captura, que usted proyecta para el próximo censo, ¿cuáles serían?

Pregunta 1:

¿Como se asegura la vinculación entre los distintos tipos de métodos de recolección cuando se usa métodos mixtos en especial para la medición de cobertura?

Pregunta 2:

¿Como fueron las pruebas realizadas para la elección del dispositivo móvil y el diseño de la recolección censal?

Pregunta 3:

¿Cuáles serían los principales elementos que deberíamos considerar para el correcto monitoreo y control de los sistemas y transferencias de datos durante la fase de recolección?

Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe
Grupo de Trabajo Censos de la Conferencia Estadística de las Américas – CEPAL
Documento Lineamientos generales para la captura de datos censales

En el marco del plan de trabajo del **Grupo de censos de la CEA CEPAL** para el **bienio 2020-2021** se comprometió la realización de un documento que abordara lineamientos generales para la captura de datos censales para los países de América Latina y el Caribe. Uno de los objetivos específicos es la recopilación de experiencias por parte de los países respecto de la utilización de algunos de los métodos de recolección y captura de datos.

La presente ficha tiene por objetivo sistematizar la experiencia nacional respecto del método de captura de datos censales que ha sido solicitada a usted como representante de su país por parte del equipo de coordinación del Grupo de Trabajo de Censos de CEA Cepal.

La información aquí recopilada será utilizada como insumo para la redacción del documento, durante febrero del 2021 le enviaremos la versión preliminar del documento que será generado con su aporte para que sea previamente aprobada por la dirección de la oficina nacional de estadísticas de su país.

La ficha que se presenta consta de **5 partes**: Información de contacto; Características generales del último censo realizado en su país; Información del Método de Captura; Síntesis de aprendizajes; y Planes para el próximo censo.

Le pedimos enviar la ficha completa a más tardar el día **16 de octubre del 2020**. Su contenido será revisado por el equipo de coordinación del Grupo de Trabajo de Censos y de ser necesario le solicitará una reunión para profundizar en alguno de sus aspectos.

Agradecemos desde ya su colaboración en este proceso y ante cualquier duda o consulta por favor contactarse con Olga Barquero del INE-Chile: olga.barquero@ine.cl y Alejandra Silva de CELADE: alejandra.silva@cepal.org

1. Información de Contacto.

Por favor llenar con su información de contacto como responsable del llenado de la ficha por parte de su país.

Nombre	EDNA MARGARITA VALLE CABRERA
Correo electrónico	emvallec@dane.gov.co
Cargo	Directora de Censos y Demografía (E)
Institución	DANE
País	COLOMBIA

Nombre	MARIANA FRANCISCA OSPINA BOHÓRQUEZ
--------	------------------------------------

Correo electrónico	mfospinab@dane.gov.co
Cargo	Coordinadora GIT Proyecciones de población y análisis demográfico - Dirección de Censos y Demografía
Institución	DANE
País	COLOMBIA

Nombre	SILVIA ALEJANDRA HERNÁNDEZ OSORIO
Correo electrónico	sahernandezo@dane.gov.co
Cargo	Contratista GIT Relacionamento Nacional e Internacional – Dirección General
Institución	DANE
País	COLOMBIA

2. Características generales del **último censo** realizado en el país.

2.1. Marque con una X el Tipo de Censo para el último censo realizado en su país:

De Hecho De Derecho

2.2. Indique la fecha de inicio de la recolección del último censo en su país:

Día Mes Año

2.3. Indique la duración de la recolección del último censo en su país en semanas, se debe incluir el período oficial además de las semanas que se extendió la recolección más allá de ese período hasta el cierre del operativo censal.

a. Duración del período oficial: semanas. Si su censo fue de hecho anotar 0.

En caso de que quiera especificar condiciones especiales para el período oficial de recolección, como, por ejemplo, diferencia de duración entre urbano y rural, especifique en el cuadro siguiente:

DEL 9 DE ENERO AL 12 DE ABRIL SE HIZO EL E-CENSO; DEL 26 DE ABRIL AL 30 DE OCTUBRE RECOLECCIÓN PRESENCIAL

La fase de recolección presencial se extendió por los siguientes motivos:

- Factores climáticos.
- Problemas de orden público.
- Resistencia por parte de actores sociales.
- Problemas administrativos relacionados con el personal y el transporte.
- Mayor cantidad de unidades a visitar con respecto a las programadas.
- Rediseño de rutas que se facilitaban desde municipios diferentes a los inicialmente programados.
- Conflictos o rechazos al censo en las áreas limítrofes de algunos municipios.
- Bajos rendimientos diarios, especialmente del personal étnico contratado (indígena y afrodescendientes) que exigían la prolongación del censo.

b. ¿Cuántas semanas se extendió la recolección después del período oficial?: semanas.

2.4. Indique los métodos de recolección o captura utilizados en el último censo en su país con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas cubiertas mediante ese método.

		% de Viviendas	% de Población
<input checked="" type="checkbox"/>	Papel/Digitación	1,45%	2,44%
<input type="checkbox"/>	Papel/Reconocimiento Óptico		
<input checked="" type="checkbox"/>	DMC / CAPI	86,49%	85,17%
<input checked="" type="checkbox"/>	Web / CAWI	12,05%	12,38%

2.5. *Describa de forma general la estrategia operativa del censo. Incluya si los métodos de recolección o captura se aplicaron de forma secuencial o en paralelo; qué métodos de recolección o captura fueron utilizados para operativos especiales (viviendas colectivas, situación de calle, población en tránsito u otro); indicar si la captura de datos fue centralizada o descentralizada, o cualquier otra información relevante para el contexto de su experiencia.*

En la estrategia operativa del censo, se consideró recolectar la información mediante cuatro (4) estrategias de recolección:

- **ELECTRÓNICO:** Recolección mediante el diligenciamiento del cuestionario censal por medio de una aplicación web (eCenso).
- **BARRIDO:** Recorrido sistemático de los equipos de recolección que permitía regresar diariamente a entregar reportes, visita vivienda a vivienda.
- **RUTA:** Recorrido de territorios cuyo acceso geográfico no permitía el regreso diario del equipo de trabajo al centro operativo a entregar reportes, esto solo se realizaba al finalizar cada ruta.
- **MIXTO:** Recorrido que contemplaba los dos métodos de recolección (barrido y ruta) con sus respectivas estrategias.

La captura de la información se hizo principalmente en DMC; cuando se recolectaba en cuestionarios a papel, se digitaba luego en los DMC en los centros operativos, esto se realizaba en forma paralela.

La recolección en las viviendas colectivas (Lugares Especiales de Alojamiento _LEA), se realizó con la metodología de “focalizado”, por medio de una plataforma digital donde las instituciones registraban su información. Se apoyó en las entidades que las regulan, El Instituto Nacional Penitenciario y Carcelario (INPEC); el Instituto Colombiano de Bienestar Familiar (ICBF), el Ministerio de Defensa, la Agencia Colombiana para la Reintegración, entre otros. Esta información se captó en cuestionario análogo y a través de una plantilla en la web.

Para los Habitantes de la Calle, siguiendo los lineamientos de la Ley 1641 de 2013, el DANE procedió a adelantar una serie de convenios con las alcaldías municipales y las instituciones que tienen programas de atención a dicha población, para adelantar el censo de los Habitantes de la Calle. El contenido del cuestionario para recolectar su información fue diferente al del Censo Nacional, contenía preguntas sociodemográficas tradicionales y otras específicas para atender los lineamientos de política contenidos en dicha Ley.

La captura de la información fue descentralizada in situ, pero una vez cerradas las encuestas completas, se realizaba un reporte automático desde los DMC mediante mensajes de texto al DANE central para evidenciar el avance de los recorridos; la información de los cuestionarios era transmitida a diario desde los centros operativos vía internet al DANE central, con los respectivos protocolos de seguridad.

El levantamiento de la información se realizó por fases. La primera fase del Censo 2018, en la que se implementó el método electrónico, denominado eCenso, se realizó entre el 9 de enero y el 12 de abril de 2018, mientras que las fases 2, 3 y 4, donde se aplicaron los métodos de recolección de información de forma presencial se desarrolló entre el 26 de abril y el 30 de octubre de 2018.

Segunda fase: Corresponde a la recolección de información de manera presencial o puerta a puerta en las zonas del país que concentran el mayor volumen de población y registran también la mayor dinámica, estas son las capitales de los departamentos más grandes y sus respectivos departamentos: Bogotá, Cundinamarca, Medellín y Antioquia, Cali y Valle del Cauca, Bucaramanga y Santander, Cúcuta y Norte de Santander, Cartagena y Bolívar, y el Archipiélago de San Andrés, Providencia y Santa Catalina.

Tercera fase: Corresponde a la recolección de información de manera presencial o puerta a puerta en las zonas del país que concentran un volumen de población importante y registran una dinámica moderada de población, esto son las capitales de departamento o ciudades medianas y pequeñas con sus respectivos departamentos.

Cuarta fase: Corresponde a la recolección de información de manera presencial en las zonas del país que concentran un volumen de población menor y disperso, registran una dinámica baja y están en zonas con mucha complejidad para desplazamiento, esto son: las regiones geográficas de la Amazonia, Orinoquía, el Anden Pacífico y la Alta Guajira.

3. Características del método: **Dispositivo Móvil de Captura.**

3.2. Si usted tuviera que defender el uso de dispositivos móviles de captura frente a sus pares de otros países, ¿podría sintetizar el proceso de decisión para la elección de este método como parte del proceso de recolección censal?

El uso del DMC para captar información en encuestas o censos es de vital importancia dado que mejoran los rendimientos en la toma de información, se disminuyen los errores que se cometen al tomar la información dado que las preguntas se validan previamente con manejo de flujos, rangos, universos y consistencia entre variables, conformación y consolidación de resultados en tiempo oportuno, de forma que sus resultados se pueden entregar una vez culmine el operativo una vez se realice la validación e imputación de la información faltante de ser necesario. Además, con el uso de los DMC se evita el proceso de crítica y captura de la información, lo que disminuye costos en este proceso y se obtiene una información con mayor calidad.

En la siguiente tabla se muestra el porcentaje de datos faltantes e inconsistentes para el DMC:

Método de Recolección	Porcentaje de Datos Faltantes	Porcentaje de Datos Inconsistentes
Por Dispositivo Móvil de Captura (DMC)	0.37%	0.39%

Fuente: DANE – Dirección de Censos y Demografía

En resumen, se defiende el uso de DMC por las siguientes razones:

- Disminuye los tiempos de entrevista.
- Mejora el manejo de los cuestionarios en lo relacionado con los flujos, los filtros y las normas de validación que se incluyen en los programas de recolección.
- Al poder ser reutilizados en otras operaciones estadísticas, los presupuestos de materiales se disminuyen, por cuanto se puede disminuir sustancialmente el uso del papel; este beneficio también abarca el tema del manejo ambiental.
- Se pueden utilizar procesos de control de consistencia y de validación en forma automática, así como también el de codificación.
- La transmisión de la información es más eficiente en cuanto al tiempo que demora en llegar a las oficinas del DANE central y pueden incluirse correctivos a corto plazo durante la recolección.

3.3. Indique las pruebas realizadas durante la preparación de la recolección con dispositivos móviles de captura, incluir todas las pruebas tecnológicas, de cuestionario, operativas y censo experimental, donde se haya evaluado este método. Identificar los principales hallazgos en esas pruebas.

Además de las pruebas de escritorio realizadas en el DANE con diferentes funcionarios de la entidad, la mayor prueba se realizó en el Censo experimental de Jamundí, Valle del Cauca en 2016, donde se evidenciaron las ventajas de su uso y las fallas (recalentamiento de máquinas y bloqueos del cuestionario) y limitaciones en el Software, situaciones que fueron corregidas por el equipo de sistemas.

En este orden de ideas, se realizaron:

- Pruebas a los cuestionarios análogos
- Pruebas temáticas de las preguntas incluidas y excluidas
- Pruebas de escritorio para identificar la comprensión de las preguntas
- Pruebas de campo con el fin de captar comprensión y tiempos de diligenciamiento

Para el DMC se elaboran las normas de validación de flujos, rangos y consistencia del cuestionario análogo y se gestiona con la Oficina de Sistemas, una vez se desarrolla el programa de captura del cuestionario análogo se procede a validar que todas las normas se hayan incluido mediante pruebas de escritorio, al igual que se prueba la capacidad del dispositivo y tiempo de carga de las baterías para evitar que inconvenientes en el desarrollo de la encuesta.

Estas pruebas se describen en una matriz que identifica los errores que se presenten, para que sean tenidos en cuenta por la Oficina de Sistemas y así, una vez se superen todos los inconvenientes se da vía libre al programa de captura diseñado para tal fin.

Por otra parte, una vez se realizaron las pruebas de campo, se evalúa y analiza la captura con datos reales, se valida tanto el programa como los tiempos en el diligenciamiento y manejo del dispositivo.

3.4. Sintetice el desarrollo TI de la recolección con dispositivos móviles de captura. Indique si el software de recolección fue desarrollado in house o se utilizó un software externo. Se solicita describir procesos de respaldo, transferencia de datos, elementos de seguridad y protección de datos personales (encriptaciones, manejo de claves y accesos, etc.); Incluir consideraciones para la adquisición (licitaciones de compra, arriendo o provisión de servicios).

El desarrollo de una de las herramientas de captura, como es la utilización de DMC, se realizó in house, bajo plataforma Android.

Una vez el encuestador finaliza cada encuesta, esta es almacenada en el dispositivo y genera un mensaje de texto que es enviado a través de un operador móvil celular (bajo un proceso de licitación se provee el servicio) y es consolidada en el esquema de base de datos. El mensaje de datos no pudo superar los 160 caracteres de texto por lo que dichos mensajes contendrán una codificación que permite identificar el departamento, municipio, encuestador, supervisor, encuesta, total de hogares, total de viviendas, total de personas residentes, estado de la encuesta y estado de la vivienda.

Al finalizar la jornada de recolección, el supervisor disponía de una opción para generar las encuestas finalizadas, las cuales eran almacenadas localmente encriptadas a través de un algoritmo asimétrico RSA con una longitud de llave de 2048 bits en formato de texto en una estructura Json y comprimida en formato zip, las encuestas generadas son cambiadas de estado para evitar el reproceso de validación en base de datos local del DMC y el archivo generado es almacenado como copia en la tarjeta SD.

Una vez el apoyo informático descarga la información generada, era transmitida a través de una red virtual privada (VPN Virtual Private Network, por sus siglas en inglés), en dicho proceso cada archivo es almacenado en un servidor y calendarizado por los componentes de software encargados de realizar la descompresión del archivo, la desencriptación y validación para efectuar el cargue a la base de datos. La información se tenía disponible en un transcurso no mayor a 8 horas una vez es realizado el proceso por el apoyo informático.

Estructura de árbol para el almacenamiento de archivos:

Con el propósito de facilitar la ubicación de los archivos se definió una estructura en la que se creó como raíz una carpeta CNPV en la cual se definieron dos subcarpetas:

- CNPV2018 – Hogares / LEA

CNPV2018 para almacenar los archivos que contendrán la información de los hogares, en la cual se creó una subcarpeta por cada departamento, una subcarpeta por cada uno de sus municipios y a su vez una carpeta con cada una de las unidades de cobertura y finalmente una carpeta con la fecha en la cual fue realizada la transmisión de los archivos.

Una vez finalizado el operativo de recolección se realizará una copia de respaldo de todos los archivos generados durante el proceso y almacenados en cada DMC como respaldo y verificaciones y validaciones posteriores. Este respaldo es enviado a través del protocolo de transferencia de archivos (FTP) y por medio de la VPN.

Para el desarrollo del software de captura se requerían las normas de validación y consistencia siguiendo el cuestionario aplicado, así mismo, la transmisión de la información se realizaba por parte de los supervisores y apoyos informáticos de acuerdo con los lineamientos anteriormente descritos.

3.5. Identifique elementos de logística operativa asociados a los dispositivos móviles de captura, por ejemplo, distribución de material para la recolección, retorno de material de recolección, manejo de contingencias en caso de pérdidas, destrucción o fallas en los materiales y cualquier otro elemento relevante en este sentido.

Se emplearon formatos para la entrega y recepción de los dispositivos al personal de campo, al igual que hay formatos en caso de daños o robos.

No se emplearon formatos que identifiquen errores cometidos en diligenciamiento o sobre la forma de corrección en campo.

En la eventualidad en que se detectaron errores en la validación de la información se devuelven en un formato para que la fuente primaria hiciera las correcciones correspondientes, en este caso, se debe restablecer la encuesta que se capturo antes de ser transmitida, es decir, se hacían los ajustes correspondientes y se volvían a enviar avisando en el formato correspondiente a la Oficina de Sistemas para eliminar la anterior encuesta y cargar la actualización.

3.6. *Describa la forma de monitoreo de la recolección con dispositivos móviles de captura. Indique como se realizó el monitoreo y control de la transferencia de datos; la forma de control y monitoreo del avance de la recolección en términos de cobertura territorial y poblacional. Qué validaciones se realizaron para evitar duplicados (incluir algunos criterios de identificación), pérdida de datos durante la captura, o datos erróneos por mal funcionamiento del método.*

Control de cobertura

Se realizó a partir de los formatos operativos diligenciados diariamente por los censistas para cada unidad de cobertura asignada. Allí registraban los resultados sobre las unidades visitadas (Viviendas, unidades mixtas y no residenciales): tipo de ocupación de la vivienda y resultado de entrevista. Esta información fue consolidada por los supervisores diariamente y enviada a los coordinadores, quienes a su vez revisaban y consolidaban la información, para enviarla al siguiente nivel para su consolidación hasta llegar al coordinador departamental, permitiendo ejercer el seguimiento y monitoreo permanente directamente desde los municipios.

El control de la cobertura en campo del CNPV 2018 mediante el método de barrido fue complementario a través del visor geográfico, en el cual los coordinadores departamentales podían observar que Unidad de Cobertura Urbana o Unidad de Cobertura Rural habían trabajado y cuáles les faltaban. Posteriormente, desde el DANE Central, en cabeza de la Dirección de Geoestadística, se le enviaba un reporte a los coordinadores departamentales que permitía tener un control de la cobertura efectivo.

Se dispuso este visor geográfico con fines de seguimiento geográfico a los reportes de la herramienta SPOT (dispositivo que incluye tecnología GPS, con la funcionalidad de enviar reportes sobre la posición geográfica del personal asignado a los grupos de trabajo del método de rutas y que dentro de sus funciones estuvo el manejo de dicho dispositivo) en los grupos de trabajo de rutas. Esta herramienta permitió observar en tiempo real el avance de los grupos de trabajo asignados en las áreas operativas de rutas, como una alternativa para que desde el nivel central se evidenciara que el personal estuviera recorriendo las áreas de trabajo.

El marco de referencia utilizado para el seguimiento geográfico a las rutas mediante esta herramienta corresponde al marco de las áreas operativas diseñadas para el método de rutas, junto con la información oficial, como los límites político-administrativos departamentales y municipales, el cual se complementó con la ubicación referencial de comunidades-veredas-rancherías y demás formas de concentración de la población residente en estos territorios.

Mediante la revisión del visor geográfico dispuesto con los reportes de la herramienta SPOT, la cual fue adelantada por personal del nivel central, se evidenció el avance de los grupos de trabajo para los cuales se observaron reportes en el visor, lo cual complementó los informes provenientes por parte del personal responsable del operativo en estos territorios (coordinadores departamentales, jefes municipales, profesionales operativos de rutas).

Seguimiento desde nivel central – CONAL: el Centro Operativo Nacional (CONAL) ejerció una regulación de atención y solución de novedades que surgieron en el desarrollo operativo, soportado por personal idóneo de diferentes áreas como temática (censos y demografía, geoestadística, de sistemas y administrativa, entre otras que impactaron al operativo. Desde el nivel central, se entabla comunicación directa con la estructura organizacional del orden departamental encabezado por el coordinador departamental y su línea de colaboradores, con el fin de conocer los pormenores de la ejecución en el departamento, los posibles inconvenientes que se presenten, la gestión y solución de estos siempre teniendo como marco la metodología técnica censal. En este sentido, la transmisión desde los DMC como proceso de transmisión de los cuestionarios de hogar y se divide en dos partes que cuentan con controles que se describen a continuación, para cada uno de estos subprocesos:

Generación de archivos para transmitir: una vez los supervisores generan los archivos a transmitir. Cuando existen encuestas para transmitir, la aplicación informa que se está generando el archivo (con extensión .Zip) para la transmisión y muestra la ruta de almacenamiento. En la transmisión las encuestas con estado Cerrada desaparecen de la lista.

Transmisión de información por parte del Apoyo informático: al realizar el envío exitoso se muestra un mensaje informativo en la parte superior del formulario y se actualiza el estado para cada uno de los archivos cargados.

Para llevar a cabo el adecuado reporte y seguimiento de los envíos de la información recolectada en terreno del CNPV-2018 al DANE central, se diseñaron las metodologías para el cálculo de indicadores con los que se podía evidenciar el avance del operativo de recolección (cobertura territorial y de población) y la calidad en la transmisión de los datos, medido con dos instrumentos:

- Con las encuestas “notificadas”, que correspondían a las entrevistas cerradas por el censista cada vez que terminaba el diligenciamiento completo de un cuestionario y se notificaban en forma automática mediante mensajes de texto (SMS) reportados a la oficina de Sistemas del DANE central.
- Con la “transmisión” de las encuestas con los archivos que contenían los datos de las viviendas, los hogares y las personas censadas; esta labor se realizaba a diario en los centros de acopio en cada municipio una vez terminada la jornada de recolección y eran reportados a la Dirección de Censos y Demografía del DANE.

Para cumplir con lo anterior, se diseñaron tres (3) indicadores básicos, a saber:

- INDICADOR DE EFECTIVIDAD DE TRANSMISIÓN –IET., cuyo objetivo era mostrar el porcentaje de encuestas transmitidas con relación a las encuestas notificadas.
- INDICADOR DE RENDIMIENTO DIARIO –IRD., este indicador se desarrolló para mostrar el número de encuestas que realizaba cada uno de los censistas o supervisores durante un periodo determinado. Se calculaba tanto para encuestas transmitidas como notificadas por separado por cuanto provenían de dos bases de datos diferentes.
- INDICADOR DE PERSONAL EFECTIVO –IPEF., Este indicador se desarrolló para mostrar el número de encuestadores transmitiendo o notificando encuestas por municipio, como proporción del total de encuestadores contratados por municipio cada día en un periodo de reporte determinado. Se calculaba tanto para encuestas transmitidas como notificadas por separado porque provenían de dos bases de datos diferentes.

4. Síntesis del aprendizaje sobre el método de captura.

4.1. Si usted tuviera la oportunidad de entregar tres recomendaciones sobre este método de captura a otro país, ¿qué elementos a partir de su experiencia deberían ser críticos para que el método funcione de forma adecuada?

Recomendación 1:

Deben hacerse suficientes pruebas para determinar la robustez de software y la resistencia del hardware. Los programas incluidos para diligenciamiento de los cuestionarios deben funcionar con todas las consideraciones técnicas, metodológicas y temáticas, y con cartografía actualizada.

Por calidad de la información, no se recomienda la recolección en papel, por lo cual el uso de cuestionarios análogos debe ser excepcional y muy restringido.

Recomendación 2:

En el proceso de aprendizaje, utilizar más tiempo en la navegabilidad del aplicativo, en este proceso permitir que el personal de campo entre y salga en diferentes momentos, cierre de encuestas y retoma de las mismas, qué hacer al equivocarse del registro de personas, lo importante es que antes de salir a campo tenga la navegabilidad completa del aplicativo y sepan los puntos clave para su diligenciamiento

Recomendación 3:

Utilizar DMC de última tecnología. Así mismo, se recomienda para el monitoreo el uso de GPS u otros instrumentos de localización para controlar los recorridos y asegurar a los equipos de trabajo, con un adecuado plan de datos para los envíos de información a través de internet.

Por último, la comunicación entre áreas internas del operativo censal se basa en la correcta interacción con los equipos de sistemas, que son fundamentales para la articulación de los procesos y el monitoreo no sólo técnico sino en especial con las áreas administrativas y de contratación de personal, servicios de logística y transporte, así mismo, ellos deben garantizar la implementación de llaves compuestas y únicas para el seguimiento de toda la información recolectada que facilite los procesamientos y la difusión de resultados.

5. Planes para el próximo censo de población y vivienda.

5.1. Marque con una X el Tipo de Censo que espera realizar para la próxima ronda censal en su país:

De Hecho De Derecho

5.2. Indique la duración de la recolección censal que espera para su próximo censo en el país, en semanas.

semanas. Si su censo será de hecho anotar 0.

5.3. Indique los métodos de recolección o captura que espera utilizar en su país para el próximo censo con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas que espera cubrir con cada método.

		% de Viviendas	% de Población
<input type="checkbox"/>	Papel/Digitación		
<input type="checkbox"/>	Papel/Reconocimiento Óptico		
<input checked="" type="checkbox"/>	DMC / CAPI	85%	85%
<input checked="" type="checkbox"/>	Web / CAWI	15%	15%

5.4. Identificar con una X todos los tipos de cuestionario censal que va a considerar en el próximo censo en su país.

<input checked="" type="checkbox"/>	Viviendas particulares
<input checked="" type="checkbox"/>	Viviendas Colectivas
<input type="checkbox"/>	Personas en Situación de Calle
<input type="checkbox"/>	Personas en Tránsito o Viajeros
<input checked="" type="checkbox"/>	Otros

Si marca "otros", especificar en este recuadro:

De equipamiento y entorno urbanístico

5.5. Describa de forma breve los métodos a utilizar en el próximo censo, explicita de forma general las razones de mantener o cambiar los métodos utilizados. Además, incorpore los principales desafíos y avances en la preparación del censo en esta materia.

En general, se mantienen los mismos métodos aplicados en el más reciente operativo censal de 2018.

Para el próximo censo se debe pensar en continuar en un censo de tiempo ampliado, pero hacer las estrategias de contratación necesarias para poder contar con todo el personal requerido en los tiempos previstos y que este no sea el impedimento para poder cerrar el operativo censal puerta a puerta en no más de tres (3) meses.

Se recomienda tener en cuenta con relación a los métodos a emplear en el próximo censo serian: el electrónico (eCenso), con una campaña masiva de forma tal que se logre una mayor cobertura por este medio teniendo en cuenta no solo el sector público, sino además contar con el acompañamiento de los gremios, las empresas, la academia y campañas masivas por medios de comunicación, entre otros. Revisar si se pueden emplear DMC con mayor capacidad, de forma tal que se puedan establecer mayores controles o cruces de variables mejorando la validez directamente con la fuente y dejar los tiempos necesarios en la programación del operativo para que en los casos de emplear cuestionarios análogos estos sean capturados cerca de la fuente por si se requieran validaciones adicionales estas permitan revisitas para regresar a corregir fallas directamente con la fuente antes de cerrar el operativo de campo.

5.6. Si tuviera la oportunidad de realizar tres preguntas a otros países de América Latina y el Caribe que hayan implementado el método captura, o los métodos de captura, que usted proyecta para el próximo censo, ¿cuáles serían?

Pregunta 1:

Si desarrollan el proceso del censo electrónico, cuáles serían los planes de socialización para una mejor aceptación por parte de la ciudadanía.

Pregunta 2:

Cómo sería la forma de contratación del personal que desarrollaría el proceso de recolección en un periodo ampliado, evitando las deserciones; ¿y que su contratación se haga de forma ágil en los tiempos previstos para evitar demoras en la contratación? ¿cómo sistematizaron el seguimiento a la contratación y cómo controlaron la disposición de personal efectivo en campo?

Pregunta 3:

¿Cuál cree que es el mecanismo de transmisión de información más efectivo en términos de seguridad de la información y rapidez de transmisión? Es decir, ¿En zonas con dificultades de conectividad, ¿cómo hicieron la transmisión de la información?, ¿Cómo medían la calidad de la transmisión de los datos en cuanto a la completitud, que lo que se mandara fuera exactamente lo que llegaba?

Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe
Grupo de Trabajo Censos de la Conferencia Estadística de las Américas – CEPAL
Documento Lineamientos generales para la captura de datos censales

En el marco del plan de trabajo del **Grupo de censos de la CEA CEPAL** para el **bienio 2020-2021** se comprometió la realización de un documento que abordara lineamientos generales para la captura de datos censales para los países de América Latina y el Caribe. Uno de los objetivos específicos es la recopilación de experiencias por parte de los países respecto de la utilización de algunos de los métodos de recolección y captura de datos.

La presente ficha tiene por objetivo sistematizar la experiencia nacional respecto del método de captura de datos censales que ha sido solicitada a usted como representante de su país por parte del equipo de coordinación del Grupo de Trabajo de Censos de CEA Cepal.

La información aquí recopilada será utilizada como insumo para la redacción del documento, durante febrero del 2021 le enviaremos la versión preliminar del documento que será generado con su aporte para que sea previamente aprobada por la dirección de la oficina nacional de estadísticas de su país.

La ficha que se presenta consta de **5 partes**: Información de contacto; Características generales del último censo realizado en su país; Información del Método de Captura; Síntesis de aprendizajes; y Planes para el próximo censo.

Le pedimos enviar la ficha completa a más tardar el día **16 de octubre del 2020**. Su contenido será revisado por el equipo de coordinación del Grupo de Trabajo de Censos y de ser necesario le solicitará una reunión para profundizar en alguno de sus aspectos.

Agradecemos desde ya su colaboración en este proceso y ante cualquier duda o consulta por favor contactarse con Olga Barquero del INE-Chile: olga.barquero@ine.cl y Alejandra Silva de CELADE: alejandra.silva@cepal.org

1. Información de Contacto.

Por favor llenar con su información de contacto como responsable del llenado de la ficha por parte de su país.

Nombre	Sofía Mora Steiner
Correo electrónico	sofia.mora@inec.go.cr
Cargo	Coordinadora Unidad de Diseño, Análisis y Productos, Censo 2021.
Institución	Instituto Nacional de Estadística y Censos
País	Costa Rica

2. Características generales del **último censo** realizado en el país.

2.1. Marque con una X el Tipo de Censo para el último censo realizado en su país:

De Hecho De Derecho

2.2. Indique la fecha de inicio de la recolección del último censo en su país:

Día Mes Año

2.3. Indique la duración de la recolección del último censo en su país en semanas, se debe incluir el período oficial además de las semanas que se extendió la recolección más allá de ese período hasta el cierre del operativo censal.

a. Duración del período oficial: semanas. Si su censo fue de hecho anotar 0.

En caso de que quiera especificar condiciones especiales para el período oficial de recolección, como, por ejemplo, diferencia de duración entre urbano y rural, especifique en el cuadro siguiente:

Se realizó en 5 días, de lunes a viernes.

b. ¿Cuántas semanas se extendió la recolección después del período oficial?: semanas.

2.4. Indique los métodos de recolección o captura utilizados en el último censo en su país con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas cubiertas mediante ese método.

		% de Viviendas	% de Población
<input type="checkbox"/>	Papel/Digitación		
<input checked="" type="checkbox"/>	Papel/Reconocimiento Óptico	100	100
<input type="checkbox"/>	DMC / CAPI		
<input type="checkbox"/>	Web / CAWI		

2.5. Describa de forma general la estrategia operativa del censo. Incluya si los métodos de recolección o captura se aplicaron de forma secuencial o en paralelo; qué métodos de recolección o captura fueron utilizados para operativos especiales (viviendas colectivas, situación de calle, población en tránsito u otro); indicar si la captura de datos fue centralizada o descentralizada, o cualquier otra información relevante para el contexto de su experiencia.

En el Censo 2011 se utilizó un solo método de recolección, que fue a través de cuestionario en papel. Todos los cuestionarios fueron procesados por medio de reconocimiento óptico, de forma centralizada en las oficinas del INEC.

En los operativos especiales se utilizó el mismo método. En algunos centros penitenciarios se captó cierta información usando los registros administrativos de la institución, pero los datos eran transferidos al cuestionario censal en papel por una persona censista.

Hubo algunos casos de entrevistas por medio telefónico (muy esporádicas), pero del mismo modo, la información se levantó en el papel y posteriormente se procesó en el lector óptico con todo el material.

El operativo se realizó simultáneamente en todo el país, excepto en un territorio indígena, donde se realizó una semana antes porque el ingreso debió ser por vía aérea. En ese lugar también se levantó la información en papel.

3. Características del método: **Dispositivo Móvil de Captura.**

3.2. Si usted tuviera que defender el uso de dispositivos móviles de captura frente a sus pares de otros países, ¿podría sintetizar el proceso de decisión para la elección de este método como parte del proceso de recolección censal?

Costa Rica no ha realizado aún ningún censo nacional utilizando DMC. Sin embargo, el INEC se prepara para implementar el Censo 2021 con DMC.

Para tomar la decisión se analizaron los siguientes factores:

- La experiencia del INEC, desde 2010, con el uso de DMC en otras operaciones estadísticas como encuestas y registros.
- La inversión de la compra de DMC y desarrollo tecnológico frente a la impresión de papel y contratación de lectores ópticos.
- La automatización de los flujos de las entrevistas y los controles que se realizan durante la captura de datos para mejorar la calidad de los datos.
- La oportunidad de mejorar la cobertura censal.
- La oportunidad de validar parte de la información en tiempo real y disminuir los tiempos de procesamiento de datos
- La posibilidad de brindar los resultados censales de una forma más oportuna.
- La accesibilidad y usabilidad del sistema de captura de datos, para disminuir los posibles errores de aplicación del cuestionario, en los diferentes flujos posibles.
- Las condiciones del país respecto a cobertura de internet e infraestructura tecnológica
- Las habilidades tecnológicas, en cuanto al uso de dispositivos móviles, por parte de la población como posibles censistas.

A partir de esas valoraciones, se determinó que, el Censo 2021 de Costa Rica se lleve a cabo utilizando DMC.

3.3. Indique las pruebas realizadas durante la preparación de la recolección con dispositivos móviles de captura, incluir todas las pruebas tecnológicas, de cuestionario, operativas y censo experimental, donde se haya evaluado este método. Identificar los principales hallazgos en esas pruebas.

El desarrollo tecnológico aún está en proceso, por lo que actualmente se están realizando aún prueba.

Al día de hoy se han realizado nueve pruebas de cuestionario, en diferentes contextos del país, donde se utilizaron DMC, no obstante, se probó un sistema de captura distinto al que se utilizará en el Censo Piloto.

Por parte del desarrollador, se han realizado diferentes pruebas de usabilidad con testers de diferentes perfiles para validar la accesibilidad de la aplicación.

Además, se están realizando pruebas sobre la plataforma tecnológica tales como pruebas de rendimiento y de la capacidad de los recursos (almacenamiento, enlaces de internet, memoria, procesamiento entre otros) pruebas de transferencia de datos, pruebas de seguridad (seguridad perimetral, encriptación de los datos y de las aplicaciones) y pruebas de las características técnicas de los dispositivos móviles de captura.

En el mes de noviembre de 2020 se llevará a cabo el Censo piloto, donde se llevarán a prueba todos los sistemas y la infraestructura tecnológica propuesta.

3.4. Sintetice el desarrollo TI de la recolección con dispositivos móviles de captura. Indique si el software de recolección fue desarrollado in house o se utilizó un software externo. Se solicita describir procesos de respaldo, transferencia de datos, elementos de seguridad y protección de datos personales (encriptaciones, manejo de claves y accesos, etc.); Incluir consideraciones para la adquisición (licitaciones de compra, arriendo o provisión de servicios).

El desarrollo de TI para la recolección de datos se está realizando de forma conjunta con el equipo de TI del INEC y el Instituto Tecnológico de Costa Rica (TEC, una universidad estatal).

El TEC fue contratado por el INEC para el desarrollo de tres grandes sistemas:

- Gestor de operativo de campo: sistema que permite el reclutamiento de personal, asignación de cargas de trabajo, seguimiento de reportes y cobertura, asignación de pagos.
- App Censista: es la aplicación de captura de datos, que incluye la digitalización de las áreas de empadronamiento y el enlace del punto de vivienda con la información del cuestionario censal.
- Autocenso: aplicación web que permitirá a las viviendas que no pudieron censarse de manera presencial, hacerlo por medio de la web, utilizando un código que la persona censista le asignará a la vivienda.

Por su parte, el INEC desarrolló los sistemas para la actualización cartográfica, el procesamiento de datos y la gestión de campo desde las agencias censales. Además, trabaja en conjunto con el TEC en la infraestructura que soportará todos los sistemas y los aspectos de replicación de la base de datos.

La transferencia de datos se realizará de forma inmediata a un servidor en nube, una vez que la persona censista finalice la entrevista. En caso de que, al momento no tenga acceso a internet, la transferencia se realizará una vez que el dispositivo se conecte a una red.

Diariamente se replicará la base de datos a un servidor del INEC, la cual pasará al procesamiento de información.

Para la seguridad u protección de datos se trabajará con encriptación y uso de claves de acceso.

3.5. Identifique elementos de logística operativa asociados a los dispositivos móviles de captura, por ejemplo, distribución de material para la recolección, retorno de material de recolección, manejo de contingencias en caso de pérdidas, destrucción o fallas en los materiales y cualquier otro elemento relevante en este sentido.

La logística para el Censo 2021 está determinada de la siguiente manera:

El nivel central (INEC) distribuirá los paquetes de material censal, debidamente armados, a 235 centros operativos en todo el país, a través de Correos de Costa Rica.

En los centros operativos, que serán a su vez, las sedes de capacitación, se distribuirán los paquetes a cada una de las personas censistas y supervisoras asignadas a ese centro. La distribución la harán los coordinadores de centro operativo y los coordinadores de distrito.

Se estima la contratación de 14500 personas en total para el operativo, que en 2021 se realizará en 18 días.

Cada centro operativo contará con un inventario de materiales en caso de que se deba reponer algo, incluyendo los DMC. Esto se gestionará a través de la jerarquía establecida para el operativo de campo (censista solicita a su supervisión y este lo gestiona con la coordinación de distrito). En caso de que no haya dispositivos disponibles de forma oportuna, la persona censista utilizará papel y posteriormente deberá trasladar la información al DMC.

La devolución del material se realizará de la misma forma, el personal censista lo entregará en su centro operativo, donde será organizado para su devolución al nivel central, a través de Correos de Costa Rica.

Con Correos de Costa Rica (institución estatal), se hará una contratación para el servicio de traslado de material.

3.6. Describa la forma de monitoreo de la recolección con dispositivos móviles de captura. Indique como se realizó el monitoreo y control de la transferencia de datos; la forma de control y monitoreo del avance de la recolección en términos de cobertura territorial y poblacional. Qué validaciones se realizaron para evitar duplicados (incluir algunos criterios de identificación), pérdida de datos durante la captura, o datos erróneos por mal funcionamiento del método.

Aún en Costa Rica no se implementa el uso de los DMC, se están realizando pruebas y análisis de la infraestructura.

La identificación de las viviendas surge de la división territorial de Costa Rica (Provincia, Cantón y Distrito) y el levantamiento del marco geoestadístico que realizó el INEC dividiendo el país en UGMs (Unidad Geoestadística Mínima) además se realiza una planificación de las cargas de trabajo de cada Censista y se crean las AE (áreas de empadronamiento) que corresponde a la carga de trabajo de un censista durante el operativo censal. Esto permite tener un control sobre el avance en campo y sobre lo que se espera obtener.

Al utilizar DMC con capacidad de conexión a internet y la utilización de mapas digitales en el proceso de recolección de las entrevistas permite conocer las coordenadas geográficas exactas de los diferentes puntos de las viviendas. Todo esto facilita el monitoreo día con día durante el operativo censal.

El monitoreo se apoya con herramientas de software instaladas en la plataforma para conocer la cobertura censal, la calidad de los datos, determinar patrones de comportamiento, entre otros. Para evitar la duplicidad de la información se busca automatizar lo máximo posible el proceso recolección y tener la menor intervención de las personas en la asignación de la llave primaria y los diferentes medios de recolección existentes tales como el autocenso. Para ello tenemos una base de datos centralizada que gestiona esta llave primaria y gestiona por estados los diferentes resultados de las entrevistas, pero además durante el procesamiento de los datos se utilizan diferentes mecanismos de inconsistencias para detectar los datos duplicados que a pesar de los controles existentes aún se puedan pasar a los procesos siguientes.

4. Síntesis del aprendizaje sobre el método de captura.

4.1. Si usted tuviera la oportunidad de entregar tres recomendaciones sobre este método de captura a otro país, ¿qué elementos a partir de su experiencia deberían ser críticos para que el método funcione de forma adecuada?

Recomendación 1:

Establecer con antelación los requerimientos de los diferentes sistemas involucrados. El equipo técnico debe tener claridad de lo que requiere y una comunicación constante con el desarrollador para que los productos cumplan con las expectativas y necesidades.

Recomendación 2:

Una correcta y oportuna planificación, con investigación previa, para determinar los diferentes elementos que se deben considerar, especialmente en términos de la infraestructura tecnológica que deberá soportar todos los desarrollos.

Recomendación 3:

Comunicación asertiva y oportuna entre las diferentes partes involucradas, con el fin de que se tenga una sintonía, se informe adecuadamente de todas las actividades y se dé seguimiento para cumplir en tiempo y forma con los objetivos.

5. Planes para el próximo censo de población y vivienda.

5.1. Marque con una X el Tipo de Censo que espera realizar para la próxima ronda censal en su país:

De Hecho De Derecho

5.2. Indique la duración de la recolección censal que espera para su próximo censo en el país, en semanas.

semanas. Si su censo será de hecho anotar 0.

5.3. Indique los métodos de recolección o captura que espera utilizar en su país para el próximo censo con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas que espera cubrir con cada método.

		% de Viviendas	% de Población
<input type="checkbox"/>	Papel/Digitación		
<input type="checkbox"/>	Papel/Reconocimiento Óptico		
<input checked="" type="checkbox"/>	DMC / CAPI	90	90
<input type="checkbox"/>	Web / CAWI	10	10

5.4. Identificar con una X todos los tipos de cuestionario censal que va a considerar en el próximo censo en su país.

<input checked="" type="checkbox"/>	Viviendas particulares
<input checked="" type="checkbox"/>	Viviendas Colectivas
<input checked="" type="checkbox"/>	Personas en Situación de Calle
<input type="checkbox"/>	Personas en Tránsito o Viajeros
<input type="checkbox"/>	Otros

Si marca "otros", especificar en este recuadro:

5.5. Describa de forma breve los métodos a utilizar en el próximo censo, explicita de forma general las razones de mantener o cambiar los métodos utilizados. Además, incorpore los principales desafíos y avances en la preparación del censo en esta materia.

Se utilizará DMC y autocenso vía web. Las razones se indicaron anteriormente, pero se resumen a continuación:

- Posibilidad de mejorar la calidad de la información al poder validar algunos datos en tiempo real.
- Mayor oportunidad al brindar los resultados pues los datos se procesarán de forma paralela a la recolección.
- Posibilidad de mejorar el entendimiento de las personas censistas sobre la aplicación del cuestionario, al utilizar una aplicación digital de fácil uso.
- Posibilidad de tener un enlace con la información del cuestionario, directamente con un punto georreferenciado en el mapa.

Desafíos:

- La definición adecuada de los requerimientos tecnológicos.
- La elección adecuada de un aliado para el desarrollo de los sistemas.
- Identificar los posibles riesgos que conllevan un cambio en el método de captura.

Avances:

- Se cuenta con sistemas lo suficientemente avanzados como para ser llevados a prueba en un Censo Piloto.

5.6. Si tuviera la oportunidad de realizar tres preguntas a otros países de América Latina y el Caribe que hayan implementado el método captura, o los métodos de captura, que usted proyecta para el próximo censo, ¿cuáles serían?

Pregunta 1:

¿Cuánto tiempo antes inició la preparación en términos de establecer la infraestructura tecnológica y el desarrollo de sistemas?

Pregunta 2:

¿Cuáles fueron los principales problemas que se debieron solucionar durante el trabajo de campo y la transferencia de información? ¿cómo gestionaron los controles de calidad y cobertura?

Pregunta 3:

¿Cuáles son las lecciones aprendidas en la realización de un censo utilizando DMC?

Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe
Grupo de Trabajo Censos de la Conferencia Estadística de las Américas – CEPAL
Documento Lineamientos generales para la captura de datos censales

En el marco del plan de trabajo del **Grupo de censos de la CEA CEPAL** para el **bienio 2020-2021** se comprometió la realización de un documento que abordará lineamientos generales para la captura de datos censales para los países de América Latina y el Caribe. Uno de los objetivos específicos es la recopilación de experiencias por parte de los países respecto de la utilización de algunos de los métodos de recolección y captura de datos.

La presente ficha tiene por objetivo sistematizar la experiencia nacional respecto del método de captura de datos censales que ha sido solicitada a usted como representante de su país por parte del equipo de coordinación del Grupo de Trabajo de Censos de CEA Cepal.

La información aquí recopilada será utilizada como insumo para la redacción del documento, durante febrero del 2021 le enviaremos la versión preliminar del documento que será generado con su aporte para que sea previamente aprobada por la dirección de la oficina nacional de estadísticas de su país.

La ficha que se presenta consta de **5 partes**: Información de contacto; Características generales del último censo realizado en su país; Información del Método de Captura; Síntesis de aprendizajes; y Planes para el próximo censo.

Le pedimos enviar la ficha completa a más tardar el día **16 de octubre del 2020**. Su contenido será revisado por el equipo de coordinación del Grupo de Trabajo de Censos y de ser necesario le solicitará una reunión para profundizar en alguno de sus aspectos.

Agradecemos desde ya su colaboración en este proceso y ante cualquier duda o consulta por favor contactarse con Olga Barquero del INE-Chile: olga.barquero@ine.cl y Alejandra Silva de CELADE: alejandra.silva@cepal.org

1. Información de Contacto.

Por favor llenar con su información de contacto como responsable del llenado de la ficha por parte de su país.

Nombre	Willam Tirira
Correo electrónico	Willam tirira@inec.gob.ec / williamtirira@yahoo.es
Cargo	Responsable nacional de procesamiento CPV
Institución	Instituto Nacional de Estadística y Censos
País	Ecuador

Nombre	Mario Herrera
Correo electrónico	Mario_herrera@inec.gob.ec ; mherrera58@hotmail.com
Cargo	Analista de imputación y validación de información de los censos de población y vivienda
Institución	Instituto Nacional de Estadística y Censos
País	Ecuador

2. Características generales del **último censo** realizado en el país.

2.1. Marque con una X el Tipo de Censo para el último censo realizado en su país:

De Hecho De Derecho

2.2. Indique la fecha de inicio de la recolección del último censo en su país:

Día Mes Año

2.3. Indique la duración de la recolección del último censo en su país en semanas, se debe incluir el período oficial además de las semanas que se extendió la recolección más allá de ese período hasta el cierre del operativo censal.

a. Duración del período oficial: semanas. Si su censo fue de hecho anotar 0.

En caso de que quiera especificar condiciones especiales para el período oficial de recolección, como, por ejemplo, diferencia de duración entre urbano y rural, especifique en el cuadro siguiente:

En el área amanzanada se recolectó la información en un solo día, mientras en el área dispersa, el operativo de campo fue de una semana.

b. ¿Cuántas semanas se extendió la recolección después del período oficial?: semanas.

2.4. Indique los métodos de recolección o captura utilizados en el último censo en su país con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas cubiertas mediante ese método.

		% de Viviendas	% de Población
<input type="checkbox"/>	Papel/Digitación		
<input checked="" type="checkbox"/>	Papel/Reconocimiento Óptico	100%	100%
<input type="checkbox"/>	DMC / CAPI		
<input type="checkbox"/>	Web / CAWI		

2.5. Describa de forma general la estrategia operativa del censo. Incluya si los métodos de recolección o captura se aplicaron de forma secuencial o en paralelo; qué métodos de recolección o captura fueron utilizados para operativos especiales (viviendas colectivas, situación de calle, población en tránsito u otro); indicar si la captura de datos fue centralizada o descentralizada, o cualquier otra información relevante para el contexto de su experiencia.

Como estrategia de levantamiento, se planificaron y ejecutaron los siguientes operativos censales, en el Censo 2010:

Operativo de viviendas particulares en áreas amanzanadas: el levantamiento se realizó en un solo día a escala nacional en todas las áreas con características de amanzanamiento en el país, fueron censistas los estudiantes de bachillerato o de universidad, como jefes de sector (con una carga de 12 censistas aproximadamente) se designaron a los profesores de colegio o universidad, también se designaron Jefes de Zona (directores o personal de los colegios o universidad) que tuvieron a cargo alrededor de 10 Jefes de Sector.

Operativo de viviendas particulares en áreas dispersas: el operativo se planificó para una semana, en la que participaron como censistas los profesores de las escuelas ubicadas en el área rural, cada uno a cargo de un sector disperso (80 viviendas aproximadamente).

Operativo de viviendas colectivas y móviles: este operativo especial tuvo una duración de una semana, fueron censistas, delegados de dichas viviendas, debidamente capacitados, con el apoyo de personal INEC.

Operativo de personas sin vivienda: se censaron a las personas en situación de calle en un operativo especial realizado desde la noche anterior al día del censo y que duró generalmente algunas horas alrededor del momento censal (00H00 del día del censo). Lo realizó personal INEC.

Operativo en la frontera norte: por las dificultades de acceso y seguridad en esta área del país, se realizó el levantamiento con personal de las Fuerzas Armadas, la duración fue similar al área dispersa.

En todos los operativos se utilizaron cuestionarios físicos tipo cuadernillo. La responsabilidad de la captación de datos fue descentralizada, su ejecución recayó sobre las cuatro coordinaciones zonales a través de las cuales desconcentra la estructura institucional, cada una de ellas a cargo de un número determinado de provincias, cantones y parroquias rurales, que conforman la División Político Administrativa del país.

Al concluir el operativo, se recopilan los cuestionarios, agregados por área de empadronamiento (carga de trabajo del censista amanzanado), luego agregados por sector censal (a cargo del Jefe de Sector) y zona censal (a cargo del Jefe de Zona), el material se concentró en las ciudades o parroquias rurales, luego se lo trasladó a las capitales cantonales y finalmente a las capitales provinciales, desde allí se envió el material consolidado a las sedes zonales INEC, donde se conformó el archivo físico, para su entrega a la siguiente fase que es el procesamiento de datos.

3. Características del método: Escaneo in house.

3.2. Si usted tuviera que defender el método de **escaneo in house** frente a sus pares de otros países, ¿podría sintetizar el proceso de decisión para la elección de este método como parte del proceso de recolección censal?

EXPERIENCIAS 1990 Y 2001

Se realizó la comparabilidad entre tres tipos de métodos de captura: manual (**Censos 1990 y 2001**), escaneado (**Proyecto Registro Único de Beneficiarios 2018 - RUB**) y captura a través de dispositivos móviles (**Sin experiencia en proyectos censos**).

Captura manual: Susceptible a generar más errores en la digitación, requerimiento de amplio espacio físico, computadores de escritorio, contratación de personal con la dificultad de administrar el mismo

Dispositivos móviles: En un censo de hecho se descarta el uso de esa tecnología debido a costos muy altos en la adquisición de equipos (Uso por única vez) y tecnología adicional (especialmente comunicaciones)

DECISION ESCANEADO IN-HOUSE (2010)

Razones:

- **Experiencia del proyecto RUB**
 - *Equipo técnico del INEC con la experticia de escaneado y procesos posteriores*
 - *Conocimiento de la infraestructura tecnológica utilizada en el proyecto y proyección de lo requerido para el censo 2010*
 - *Conocimiento de la plataforma de software utilizado en el proyecto RUB y la posible utilización en el censo 2010*
- *Avance tecnológico en el procesamiento censal.*
- *Control del proceso por parte del equipo técnico del censo*

3.3. Indique las pruebas realizadas durante la preparación de la recolección con **escaneo in house**, incluir todas las pruebas tecnológicas, de cuestionario, operativas y censo experimental, donde se haya evaluado este método. Identificar los principales hallazgos en esas pruebas.

EXPERIENCIA 2010

Experiencia previa

El INEC tenía la experiencia previa de escaneo en la investigación del Registro Único de Beneficiarios (cuestionario de una sola hoja - 2 páginas)

Experiencia de pruebas tecnológicas y cuestionario

No se realizaron pruebas de escaneo en pruebas piloto.

Experiencia del censo experimental

En el censo experimental, se escaneo con un cuestionario totalmente diferente al definitivo tanto en el diseño (horizontal) y contenido (preguntas) y diferente tipo de papel; por lo tanto esta prueba no permitió obtener ninguna observación y/o recomendación.

El cuestionario para el censo 2010 fue aprobado en el mes de julio del año indicado, luego de pruebas de campo para comprobar el diseño (entre cuestionario horizontal vs vertical), definiendo al cuestionario vertical como definitivo, luego de esta prueba de campo se diseñó el cuestionario definitivo y pasó al proceso de impresión.

En el proceso de impresión se realizó un control de calidad específicamente sobre la lectura al código de barras, marcas de escáner y calidad de impresión.

3.4. Sintetice el desarrollo del método de reconocimiento óptico. Indique si el proceso se realizó in house o a través de un tercero; se solicita describir proceso de respaldos de imágenes, transferencia de datos, elementos de seguridad y protección de datos personales (encriptaciones, manejo de claves y accesos, etc.); Incluir consideraciones para la adquisición (licitaciones de compra, arriendo o provisión de servicios). Debe indicar qué softwares se utilizaron para el reconocimiento óptico y gestión del proceso de digitalización.

EXPERIENCIA CPV 2010

Indique si el proceso se realizó in house o a través de un tercero

El proceso se realizó en la modalidad in house, por tanto, el escaneo y actividades relacionadas (interpretación, verificación y transferencia) fue realizado a través de un software contratado para el efecto, el INEC instaló la infraestructura tecnológica (red de datos, servidores de datos e imágenes, computadoras de escritorio y escáneres), contrato el personal para todas las actividades indicadas y ejecuto el proceso con su propio personal (experiencia 2010).

Se solicita describir proceso de respaldos de imágenes, transferencia de datos

Al ser el proceso de forma descentralizada, el INEC instaló en cada una de las coordinaciones zonales servidores con espacio en disco para la grabación de las imágenes a través de unidades de cinta, el proceso de respaldo fue periódico, no se realizó la transferencia de datos a administración central en ningún momento.

Elementos de seguridad y protección de datos personales (encriptaciones, manejo de claves y accesos, etc.)

Toda la información se encontraba dentro del servidor, no se realizó ningún proceso de encriptación dado que la información se encontraba en una BDD, además de que no se investigó datos identificativos de la población, las claves de acceso eran de responsabilidad exclusiva de los administradores del sistema y del responsable zonal de procesamiento. Se configuro una red con bloqueo a internet, bloqueo a puertos de los computadores conectados en red y políticas de seguridad del sistema operativo (Active Directory).

Incluir consideraciones para la adquisición (licitaciones de compra, arriendo o provisión de servicios)

Previa a la adquisición de infraestructura tecnológica y servicios, el componente Procesamiento del Censo 2010 elaboró el documento Lineamientos de Procesamiento del Censo de Población y Vivienda 2010 (Presupuestos, cronogramas, especificaciones técnicas y consideraciones metodológicas, documento que fue aprobado por la Comisión Técnica del Censo.

El lineamiento aprobado era el documento habilitante para iniciar los procesos de adquisición de la infraestructura tecnológica requerida para el procesamiento.

Los servicios más relevantes contratados fueron: comunicaciones, alquiler de software y mantenimiento de infraestructura (guillotinas y escáneres), la adquisición de infraestructura tecnológica (servidores y computadores de escritorio) estaban bajo garantía del fabricante, este proceso de adquisición fue realizado sin novedad alguna.

Debe indicar qué softwares se utilizaron para el reconocimiento óptico y gestión del proceso de digitalización

Para la gestión del proceso de digitalización (administración de la documentación censal) se realizó un sistema que permitía la distribución y registro a nivel de sector censal en las fases de crítica-codificación, corte de cuestionarios y escaneo; con este sistema se controlaba la producción además de permitir obtener indicadores de avance de la misma.

El software fue alquilado a una empresa ecuatoriana con licencia de distribución, el nombre: Readsoft Forms Version 5.3, el alquiler fue de tipo modular (licencia por cada módulo): Manager, Scan, Interpret, Verify y Transfer) distribuidos en cada zonal de acuerdo al número de cuestionarios por cada coordinación zonal.

De la empresa proveedora se recibieron los siguientes servicios adicionales como parte del contrato: instalación y configuración de los módulos, capacitación al personal técnico, afinamiento previo al proceso y acompañamiento al inicio de la producción.

3.5. Identifique elementos de logística operativa asociados al reconocimiento óptico, por ejemplo, distribución de material para la recolección, retorno de material de recolección, manejo de contingencias en caso de pérdidas, destrucción o fallas en los materiales asociados al método de captura y cualquier otro elemento relevante en este sentido.

EXPERIENCIA CPV 2010

Identifique elementos de logística operativa asociados al reconocimiento óptico, por ejemplo, distribución de material para la recolección, retorno de material de recolección

Se planificó la identificación de los sectores censales (cajas) con etiquetas (código del sector y código de barras) tanto para la distribución para el empadronamiento y retorno del mismo, lamentablemente debido a problemas administrativos (adquisición errónea de etiquetas) y la fallida adquisición de lectores de código de barras (lectores muy básicos), por lo tanto la logística operativa asociada al reconocimiento óptico no fue implementada.

Manejo de contingencias en caso de pérdidas.

Se tuvo dos problemas en las pérdidas de material censal:

1. En una coordinación zonal, el alquiler continuo de bodegas causó la pérdida de varias viviendas colectivas
2. La falta de control de recepción de material en las sedes censales, causó la pérdida de 10 sectores censales (de un total de 42.000 sectores)

Para el primer caso no hubo solución y en el segundo se imputó con información de sectores aledaños.

Destrucción o fallas en los materiales asociados al método de captura

1. Cuestionarios censales deteriorados (agua y papel deteriorado): Para estos casos se transcribió la información a nuevos cuestionarios.
2. En infraestructura tecnológica se debe planificar y contratar todos los servicios y mantenimientos:
 - a. Guillotinas: Acometidas eléctricas a cargo de la empresa proveedora, adquisición de cuchillas adicionales, mantenimiento de las guillotinas y afilamiento (mensual) de cuchillas.
 - b. Escáneres: Mantenimiento periódico (semanal o quincenal), kits de suministros adquiridos a tiempo.

3.6. Describa la forma de monitoreo del funcionamiento del reconocimiento óptico. Indique como se realizó el monitoreo y control del scanner; la forma de control y monitoreo de los cuestionarios en papel para asegurar que todos los cuestionarios fueron escaneados. Qué validaciones se realizaron para evitar duplicados (incluir algunos criterios de identificación), pérdida de datos durante la captura, o datos erróneos por mal funcionamiento del método.

EXPERIENCIA 2010

Indique como se realizó el monitoreo y control del scanner; la forma de control y monitoreo de los cuestionarios en papel para asegurar que todos los cuestionarios fueron escaneados

El monitoreo de la fase de escaneo se realizó a través del propio software, unido a las características propias de un escáner de alta producción. Cada sector escaneado el software (módulo SCAN) entregaba un reporte de hojas escaneadas en función del número de hoja. En caso de que algún cuestionario esté incompleto, el proceso se detenía hasta solucionar el error.

Qué validaciones se realizaron para evitar duplicados (incluir algunos criterios de identificación)

En caso de duplicación en el escaneado de cuestionarios, se realizó una aplicación por el equipo de procesamiento que eliminaba el cuestionario en función del código de barras (código único por cuestionario).

Pérdida de datos durante la captura, o datos erróneos por mal funcionamiento del método

En el CPV se tuvo un grave problema por no incluir una fase de control de calidad después de la fase de interpretación, este problema obligó a reescanear al menos un 20% de sectores censales.

La elaboración de la plantilla (en el software) de la página 2 del cuestionario censal, no escaneaba las marcas de una columna de variables, la solución fue actualizar la plantilla, reescanear las imágenes almacenadas en los servidores, reinterpretar las marcas e insertar las respuestas en las variables erróneas en función del código de barras del cuestionario censal.

Con la modificación y revisión de las plantillas, el 80% de sectores censales fueron escaneados e interpretados sin ninguna novedad posterior.

4. Síntesis del aprendizaje sobre el método de captura.

4.1. Si usted tuviera la oportunidad de entregar tres recomendaciones sobre este método de captura a otro país, ¿qué elementos a partir de su experiencia deberían ser críticos para que el método funcione de forma adecuada?

Recomendación 1:

Pruebas Pilotos y Censo Experimental deben ser probados con todos los elementos metodológicos e infraestructura física y tecnológica diseñados para el censo definitivo.

Recomendación 2:

*Adquisición de equipamiento en función del rendimiento de las pruebas pilotos y censo experimental, **aquí se incluye el tipo de papel de acuerdo a las recomendaciones internacionales.***

Recomendación 3:

Implementar control de calidad y supervisión en cada una de las actividades del procesamiento censal, para detectar imprevistos a tiempo.

5. Planes para el próximo censo de población y vivienda.

5.1. Marque con una X el Tipo de Censo que espera realizar para la próxima ronda censal en su país:

De Hecho De Derecho

5.2. Indique la duración de la recolección censal que espera para su próximo censo en el país, en semanas.

semanas. Si su censo será de hecho anotar 0.

5.3. Indique los métodos de recolección o captura que espera utilizar en su país para el próximo censo con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas que espera cubrir con cada método.

		% de Viviendas	% de Población
<input type="checkbox"/>	Papel/Digitación		
<input checked="" type="checkbox"/>	Papel/Reconocimiento Óptico	100%	100%
<input type="checkbox"/>	DMC / CAPI		
<input type="checkbox"/>	Web / CAWI		

5.4. Identificar con una X todos los tipos de cuestionario censal que va a considerar en el próximo censo en su país.

- Viviendas particulares
- Viviendas Colectivas
- Personas en Situación de Calle
- Personas en Tránsito o Viajeros
- Otros

Si marca "otros", especificar en este recuadro:

Se utilizará un solo cuestionario para todos los operativos.

5.5. Describa de forma breve los métodos a utilizar en el próximo censo, explicita de forma general las razones de mantener o cambiar los métodos utilizados. Además, incorpore los principales desafíos y avances en la preparación del censo en esta materia.

Para el VIII Censo de Población y VII de Vivienda 2021, de Hecho, se utilizará como medio de captura el reconocimiento óptico de imágenes, marcas y caracteres. El proceso de reconocimiento óptico de los cuestionarios lo realizará el INEC a través de software que será contratado a un proveedor externo.

Como breve descripción de las fases que contempla el procesamiento tenemos:

Archivo inicial: Recepción de la información enviada luego del empadronamiento y ordenamiento y ubicación en las estanterías.

Pre-Critica: Separación de cuestionarios en blanco y otros materiales, revisión y corrección de la identificación geográfica de cada cuestionario, revisión y ordenamiento de los cuestionarios por área de empadronamiento y vivienda en orden ascendente, revisión de la estructura vivienda-hogar-población.

Guillotina: Corte del lote censal en la línea de corte (impreso en cada cuestionario censal).

Escaneo: Escaneo de cada lote censal (sector censal), no se deben escanear sectores unidos o parciales, verificación en el monitor del sistema que el escaneo sea completo, verificación que las imágenes de cada sector censal vayan a su correspondiente carpeta.

Interpretación: El software especializado interpreta el contenido de las imágenes y transforma en valores numéricos o alfanuméricos y almacena en una base de datos.

Verificación: Verificación de las variables interpretadas (números y caracteres alfanuméricos), corrección manual (teclado) entre la imagen de la variable y lo contenido en la base de datos.

Control de cobertura: Corrección de identificaciones de cuestionarios (ubicación geográfica errada o diferente a la base precensal, matching entre la base precensal y la base censal a niveles geográficos (zona, sector y áreas de empadronamiento), re escaneo de sectores con errores.

Codificación: Codificación automática de las variables con respuesta alfanumérica (geográficas y económicas), codificación asistida de las variables que no pudo realizar la codificación automática. Previamente se valida y se añaden variables de Registros Administrativos.

Validación e imputación: Consiste en obtener reportes de validación e imputar o asignar valores a variables con información inconsistente, nula o en blanco; hasta obtener una base de datos consistente e íntegra para obtener reportes de salida para uso de instituciones del país y usuarios en general. Imputación Estática (Hard Coded o Cold Deck) / Imputación dinámica (Hot Deck). Previamente se valida y se añaden variables de Registros Administrativos.

En el Censo experimental se probó la digitación de formularios existiendo muchos errores, entre más manual sea el proceso más riesgo existe de cometer errores. Se podría externalizar el servicio de digitalización, sin embargo, si existe algún problema en el proceso de adquisición, existe el alto riesgo de no presentar resultados del Censo. Lo mencionado y a esto aumentando la experiencia que tiene el INEC nos ha llevado a realizar el Censo con digitalización in house.

Desafíos y avances:

Implementar machine learning en la codificación

Contar con un sistema tracking, desarrollado en el INEC, de tal forma que se monitoree de forma permanente en donde se encuentra las cajas con los sectores censales y el estado del procesamiento de los sectores censales

NOTA:

5.6. Si tuviera la oportunidad de realizar tres preguntas a otros países de América Latina y el Caribe que hayan implementado el método captura, o los métodos de captura, que usted proyecta para el próximo censo, ¿cuáles serían?

Pregunta 1:

¿Qué beneficios frente a otros métodos de captura han experimentado con la utilización de escaneo in house?

Pregunta 2:

¿Qué tipo de motor de base de datos fue utilizada con el software especializado en captura inteligente de caracteres?

Pregunta 3:

¿Cuál fue la arquitectura para la implementación del software de reconocimiento óptico? (centralizado o descentralizado) ventajas y desventajas.

Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe
Grupo de Trabajo Censos de la Conferencia Estadística de las Américas – CEPAL
Documento Lineamientos generales para la captura de datos censales

En el marco del plan de trabajo del **Grupo de censos de la CEA CEPAL** para el **bienio 2020-2021** se comprometió la realización de un documento que abordara lineamientos generales para la captura de datos censales para los países de América Latina y el Caribe. Uno de los objetivos específicos es la recopilación de experiencias por parte de los países respecto de la utilización de algunos de los métodos de recolección y captura de datos.

La presente ficha tiene por objetivo sistematizar la experiencia nacional respecto del método de captura de datos censales que ha sido solicitada a usted como representante de su país por parte del equipo de coordinación del Grupo de Trabajo de Censos de CEA Cepal.

La información aquí recopilada será utilizada como insumo para la redacción del documento, durante febrero del 2021 le enviaremos la versión preliminar del documento que será generado con su aporte para que sea previamente aprobada por la dirección de la oficina nacional de estadísticas de su país.

La ficha que se presenta consta de **5 partes**: Información de contacto; Características generales del último censo realizado en su país; Información del Método de Captura; Síntesis de aprendizajes; y Planes para el próximo censo.

Le pedimos enviar la ficha completa a más tardar el día **16 de octubre del 2020**. Su contenido será revisado por el equipo de coordinación del Grupo de Trabajo de Censos y de ser necesario le solicitará una reunión para profundizar en alguno de sus aspectos.

Agradecemos desde ya su colaboración en este proceso y ante cualquier duda o consulta por favor contactarse con Olga Barquero del INE-Chile: olga.barquero@ine.cl y Alejandra Silva de CELADE: alejandra.silva@cepal.org

1. Información de Contacto.

Por favor llenar con su información de contacto como responsable del llenado de la ficha por parte de su país.

Nombre	Edgar Vielma Orozco
Correo electrónico	Edgar.vielma@inegi.org.mx
Cargo	Director General de Estadísticas Sociodemográficas
Institución	Instituto Nacional de Estadística y Geografía, INEGI
País	México
Nombre	Sabino Navarro Campos
Correo electrónico	sabino.navarro@inegi.org.mx

Cargo	Subdirector de Desarrollo de Sistemas en la Dirección General de Estadísticas Sociodemográficas
Institución	Instituto Nacional de Estadística y Geografía, INEGI
País	México

2. Características generales del **último censo** realizado en el país.

2.1. Marque con una X el Tipo de Censo para el último censo realizado en su país:

De Hecho De Derecho

2.2. Indique la fecha de inicio de la recolección del último censo en su país:

Día Mes Año

2.3. Indique la duración de la recolección del último censo en su país en semanas, se debe incluir el período oficial además de las semanas que se extendió la recolección más allá de ese período hasta el cierre del operativo censal.

a. Duración del período oficial: semanas. Si su censo fue de hecho anotar 0.

En caso de que quiera especificar condiciones especiales para el período oficial de recolección, como, por ejemplo, diferencia de duración entre urbano y rural, especifique en el cuadro siguiente:

Al cierre del operativo de enumeración se habían censado el 95% de las viviendas particulares habitadas identificadas. No obstante, debido a la contingencia sanitaria derivada de la pandemia por COVID-19, el 31 de marzo la Secretaría de Salud emitió un Acuerdo por el que se suspendió toda actividad no esencial incluyendo el levantamiento de censos y encuestas mediante entrevistas cara a cara; por lo que la etapa de verificación, que involucró la verificación de viviendas deshabitadas, así como la recuperación de pendientes y viviendas omitidas, se suspendió temporalmente, mientras que se amplió el periodo para que la población respondiera el cuestionario en línea por el método de Autoenumeración (Internet).

El 17 de julio, la Secretaría de Salud emitió el Acuerdo por el que se reanudan los Censos y Encuestas que involucran la movilización de personas y su interacción física, sujeto a un protocolo de actuación elaborado por el INEGI y aprobado por la misma Secretaría. Con base en esas directrices, las actividades de la fase de verificación se reiniciaron, pero se agregaron al instrumento de captación preguntas para identificar a los residentes de la vivienda en marzo de 2020.

La encuesta de Posenumeración se canceló debido a la distancia temporal de su aplicación, respecto de la etapa de Enumeración, sin embargo, fue sustituida por la Encuesta de Muestreo para la Estimación Focalizada de Cobertura (ENCOBER).

Por su parte, el calendario de difusión de resultados se modificó, por lo que los resultados básicos se publicarán este mismo año dependiendo de la evolución de la pandemia.

b. ¿Cuántas semanas se extendió la recolección después del período oficial?: Para la Enumeración no se extendió ningún día, sin embargo, para la etapa de Verificación que había iniciado el 23 de marzo, y fue suspendida el día 31 del mismo mes en su etapa de levantamiento de campo, esta se reanudó a partir del 8 de junio de manera escalonada en cada coordinación estatal, con base en el semáforo de riesgo epidemiológico COVID-19, y de manera generalizada en todo el territorio nacional, a partir del 17 de julio, de conformidad con el Acuerdo emitido por la Secretaría de Salud.

2.4. Indique los métodos de recolección o captura utilizados en el último censo en su país con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas cubiertas mediante ese método.

		% de Viviendas	% de Población
<input checked="" type="checkbox"/>	Papel/Digitación	2.00	2.00
<input type="checkbox"/>	Papel/Reconocimiento Óptico		
<input checked="" type="checkbox"/>	DMC / CAPI	97.7	97.7
<input checked="" type="checkbox"/>	Web / CAWI	0.30	0.30

2.5. Describa de forma general la estrategia operativa del censo. Incluya si los métodos de recolección o captura se aplicaron de forma secuencial o en paralelo; qué métodos de recolección o captura fueron utilizados para operativos especiales (viviendas colectivas, situación de calle, población en tránsito u otro); indicar si la captura de datos fue centralizada o descentralizada, o cualquier otra información relevante para el contexto de su experiencia.

La recolección de la información del Censo cubrió el conjunto de actividades para obtener los datos de cada unidad de observación, de acuerdo con el programa y los procedimientos de trabajo establecidos, incluyendo una estructura operativa y controles que promovieron la eficacia en cada una de las acciones. Esto implicó ejecutar los esquemas para la captación de datos diseñados previamente, como son: la preparación y distribución de materiales de apoyo, la integración de recursos humanos, la comunicación y concertación, bajo una detallada programación de actividades, una estructura orgánica y controles que se monitorearon continuamente desde la fase de planeación hasta el cierre del levantamiento.

La estrategia operativa contempló lo siguiente:

1. El **Censo**: cuyo periodo de levantamiento fue del 2 al 27 de marzo de 2020, integró la enumeración exhaustiva de viviendas y los operativos especiales para captar la información de la población residente en alojamientos colectivos, en el Servicio Exterior Mexicano y de la población sin vivienda (en situación de calle).
2. La **Verificación de cobertura**: con un periodo del 23 de marzo al 10 de abril de 2020, mismo que fue postergado debido la pandemia del Covid-19, y retomado hasta después del 17 de julio, fecha en que la Secretaría de Salud emitió el acuerdo mediante el que se permitió la reanudación de Censos y Encuestas.
3. La encuesta de **Posenumeración**: que se realizaría del 27 de abril al 8 de mayo de 2020, y que fue cancelada debido a la pandemia por Covid-19. Esta fue sustituida por la Encuesta de Muestreo para la Estimación Focalizada de Cobertura (ENCOBER).

El Censo 2020 incluyó tres métodos de recolección de datos:

1. La *entrevista directa (DCM y papel)*: fue el principal método utilizado, se realizaba utilizando dispositivos de cómputo móvil en la mayoría de los casos, los cuestionarios en papel se utilizaron en las áreas de alto riesgo o debido a alguna incidencia con el dispositivo.
2. La *Autoenumeración*: se realizó mediante la implementación de una página web para que el ciudadano, a partir de una carta invitación entregada en su domicilio por el entrevistador, proporcionara los datos de las características de su vivienda y sus residentes.
3. La *entrevista asistida por teléfono*: el ciudadano marcaba al teléfono señalado en la carta invitación para proporcionar los datos de su vivienda y los habitantes.

Para la **Autoenumeración** y la **Entrevista asistida por teléfono** se generaron cartas invitación que contenían las credenciales únicas para responder el cuestionario básico vía Internet, y fue entregada por el entrevistador, bajo demanda de los informantes, o si en la tercera visita a la vivienda no se logró realizar la entrevista cara a cara; durante el recorrido de levantamiento el entrevistador iba llenando el *Listado de inmuebles*, procedimiento que incluyó la lectura en campo del código QR de las etiquetas que pegaba en los inmuebles, así como el de las invitaciones, de esta manera la vivienda quedaba asociada a estos datos, lo que permitió el control de la captación de información en el sitio Web.

El sistema *Administrador Censal* permitió aplicar los **cuestionarios** durante la entrevista directa a los informantes e incluía las siguientes funciones:

- Visualizar y administrar las cargas de trabajo de cada figura operativa.
- El Módulo de Captura para el registro de cada inmueble (Listado de Inmuebles).
- Los cuestionarios básico y ampliado para las viviendas particulares habitadas.
- Los cuestionarios para el registro de información de Entorno Urbano y Localidad.

Para censar a la población residente en viviendas colectivas (cárceles, asilos, centros de acogida, etcétera), la estructura de operativos especiales utilizó la aplicación para la captura de ésta información en un dispositivo móvil tipo *Netbook*.

En el caso de la población del Servicio Exterior Mexicano, se reutilizó la herramienta web de Autoenumeración, omitiendo la sección de vivienda y realizando ajustes a las preguntas de *Residencia hace 5 años* y *Lugar de nacimiento*.

La población sin vivienda (en situación de calle) se captó integrando previamente un directorio de puntos de encuentro e implementando, en el dispositivo de cómputo móvil, un módulo para el registro de los códigos QR de los cuestionarios aplicados y el total de personas (hombres y mujeres) captados en el punto de encuentro. El levantamiento de esta población se realizó en un solo día, el 11 de marzo.

Al igual que con la población en situación de calle, para la captura de las personalidades, también se elaboró un directorio previamente, a dicha población se le aplicó el cuestionario básico utilizando el DCM y se adecuó el módulo de captura del *Administrador Censal*.

Se contó con instrumentos de captación impresos, para ser utilizados en caso de que la aplicación llegara a tener problemas durante la entrevista o por falta de energía eléctrica, así como en aquellas áreas de trabajo inseguras debido a la delincuencia, donde los entrevistadores corrieran el riesgo de ser asaltados o agredidos. Cada instrumento tuvo un código de respuesta rápida (QR) único, el cual los hizo irrepetibles, esto sirvió para controlar su distribución entre las figuras responsables del levantamiento de información.

Todos los cuestionarios levantados en papel fueron capturados de manera descentralizada, para ello, se implementaron 69 centros de captura, los cuales fueron coordinados por un responsable de captura estatal.

Para la Autoenumeración, la captación de información se realizó a partir de la entrega de una carta invitación por parte del entrevistador a los informantes, la cual incluía una clave y una contraseña, requeridos para ingresar a la aplicación Web implementada para esta estrategia. Con la captura de estos datos, iniciaba un proceso de registro, en que se le solicitó un correo electrónico personal, medio por el cual se le proporcionaron nuevas credenciales, indispensables para ingresar y proporcionar los datos de su vivienda así como la lista y datos de sus habitantes.

En cuanto a la Entrevista asistida por teléfono, en la carta-invitación venían los números a los que el informante podía llamar y un operador del Call Center del Instituto lo atendía. Para esta estrategia se hicieron adecuaciones a la aplicación de Autoenumeración. Es importante mencionar que, en caso de que el informante no contara con la clave y contraseña de la carta-invitación, podía proporcionar el dato del QR de la etiqueta pegada por el entrevistador en la fachada de su vivienda.

Para estas estrategias, la captura de información se realizó en paralelo, ya que gran cantidad de informantes pudieron acceder simultáneamente. Es importante mencionar que las respuestas a los cuestionarios se almacenaban de forma dinámica, evitando la pérdida de información, y permitiendo al informante proporcionar sus datos en varias sesiones.

3. Características del método: **Dispositivo Móvil de Captura.**

3.2. Si usted tuviera que defender el uso de dispositivos móviles de captura frente a sus pares de otros países, ¿podría sintetizar el proceso de decisión para la elección de este método como parte del proceso de recolección censal?

El uso de dispositivos de cómputo móvil (DCM) para la captación de los datos censales, forma parte de las mejores prácticas internacionales y permitió que el Censo de Población y Vivienda 2020 estuviera a la vanguardia al determinar su uso como estrategia principal para la etapa de levantamiento.

Mediante el uso de DCM se busca establecer un riguroso control operativo, agilizar el desarrollo de la entrevista, eliminar la captura de datos captados en papel y mejorar los procesos de codificación, validación y actualización cartográfica, contribuyendo a la generación oportuna y de calidad de los resultados del proyecto censal.

A continuación, se mencionan algunos aspectos por los que se promueve el uso de dispositivos móviles en proyectos censales.

El cuestionario. En el esquema tradicional el uso de papel implica gastos de impresión, transporte y almacenamiento de los instrumentos de captación, a los que se suman costos similares pero correspondientes al uso de las formas de control, seguimiento y verificación del levantamiento de la información. Por su parte, los dispositivos móviles permiten el uso de cuestionarios digitales y facilitan la automatización de las tareas de seguimiento y control.

Si bien el uso de papel no se elimina en su totalidad, ya que el entrevistador deberá contar con cierta cantidad de cuestionarios impresos para levantar la información en los casos que no sea posible hacerlo por medio del dispositivo, sí se minimiza en gran medida la impresión de materiales.

Identificación de las unidades de investigación (áreas geográficas, viviendas, personas). El registro en papel de ciertos datos implica un índice de error humano superior al que ocurre en dispositivos móviles, ya que en ellos mucha de la información requerida se encuentra precargada, principalmente claves geográficas y nombres de las personas para las preguntas correspondientes a diferentes temas y cortes de edad; además, la inclusión en la programación del cuestionario electrónico de las instrucciones de llenado como pases y filtros, así como esquemas de validación primarios, facilitan el control de los datos conforme se van captando, permitiendo verificarlos directamente con el informante.

También, mediante el uso del **GPS** del dispositivo, se tiene una mayor precisión en la ubicación de las áreas de trabajo a cubrir por el personal operativo, para ello, en el Censo 2020, se incluyó un mapa de acceso a las áreas de trabajo, en el cual se marcaba la ubicación del entrevistador durante su recorrido. Además, de manera automática, se registraron las coordenadas GPS para cada inmueble, tanto al principio como al final de la entrevista.

Preguntas abiertas. Los dispositivos móviles proveen a los cuestionarios digitales el uso de teclado y diccionarios que minimizan los errores de escritura en algunas preguntas abiertas (variables literales) o clasificadores requeridos por ciertos temas como religión, lengua indígena, entre otros. Sin embargo, con el fin de evitar sesgo en la captación de información, esta funcionalidad no se implementó en el Censo 2020, obligando al entrevistador a realizar la captura de los datos que proporcionó el informante.

Recolección de la información. En un esquema tradicional el tiempo para la recolección de la información, hasta su procesamiento y publicación implica la captación, almacenamiento, transporte de los cuestionarios e implementación de sistemas de captura, entre otros. En cambio, con el uso de dispositivos móviles, la captura de la información puede incluso llevar implícita su pre codificación y pre validación, reduciendo los casos que haya que corregir en el procesamiento central.

Confidencialidad de la información. El desarrollo de aplicaciones de software para DCM exige la implementación de esquemas de seguridad tanto en la aplicación local, como en el almacenamiento de datos locales, los mecanismos de envío e integración a la base de datos central, así como la pérdida o robo del DCM.

Para el Censo 2020, se implementaron técnicas de encriptación y cifrado, tanto para el acceso al DCM, como a nivel de base de datos. Lo mismo se hizo para el proceso de transferencia de archivos entre el entrevistador y el supervisor, y durante su envío e integración a la base de datos central.

Oportunidad. El uso de DCM, ahorra tiempo en la integración de la información en oficinas centrales, permite ser más oportunos en la toma de decisiones que garanticen una mejor cobertura de las unidades de observación, control del avance, actividades de supervisión, verificación y posenumeración. Los procesos de codificación, validación y actualización cartográfica, para el Censo 2020, comenzaron a partir de la tercera semana de levantamiento, lo que contribuye a una más oportuna generación y publicación de resultados.

Comunicación de Estructura Operativa. El uso de DCM, facilita la comunicación entre la estructura operativa, así como el intercambio de datos, video e imágenes que facilitan el desarrollo del trabajo de campo.

Características técnicas y especificaciones de los DCM. Como ya se mencionó, el **GPS** del dispositivo permite, además de lo ya mencionado, saber la ubicación en campo del personal, mediante el uso de herramientas para la gestión de equipo de cómputo móvil (MDM).

La **cámara digital** integrada, facilitó la implementación de los códigos de respuesta rápida (QR) en los instrumentos de papel, etiquetas y cartas-invitación para la Autoenumeración, además, ayudó a gestionar su distribución y recopilación entre las figuras operativas. Por otro lado, este componente permite la captura de imágenes y grabado de video, en caso de ser necesario.

También, se pueden utilizar herramientas para la lectura de preguntas (**audio**) y el registro de respuestas (**micrófono**), cuando se tienen que aplicar preguntas que sean sensibles para la población.

Finalmente, el uso de **datos móviles** del dispositivo permite la transmisión automática e inmediata de los datos censales a las bases de datos centrales, esta estrategia se implementó en el Censo 2020 para la transmisión de las entrevistas completas, asegurando así que, en caso de daño de la base de datos local, así como pérdida o robo del DCM, estos datos ya se encontraran resguardados.

Costo/beneficio del Dispositivo Móvil (DCM). En el documento *Análisis comparativo de medios de levantamiento de información. Papel vs Dispositivo de Cómputo Móvil*, se concluyó que el proyecto con dispositivos de cómputo móvil (DCM) sería financieramente viable si el precio de los dispositivos móviles era de \$6,577.50 pesos mexicanos, monto máximo para el que se estimó que el costo del proyecto censal sería igual al de un censo tradicional en papel. Es decir, la adquisición (o arrendamiento) de los DCM a un precio inferior al mencionado traería consigo ahorros, haciendo valer los principios de eficiencia y eficacia en el uso de los recursos públicos.

Como resultado de la presupuestación realizada para analizar la viabilidad de cada método de captación (con papel o con DCM), se obtuvieron los siguientes costos estimados:

Costos totales por tipo de escenario

CAPÍTULO DE GASTO	Escenario con PAPEL	Escenario con DCM	DIFERENCIA
1000 Servicios Personales	\$ 933,846,566	\$ 166,369,720	\$ 767,476,846
2000 Materiales y suministros	\$ 14,162,948	\$ 26,812,715	-\$ 12,649,767
3000 Servicios generales	\$ 275,942,949	\$ 47,748,462	\$ 228,194,487
5000 Equipo y mobiliario	\$ 293,152,706	\$ 53,916,621	\$ 239,236,085
TOTAL	\$ 1,517,105,168	\$ 294,847,518	\$ 1,222,257,651

El costo final de los DCM fue de \$322,687,092.48 MN (16,921,189.9 USD)* por un total de 185,824 unidades que se adquirieron, de manera que **el ahorro monetario del Censo por DCM** comparado con el Censo en Papel fue de **\$899,570,558 MN (47,172,027.2 USD)**.

Moneda	Escenario Papel	Escenario DCM	Diferencia/ahorro
MNX	1,517,105,168	617,534,610	899,570,558
USD	795,545,44.7	32,382,517.6	47,172,027.2

* Tipo de cambio al 16 de julio de 2019, fecha de la firma del contrato de adjudicación.

3.3. Indique las pruebas realizadas durante la preparación de la recolección con dispositivos móviles de captura, incluir todas las pruebas tecnológicas, de cuestionario, operativas y censo experimental, donde se haya evaluado este método. Identificar los principales hallazgos en esas pruebas.

Para el Censo 2020, dos de las iniciativas más importantes para el levantamiento fueron la recolección de la información mediante el uso de dispositivos de cómputo móvil (DCM) para la *Entrevista cara a cara*, en tanto que, con la Autoenumeración, se impulsó la captación de información a través del uso del internet.

El uso del DCM fue el principal método para la recolección de información, su utilización buscó establecer un riguroso control operativo, agilizar el desarrollo de la entrevista, eliminar, en la medida posible, la captura de datos y agilizar los procesos de codificación, validación y actualización cartográfica para la producción de resultados. Se planteó la realización de algunos ejercicios para definir las estrategias del operativo de campo y de capacitación, la logística operativa, así como determinar las características mínimas requeridas de los DCM.

En 2017 se realizó el primer ejercicio, llamado **Prueba de Estrategia operativa y dispositivos de cómputo móvil**. En materia informática, el principal objetivo fue determinar las características del dispositivo móvil a utilizar y licitar para la recolección de información del Censo. En dicho ejercicio se propuso la compra o préstamo de al menos 10 dispositivos móviles de características y marcas diferentes, y contar con al menos 5 unidades de cada uno, siendo este el número de sedes en donde se realizó dicha prueba. En este sentido, se buscó que los entrevistadores fueran personas mayores de edad, sin experiencia previa en operativos censales, pero familiarizados con el manejo de tablets o teléfonos celulares.

Otros objetivos fueron: definir y probar la estrategia operativa considerando el uso de DCM; probar el desarrollo, en diferente marco tecnológico, de las herramientas informáticas, y definir su operatividad en el Censo; definir y probar los esquemas de seguridad para salvaguardar la información recolectada en campo durante el periodo de producción de las herramientas; verificar las estrategias de integración de información censal a la base de datos central; y, la retroalimentación, a la estructura operativa y de dirección, mediante informes de seguimiento, avance y control del operativo.

Como resultado de este primer ejercicio, se consideró, para el desarrollo de las herramientas informáticas, la definición de una plataforma tecnológica que proporcionara esquemas orientados al desarrollo de aplicaciones móviles y la compra de DCM con sistema operativo Android, de tal forma que, se optimizaron los sistemas para la capacitación del personal, los procesos de intercambio de información entre los diferentes operativos, el módulo de captura del *Administrador Censal*, los mecanismos de sincronización para la transferencia de datos, los esquemas de seguridad desde el acceso a las herramientas informáticas, el cifrado de la información en el DCM y durante su transferencia, hasta su integración en la base de datos centralizada; también se mejoraron los módulos para la generación de informes de avance y control, se resolvieron los problemas relacionados con el uso de la cámara digital para la lectura de códigos QR y el uso del GPS para su registro en el listado de inmuebles.

El segundo ejercicio fue la **Prueba Piloto del Censo de Población y Vivienda 2018**. Para realizarla se eligieron dos ciudades, teniendo en cuenta los flujos migratorios, la presencia de alojamientos de asistencia social y que presentaran condiciones representativas de la zona rural y urbana del país; se consideró además el clima de ambos lugares, aspecto importante para poner a prueba el desempeño de los dispositivos móviles, así como el funcionamiento de las aplicaciones informáticas durante un periodo constante de uso. Esta prueba se proyectó para tener características similares al Censo, con una muestra de más de 44 mil viviendas y una duración de 20 días hábiles; las oficinas instaladas correspondieron a una coordinación municipal, y se realizaron todos los operativos planeados, incluyendo la Posenumeración y el operativo especial para la enumeración de viviendas colectivas, mismos que no se probaron en la prueba del 2017.

También se probó el proceso de licitación para la compra de los DCM requeridos para el ejercicio, con la intención de determinar aspectos importantes en el proceso de adquisición para el Censo 2020. Sólo se adquirió un tipo de marca y modelo que cumpliera con las características mínimas y especificaciones técnicas, definidas a partir de los resultados del 2017.

Los resultados de la **Prueba de Estrategia Operativa** y de la **Prueba Piloto del Censo**, permitieron establecer las características de los DCM a adquirir para el Censo, dando lugar a lo siguiente:

Características específicas.

- El tamaño de la pantalla entre 5.7 a 7.0 pulgadas con tecnología IPS.
- Sensor GPS interconstruido que permita mostrar el geoposicionamiento global del dispositivo.
- Soporte para micro USB o USB interconstruido en el dispositivo, así como cualquier accesorio necesario para conectar una tarjeta de memoria USB.

Características mínimas.

- Pantalla de 1280 x 700 pixeles.
- Procesador de 2 núcleos.
- Sistema operativo Android versión estable 6.0 o superior.
- 2 Gb memoria RAM.
- Almacenamiento interno de 32 Gb.
- Batería, que soporte 8 horas de uso continuo de trabajo en el equipo.
- Cámara posterior de 2 mega pixeles.
- Cargador de batería incluido con cable de carga y adaptador para corriente eléctrica.
- Antena para comunicación móvil de voz y datos interconstruida UMTS.
- Ranura para tarjeta SIM interconstruida en el dispositivo.
- Nivel de luminosidad de la pantalla 400 nits o su equivalente en lúmenes o candelas.
- Soporte para Wifi interconstruido en el dispositivo con que cumpla al menos con el estándar 802.11 b/g/n

Otra contribución de la Prueba Piloto fue que se determinó integrar pruebas de desempeño de los DCM para el proceso de licitación 2019, con el fin de hacer una verificación física y técnica del equipo, de su desempeño, garantizar el correcto funcionamiento para la transferencia de datos a través del puerto micro USB o USB, así como el rendimiento de la batería.

Posteriormente se realizó el **Ensayo Censal 2019**, con el objetivo de probar los procedimientos para la ejecución de cada una de las etapas y procesos, los cuestionarios y las soluciones tecnológicas y de comunicaciones que permitieran realizar los últimos ajustes a la infraestructura de TIC y herramientas informáticas, previo a la realización del Censo de Población y Vivienda 2020.

Entre los procesos puestos a prueba en dicho Ensayo, se encuentran: la captación de datos de todos los operativos mediante los diferentes instrumentos implementados, tanto electrónicos como impresos; la estrategia de Autoenumeración vía Internet y su interrelación con las herramientas de captación de los dispositivos móviles; los procedimientos para el registro de actualizaciones cartográficas en formato digital; el intercambio de información entre figuras operativas y la integración de datos censales a la base de datos central; verificación del directorio de viviendas colectivas; el proceso de seguimiento logístico para el control de la distribución de materiales; así como el seguimiento y control al levantamiento de información.

A finales del 2019, también se realizó **La Prueba Estadística**, en la que se planteó realizar un levantamiento de información con los cuestionarios del **Censo de Población y Vivienda 2010**, con la finalidad de obtener información que permitiera estimar las diferencias en los indicadores que se deriven de los cambios de redacción o estructuración de las preguntas en los cuestionarios básico y ampliado del **Censo de Población y Vivienda 2020**, principalmente por la captación en papel del primero, y en DCM en el segundo.

Los resultados del Ensayo censal y de la Prueba estadística, permitieron realizar los últimos ajustes a la estrategia operativa y a las herramientas para la captación de información.

Adicionalmente es importante mencionar que, durante el año 2019 y principios del 2020, en materia informática se realizaron **pruebas de volumen y concurrencia**, a fin de garantizar la disponibilidad de las aplicaciones WEB, así como el desempeño de las bases de datos, tanto la del levantamiento mediante el uso de dispositivos móviles, como la correspondiente a la Autoenumeración. Se contrató a una empresa externa para la aplicación de *ethical hacking* en las aplicaciones, con el objetivo de detectar y eliminar riesgos de ataque externo durante la etapa de producción, etapa en la cual la Coordinación General de Informática del INEGI fue responsable del **monitoreo de seguridad y desempeño de la capa de servicios** implementada para el proyecto censal.

3.4. Sintetice el desarrollo TI de la recolección con dispositivos móviles de captura. Indique si el software de recolección fue desarrollado in house o se utilizó un software externo. Se solicita describir procesos de respaldo, transferencia de datos, elementos de seguridad y protección de datos personales (encriptaciones, manejo de claves y accesos, etc.); Incluir consideraciones para la adquisición (licitaciones de compra, arriendo o provisión de servicios).

Durante la etapa de levantamiento de datos censales, se usó el **Administrador Censal**, herramienta informática desarrollada *in house*, la cual incluye el Módulo de Captura para el registro de cada inmueble (Listado de Inmuebles), el Cuestionario de Entorno Urbano y el Cuestionario de Localidad, además de los cuestionarios básico y ampliado para las viviendas particulares habitadas.

También se hizo uso del **Módulo Cartográfico Censal**, de desarrollo institucional (*in house*), cuya finalidad es registrar las actualizaciones cartográficas a nivel de manzana, localidad y vialidad, permitiendo digitalizar los polígonos de nuevas áreas, dar de baja las que ya no se encuentren en campo y representar gráficamente cambios como fusiones y divisiones de manzanas, permitiendo actualizar de manera inmediata la cartografía.

Otra herramienta desarrollada fue el **Kiosco INEGI**, que limita la función general del dispositivo a un entorno empresarial, restringiendo a la utilización exclusiva de las aplicaciones desarrolladas para el Censo, habilitando o deshabilitando, por políticas de seguridad, el manejo del resto de los recursos. Mediante el uso de credenciales de usuario y un patrón de habilitación del DCM, el dispositivo se iniciaba en la aplicación para Entrevistador o Supervisor de entrevistadores, manteniendo deshabilitado el acceso a la configuración, instalación de aplicaciones, así como la depuración por USB y la conexión del dispositivo a cualquier PC, entre otros. Esta herramienta también permitió generar respaldos constantes de la base de datos, así como generar los archivos para la integración de información a la base de datos central.

La transferencia de datos tuvo dos vías:

- Envío de entrevistas completas directamente desde el DCM mediante el uso de datos móviles.
- Envío de todos los inmuebles mediante el uso de memorias USB, donde el entrevistador enviaba la información diariamente a su supervisor, el cual mediante el módulo de integración del OPERA transfería, al menos una vez al día, la información de todos sus entrevistadores vía internet.

En términos de seguridad, para el almacenamiento local en los DCM, se empleó el formato SQLite, utilizando dos métodos: encriptación y cifrado de datos. La encriptación de la base de datos local se implementó mediante el algoritmo AES256, y para el cifrado de la base de datos de los envíos de información del supervisor, se emplearon cadenas Hash de 32 caracteres, lo que minimizó la extracción y explotación de los datos en caso de robo de información por hackeo.

Además, para el envío de información a la base de datos central se implementó un mecanismo de empaquetado envolvente de archivos, de esta manera la base de datos cifrada se empaquetaba en un primer archivo con una contraseña para posteriormente empaquetar el archivo resultante empleando otra, ambas contraseñas de longitud larga, entre 12 y 32 caracteres, obligando a que fueran necesarias muchas combinaciones para romper cada una de las llaves, necesitando gran poder de cómputo para lograrlo.

También se implementó un mecanismo de validación de archivos con un código de redundancia cíclica (CRC) generado mediante el algoritmo MD5, método estándar en la criptografía actual que permite verificar que los archivos no hubieran sido manipulados o corrompidos.

Tanto para la plataforma OPERA como para la Autoenumeración, se adquirió un certificado SSL (acrónimo de Secure Sockets Layer o capa de sockets seguros), que es la tecnología estándar para mantener segura una conexión a Internet. Para todos los sitios con un certificado SSL, la URL comienza con **https**: indicando a los usuarios que sus datos se encuentran protegidos.

Para el cierre del operativo censal se implementaron en el *Administrador Censal* y en el **Kiosco INEGI**, herramientas para la generación de un respaldo de información (*backup*) y de un paquete de datos finales o de cierre. Ambos archivos se enviaron a oficinas centrales, y se le dio seguimiento mediante reportes implementados en el OPERA, siguiendo una estrategia similar a la etapa de levantamiento.

Esta versión de cierre, también integró una herramienta de confronta de cifras entre supervisor y entrevistador, así como de verificación de las cifras del dispositivo del supervisor con las del OPERA. Al término de estas tareas, se debía deshabilitar el DCM para evitar que se siguiera haciendo uso de él, mediante un **botón de bloqueo**.

Por otro lado, tanto para la **Autoenumeración** como para la entrevista asistida por teléfono, se implementó un sitio web, y se desarrollaron aplicaciones responsivas con la plataforma de desarrollo de software .NET web pages, en la cual se programaron ciertas validaciones primarias, tanto del lado del cliente como en el servidor, para contribuir a la calidad e integridad de los datos proporcionados por el informante. Para el almacenamiento de la información, al igual que para el sistema OPERA, se diseñó e implementó un esquema de base de datos en Oracle 12c.

No se omite mencionar que fue indispensable el desarrollo de un **módulo de gestión de DCM (MDM)**, implementado en la plataforma OPERA, cuyo fin fue controlar las versiones de las aplicaciones móviles instaladas en los dispositivos para asegurar que cada equipo tuviera la versión correcta. Las versiones se enviaron una vez al día, así como la clave operativa de la figura operativa a la que se le asignó el equipo y, cada determinado periodo de tiempo, las coordenadas GPS de su ubicación. Con ello, se pudo realizar el seguimiento y control del equipo en campo.

3.5. Identifique elementos de logística operativa asociados a los dispositivos móviles de captura, por ejemplo, distribución de material para la recolección, retorno de material de recolección, manejo de contingencias en caso de pérdidas, destrucción o fallas en los materiales y cualquier otro elemento relevante en este sentido.

La distribución de dispositivos móviles se realizó mediante las figuras de soporte informático y auxiliares administrativos, desde las oficinas estatales a las municipales, las cuales contaron con servicio de voz y datos, además de ser seguras para el resguardo de los dispositivos. Para la entrega y la recepción de los dispositivos, el área administrativa implementó formatos de control impresos para su resguardo personalizado.

En cada una de las oficinas municipales se contó con la figura operativa de soporte informático municipal (SIM), responsable de proporcionar el servicio de soporte técnico a los entrevistadores y supervisores durante el levantamiento de información, y de apoyar en el seguimiento a incidencias reportadas con los DCM.

También se contó con una mesa de ayuda dirigida centralmente, la cual recibió y atendió toda la problemática operativa durante el levantamiento de información.

Al presentarse una pérdida o robo del dispositivo, la figura operativa levantó una denuncia ante las autoridades judiciales para entregar en el área administrativa el acta de hechos y documentar el caso. En estos casos, el entrevistador o supervisor recibió otro dispositivo o, si el área era insegura, trabajó con cuestionarios impresos.

Cuando los dispositivos presentaron algún daño físico, como en la batería, los dispositivos se registraron en una aplicación que llevó el control de estos casos y fueron sustituidos con los de reserva.

El SIM se encargó de la configuración de los insumos informáticos, servicios de Internet y red privada virtual, así como la instalación y actualización de las aplicaciones de cómputo móvil en los DCM, tanto las versiones de capacitación como las productivas, requeridas para la captación de datos, actualización cartográfica e implementación del entorno empresarial. También fue responsable de monitorear y supervisar la infraestructura y control de fallas, accesorios y redes, a fin de reportar las incidencias ocurridas en la coordinación estatal de su adscripción.

Para apoyar las labores de soporte, se implementó un mecanismo de atención a usuarios, específicamente para la atención de fallas en las aplicaciones informáticas, en la cual las figuras operativas reportaban al SIM, quién resolvía la situación o la escalaba a su jefe inmediato hasta llegar a la coordinación estatal y de ser necesario hasta a oficinas centrales, en donde las áreas de desarrollo atendían directamente la situación.

Para la sistematización del mecanismo de atención a usuarios se desarrolló el módulo de **Soporte Informático y Gestión de DCM**, integrado en el sistema en línea **OPERA**, el cual funcionaba como un registro de tickets con la redacción de la situación presentada, este medio era el utilizado por los SIM para atender de manera ordenada, ágil y oportuna a sus usuarios.

Adicionalmente se hizo uso del **Sistema para la Consulta y Trazabilidad de los Equipos de Cómputo SICTEC**, herramienta web operada por la estructura de soporte informático, y que permitió consultar, actualizar y generar información de los Activos Informáticos del Censo 2020, permitiendo conocer la ubicación, estatus, proyecto, usuario asignado y vigencia de uso, así como cantidades totales de equipamiento en cada unidad administrativa.

El Censo contó con instrumentos de captación impresos, para ser utilizados en caso de que la aplicación llegara tener problemas durante la entrevista o por falta de energía eléctrica, así como en aquellas áreas de trabajo inseguras debido a la delincuencia, donde los entrevistadores corren el riesgo de ser asaltados o agredidos.

Los códigos QR eran únicos para cada instrumento impreso y para controlar su distribución entre las figuras responsables del levantamiento de información, pues mediante el módulo de control de instrumentos del *Administrador Censal*, fueron asignados a los supervisores y entrevistadores a través de la lectura de dicho código, con el uso de la cámara digital de los dispositivos, procedimiento que también se realizaba al momento de vincular los cuestionarios impresos, a la vivienda en la que habían sido utilizados para la realización de la entrevista. Posteriormente fueron validados y empaquetados mediante el módulo mencionado, y junto con un módulo auxiliar implementado en el OPERA, se controló su flujo hasta ser recibidos en los centros de captura.

Para la **Autoenumeración** y la **Entrevista asistida por teléfono**, se generaron cartas invitación, que contenían las credenciales únicas para responder el cuestionario básico vía Internet, y fue entregada por el entrevistador, bajo demanda de los informantes, o si en la tercera visita a la vivienda no se logró realizar la entrevista cara a cara; durante el recorrido de levantamiento el entrevistador iba llenando el *Listado de inmuebles*, procedimiento que incluyó la lectura en campo del código QR de las etiquetas que pegaba en los inmuebles, así como el de las invitaciones, de esta manera la vivienda quedaba asociada a estos datos, lo que permitió el control de la captación de información en el sitio Web.

3.6. Describa la forma de monitoreo de la recolección con dispositivos móviles de captura. Indique como se realizó el monitoreo y control de la transferencia de datos; la forma de control y monitoreo del avance de la recolección en términos de cobertura territorial y poblacional. Qué validaciones se realizaron para evitar duplicados (incluir algunos criterios de identificación), pérdida de datos durante la captura, o datos erróneos por mal funcionamiento del método.

El sistema OPERA es una plataforma institucional que permite atender cualquier proyecto de generación de estadística. Este sistema apoya tareas de seguimiento y control de las diferentes etapas del proyecto. Para su acceso se requiere que el usuario sea parte del dominio institucional, es decir que se haya creado una cuenta y una contraseña, y con base a sus credenciales, el sistema determina el perfil de usuario y presenta los módulos a los que tiene acceso.

Los principales módulos definidos en el OPERA para el Censo 2020 fueron:

- Integración de datos y distribución de materiales
- Reclutamiento y selección de personal
- Contingencias e incidencias operativas
- Seguimiento logístico
- Soporte Informático y Gestión de DCM (**MDM**)
- Seguimiento operativo

El módulo de *Seguimiento operativo*, es la herramienta que permite la detección oportuna de problemas de cobertura de los operativos censales, tomando como base las principales unidades de observación: Viviendas, Personas, Manzanas y localidades que integran el Marco Geoestadístico Nacional. Particularmente, las viviendas y su condición operativa: Visitada, censada, pendiente, deshabitada y uso temporal, tienen la mayor relevancia, ya que nos indican el avance del operativo de campo y proporcionan un panorama nacional que facilita la toma de decisiones que conlleve a una mayor cobertura.

También se presentan informes detallados sobre el avance en el proceso de actualización cartográfica detectada por el personal en campo, así como de la validación que de dichas actualizaciones realiza el personal de cartografía, quienes determinan aquellas que proceden y las que no.

Para la integración de la información a la base de datos central, como ya se mencionó, se implementaron dos mecanismos:

1. Envío de entrevistas completas desde el DCM mediante el uso de datos móviles.
2. Envío de todos los inmuebles por medio de memorias USB. El entrevistador le enviaba la información diariamente a su supervisor y éste, mediante el módulo de integración del OPERA, transfería vía Internet la información de todos sus entrevistadores al menos una vez al día.

Para el monitoreo y control de la transferencia de datos del operativo de campo, se integró un reporte de envíos de las figuras responsables de esta tarea, que informaba lo siguiente:

- Paquetes esperados a nivel nacional, estatal, y por estructura operativa hasta supervisor de entrevistadores. Los paquetes esperados estuvieron determinados por el personal contratado y una clave operativa única que se le asignó a cada entrevistador y supervisor de entrevistadores.
- Paquetes recibidos hasta el momento de la generación del reporte, porcentaje con respecto al número de paquetes esperados y paquetes faltantes.
- Estatus de los paquetes recibidos, que podían ser:
 - Paquete con integración correcta.
 - Paquete con integración parcial. Se revisaba manualmente para su integración total.
 - Paquete recibido con error. Era necesario volver a generar un nuevo envío.
 - Nomenclatura errónea. Era necesario volver a generar un nuevo envío.
 - Error en clave operativa. Era necesario volver a generar un nuevo envío desde el dispositivo de la figura operativa correspondiente.
 - Fecha de integración fuera de periodo/Fecha del archivo anterior al último envío recibido. Era necesario volver a generar un nuevo envío, en algunos casos implicaba corregir la fecha del DCM.

Aunado a este reporte, diariamente se generaba un listado nacional para cada supervisor o responsable de los envíos, en la que se indicaba, por entrevistador, la última fecha en la que se había recibido información, o si hasta el momento no se había recibido ningún inmueble de su carga de trabajo, este informe se enviaba a la estructura de coordinación y dirección del Censo, para que determinaran las causas y le dieran solución.

Para la eliminación de duplicados, se definió una estrategia de consolidación de la base de datos para tener datos homogéneos y únicos, respecto a la información censal que se captó en los diferentes operativos de campo del Censo 2020. Entre sus objetivos estuvieron asegurar la conciliación histórica de mnemónicos e implementar reglas de integridad de la información, así como definir los campos requeridos para cada instrumento y operativo, para que cada equipo de desarrollo se ajustara a las reglas definidas. En la definición de campos, se determinó generar un identificador único para cada unidad de observación, así se tiene uno para cada inmueble, vivienda particular habitada y vivienda colectiva, residente de la vivienda y persona migrantes, así como para cada localidad y manzana. Lo anterior permitió que en la etapa de integración de la base de datos operativa se pudieran identificar duplicados fácilmente.

Durante la entrevista cara a cara, conforme se iban cubriendo secciones del cuestionario, los datos captados se iban guardando en una tabla temporal y, una vez concluida la entrevista, se copiaban a la base de datos productiva, evitando su pérdida.

Para los casos de robo o daño total de dispositivo móvil, se elaboró una herramienta de restauración de información del último respaldo o paquetes de información generada, de esta manera, la pérdida de información se redujo considerablemente, ya que se complementó con una serie de comunicados a la estructura operativa, solicitando el respaldo constante de información, creando así una cultura de protección de datos censales.

4. Síntesis del aprendizaje sobre el método de captura.

4.1. Si usted tuviera la oportunidad de entregar tres recomendaciones sobre este método de captura a otro país, ¿qué elementos a partir de su experiencia deberían ser críticos para que el método funcione de forma adecuada?

Recomendación 1:

Considerar que desde el momento en que la captación de datos se lleva a cabo con el uso de dispositivos de cómputo móvil, se vuelve indispensable definir una estrategia informática para sistematizar el resto de los procesos del proyecto censal, desde la planeación operativa, el mapeo de riesgos operativos y la captación de información, hasta los procesos primarios, codificación, validación y generación de resultados. Una estrategia que responda a las necesidades funcionales y contemple un esquema de seguridad robusto, tal que contribuya a la calidad de los datos captados y a la cobertura total de áreas y unidades de observación. Para ello es importante contar con un requerimiento de desarrollo de software para cada sistema a desarrollar.

La estrategia debe incluir además una estructura de soporte informático lo más cercano posible a la estructura de levantamiento, con objeto de agilizar la atención de cualquier incidencia en materia informática que se presente, así como estructurar una estrategia de capacitación tal que garantice la transmisión de conocimientos y adquisición de habilidades para el correcto desempeño de sus funciones, no sin omitir que desde su reclutamiento se debe solicitar un perfil informático.

Recomendación 2:

En términos de tecnología, para el uso de dispositivos de cómputo móvil y la Autoenumeración para la captación de datos, los aspectos a considerar son:

Características mínimas de los dispositivos de cómputo móvil. Deben cuidarse en las especificaciones que se hacen en el proceso de licitación, ya que ello determina las características funcionales de las aplicaciones informáticas a implementar.

Desarrollo de aplicaciones de software. Es importante determinar la plataforma de desarrollo para dispositivos de cómputo móvil, ya que puede enfocarse para un determinado sistema operativo (Android, IOs o Windows). Esto aplica también a las plataformas para el desarrollo de aplicaciones en ambiente web.

Infraestructura de TI. Considera la infraestructura de cómputo requerida para el hospedaje de aplicaciones web (servidores de aplicaciones), integración de la información en una base de datos centralizada (Servidores de bases de datos); almacenamiento (NAS) y respaldo de archivos (dispositivos de respaldo como cintas magnéticas), así como el equipo de escritorio y de Cómputo móvil. Con respecto a redes y comunicaciones, se deben considerar servicios de voz y datos en oficinas del Instituto y de campo, el servicio de correo electrónico institucional, datos móviles para la telecomunicación en los DCM, entre otros.

Cobertura de los servicios de internet. Es importante tener en cuenta la capacidad de la infraestructura instalada en términos de tecnologías de la información y comunicación (TIC), así como la cobertura nacional, ya que es determinante para la estrategia de Autoenumeración.

Recomendación 3:

Con el objetivo de verificar el correcto funcionamiento de toda la infraestructura de TIC, es decir el hardware, el software, los elementos de red, los diferentes sistemas operativos utilizados, y el almacenamiento de datos, se debe definir un plan de pruebas considerando.

Pruebas de sistemas. El objetivo principal es asegurar que cada aplicación informática satisface las necesidades funcionales incluidas en el(los) requerimiento(s) de desarrollo de software realizado(s) por las áreas responsables de cada uno de los procesos a sistematizar para el proyecto censal.

Pruebas de volumen. Hacen referencia a grandes cantidades de datos para determinar los límites en que se causa que un sistema falle. Sirven para identificar la carga máxima o volumen que el sistema puede manejar en un periodo determinado, principalmente el más crítico, así, se realizan para verificar que la aplicación informática funcione adecuadamente bajo los siguientes escenarios de volumen: máximo número de clientes conectados y tamaño máximo de base de datos.

Hackeo ético (pruebas de penetración). Su objetivo es identificar y reparar posibles vulnerabilidades en la capa de servicios implementada para el proyecto, a fin de fortalecer y mejorar la estrategia de seguridad implementada.

5. Planes para el próximo censo de población y vivienda.

5.1. Marque con una X el Tipo de Censo que espera realizar para la próxima ronda censal en su país:

De Hecho De Derecho

5.2. Indique la duración de la recolección censal que espera para su próximo censo en el país, en semanas.

semanas. Si su censo será de hecho anotar 0.

5.3. Indique los métodos de recolección o captura que espera utilizar en su país para el próximo censo con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas que espera cubrir con cada método.

		% de Viviendas	% de Población
<input checked="" type="checkbox"/>	Papel/Digitación	2.00	2.00
<input type="checkbox"/>	Papel/Reconocimiento Óptico		
<input checked="" type="checkbox"/>	DMC / CAPI	95.00	95.00
<input checked="" type="checkbox"/>	Web / CAWI	3.00	3.00

5.4. Identificar con una X todos los tipos de cuestionario censal que va a considerar en el próximo censo en su país.

<input checked="" type="checkbox"/>	Viviendas particulares
<input checked="" type="checkbox"/>	Viviendas Colectivas
<input checked="" type="checkbox"/>	Personas en Situación de Calle
<input type="checkbox"/>	Personas en Tránsito o Viajeros
<input checked="" type="checkbox"/>	Otros

Si marca "otros", especificar en este recuadro:

- Servicio exterior mexicano (embajadas y oficinas consulares);
- Alojamiento de asistencia social
- Entorno urbano
- Cuestionario de la localidad
- Personalidades

5.5. Describa de forma breve los métodos a utilizar en el próximo censo, explicité de forma general las razones de mantener o cambiar los métodos utilizados. Además, incorpore los principales desafíos y avances en la preparación del censo en esta materia.

Los procesos censales en México han tenido transformaciones importantes a lo largo de los años, sobre todo derivados del uso de las tecnologías, y el Censo de Población y Vivienda 2020 no es la excepción, ya que su contribución más importante fue el uso de dispositivos de cómputo móviles (DCM) para el levantamiento de información, lo cual representó grandes retos.

En términos informáticos el reto más importante fue definir una estrategia informática que considerara los **aspectos técnicos** de la sistematización de procesos del *Censo*, que garantizara la **infraestructura tecnológica** requerida para el proyecto censal, considerando aspectos como:

La adquisición de DCM y equipamiento de las oficinas de campo y personal.

- Definición de las características del equipo de cómputo móvil y de escritorio
- Elaboración de dictámenes técnicos
- Definición y ejecución del plan de pruebas técnicas a los DCM.

Infraestructura de TIC

- Servidores de aplicaciones y base de datos, ancho de banda, red y correo institucionales, certificado de seguridad (SSL), almacenamiento NAS, firewall, voz y datos en las oficinas de campo y datos móviles en los DCM.

Servicios de TI /Seguridad /Redundancia

- Administración de servidores de aplicaciones y bases de datos.
- *Testing* de aplicaciones web. Plataforma OPERA, sistema de Autoenumeración, herramienta para posenumeración-cotejo, y el sistema de seguimiento y control del procesamiento SISEG.
- *Testing* de aplicaciones locales para la captación de información.
- Implementación de esquemas robustos de seguridad de la información.
- Esquemización de pruebas de volumen, concurrencia, conectividad e integración de información a base de datos, entre otros.
- Monitoreo para el correcto desempeño de aplicaciones web, los servidores de aplicaciones y base de datos, servicios de comunicación, etc.

Desarrollo de software

- Desarrollo, mantenimiento evolutivo y correctivo de aplicaciones móviles y web.

Recursos humanos

- Personal capacitado para el desarrollo de aplicaciones móviles y web.
- Estructura de soporte informático en las oficinas de campo.
- Personal especializado en la gestión de infraestructura de TIC.

Es importante considerar que el uso de la tecnología en este proyecto, proporcionó grandes experiencias y conocimiento a todas las áreas involucradas, fortaleciendo principalmente a las áreas informáticas del INEGI, y significó un gran esfuerzo a nivel institucional. En este sentido, es importante mencionar que las áreas de desarrollo y definición de las estrategias en materia informática dependieron directamente del área responsable del proyecto, lo que favoreció la consolidación de aplicaciones robustas y funcionalmente alineadas a las necesidades.

Por otro lado, la captación por Internet también conocida como Autoenumeración, representa otro desafío: la falta de participación de la población en general. El INEGI ha sido históricamente identificado con las entrevistas cara a cara y la gente difícilmente identifica el Censo sin esta interacción humana; es así que para fortalecer la estrategia de Autoenumeración, será necesaria una campaña más agresiva de difusión y comunicación sobre la importancia de la participación activa de la población. Se debe buscar que los informantes estén lo suficientemente sensibilizados para que, llegado el momento, capturen ellos mismos su información sin la presencia física del personal del Instituto. La Autoenumeración toma unos pocos minutos y se evita el contacto cara a cara con el entrevistador, lo que en tiempos de pandemia o cuando se trata de zonas inseguras es deseable para todos.

Asimismo, es importante que el informante tenga claridad de lo valioso que es su información para la toma de decisiones en el país, se debe insistir en la utilidad social de la misma. Hace falta mayor sensibilización sobre la relevancia de la oportunidad del registro y transformar a los participantes en agentes de cambio activos. La experiencia del Censo 2020 dejó en claro que aún hay un largo trecho por recorrer en este sentido, ello debido a que un sector muy acotado de la población fue el que participó en la Autoenumeración.

Otro tema importante a considerar es que se debe garantizar que la información proporcionada por el informante pueda relacionarse con la vivienda correcta, para ello, es indispensable contar con un directorio o listado de inmuebles previamente generado y validado, en el cual se tenga un elemento, que puede ser el QR de una carta invitación o el de una etiqueta pegada en la vivienda, que permita vincular la información que proporcione el informante con el domicilio en campo. Esto implica que se realice un recorrido previo en las áreas de trabajo o el envío de una carta por medio del servicio postal a cada domicilio.

Contemplar mecanismos de Autoenumeración es fundamental para no detener la generación de información estadística en casos de fuerza mayor como las pandemias, terremotos, inundaciones, entre otros. En el caso de la pandemia por COVID-19, se observó que algunos sectores de la población fueron reacios a atender a los entrevistadores incluso con las medidas de distanciamiento social y uso de cubrebocas y gel. En este sentido, el uso de servicios de comunicación y colaboración por videoconferencia, con plataformas como Microsoft Teams, Zoom (Zoom Video Communications, Inc.), entre otros, deberán ser evaluados, ya que su uso se ha visto incrementado.

Otro tema a impulsar es el uso de registros administrativos en el Censo, pues son una fuente de información confiable y de menor costo, sin embargo, esto deberá ser a largo plazo pues hace falta robustecer los mecanismos de captación y procesamiento de las instituciones que los generan, así como robustecer el marco legal que propicie el intercambio de información interinstitucional.

5.6. Si tuviera la oportunidad de realizar tres preguntas a otros países de América Latina y el Caribe que hayan implementado el método captura, o los métodos de captura, que usted proyecta para el próximo censo, ¿cuáles serían?

Pregunta 1:

¿Cuál fue la estrategia informática implementada para atender el proyecto censal, considerando la sistematización de los procesos de planeación, reclutamiento de personal, capacitación, captación y seguimiento operativo, así como para la Integración de la base de datos central y tratamiento y explotación de información? Además, incorpore los mecanismos de seguridad implementados tanto en los dispositivos móviles, como en la transferencia de información y las aplicaciones web para el seguimiento operativo, la gestión de dispositivos de cómputo móvil (MDM) y la Autoenumeración.

Pregunta 2:

Con respecto al desarrollo de aplicaciones móviles para la captación (Administrador Censal), ¿cuáles fueron sus características funcionales (módulos informáticos), y qué estrategia siguieron para el desarrollo de las herramientas de entorno empresarial (Kiosco INEGI) así como para la actualización cartográfica en campo (MCC)? Incluir las características de los insumos cartográficos, catálogo de planeación y su conciliación, requeridos para el correcto funcionamiento y oportuna implementación de las herramientas informáticas en campo.

Pregunta 3:

Con respecto a la gestión, control y seguimiento de los dispositivos de cómputo móvil (MDM), ¿qué herramientas implementaron?, indicar si fue un desarrollo in house, y cuáles fueron sus características funcionales y alcances. Incluir consideraciones para la adquisición, si fuera el caso (licitaciones de compra, arriendo o provisión de servicios). También incluir las plataformas soportadas (Apple, Android, Windows, ...)

Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe
Grupo de Trabajo Censos de la Conferencia Estadística de las Américas – CEPAL
Documento Lineamientos generales para la captura de datos censales

En el marco del plan de trabajo del **Grupo de censos de la CEA CEPAL** para el **bienio 2020-2021** se comprometió la realización de un documento que abordara lineamientos generales para la captura de datos censales para los países de América Latina y el Caribe. Uno de los objetivos específicos es la recopilación de experiencias por parte de los países respecto de la utilización de algunos de los métodos de recolección y captura de datos.

La presente ficha tiene por objetivo sistematizar la experiencia nacional respecto del método de captura de datos censales que ha sido solicitada a usted como representante de su país por parte del equipo de coordinación del Grupo de Trabajo de Censos de CEA Cepal.

La información aquí recopilada será utilizada como insumo para la redacción del documento, durante febrero del 2021 le enviaremos la versión preliminar del documento que será generado con su aporte para que sea previamente aprobada por la dirección de la oficina nacional de estadísticas de su país.

La ficha que se presenta consta de **5 partes**: Información de contacto; Características generales del último censo realizado en su país; Información del Método de Captura; Síntesis de aprendizajes; y Planes para el próximo censo.

Le pedimos enviar la ficha completa a más tardar el día **16 de octubre del 2020**. Su contenido será revisado por el equipo de coordinación del Grupo de Trabajo de Censos y de ser necesario le solicitará una reunión para profundizar en alguno de sus aspectos.

Agradecemos desde ya su colaboración en este proceso y ante cualquier duda o consulta por favor contactarse con Olga Barquero del INE-Chile: olga.barquero@ine.cl y Alejandra Silva de CELADE: alejandra.silva@cepal.org

1. Información de Contacto.

Por favor llenar con su información de contacto como responsable del llenado de la ficha por parte de su país.

Nombre	Edgar Vielma Orozco
Correo electrónico	Edgar.vielma@inegi.org.mx
Cargo	Director General de Estadísticas Sociodemográficas
Institución	Instituto Nacional de Estadística y Geografía, INEGI
País	México
Nombre	Sabino Navarro Campos
Correo electrónico	sabino.navarro@inegi.org.mx

Cargo	Subdirector de Desarrollo de Sistemas en la Dirección General de Estadísticas Sociodemográficas
Institución	Instituto Nacional de Estadística y Geografía, INEGI
País	México

2. Características generales del **último censo** realizado en el país.

2.1. Marque con una X el Tipo de Censo para el último censo realizado en su país:

De Hecho De Derecho

2.2. Indique la fecha de inicio de la recolección del último censo en su país:

Día Mes Año

2.3. Indique la duración de la recolección del último censo en su país en semanas, se debe incluir el período oficial además de las semanas que se extendió la recolección más allá de ese período hasta el cierre del operativo censal.

a. Duración del período oficial: semanas. Si su censo fue de hecho anotar 0.

En caso de que quiera especificar condiciones especiales para el período oficial de recolección, como, por ejemplo, diferencia de duración entre urbano y rural, especifique en el cuadro siguiente:

Al cierre del operativo de enumeración se habían censado el 95% de las viviendas particulares habitadas identificadas. No obstante, debido a la contingencia sanitaria derivada de la pandemia por COVID-19, el 31 de marzo la Secretaría de Salud emitió un Acuerdo por el que se suspendió toda actividad no esencial incluyendo el levantamiento de Censos y encuestas mediante entrevistas cara a cara; por lo que la etapa de verificación, que involucró la verificación de viviendas deshabitadas, así como la recuperación de pendientes y viviendas omitidas, se suspendió temporalmente, mientras que se amplió el periodo para que la población respondiera el cuestionario en línea por el método de Autoenumeración (Internet).

El 17 de julio, la Secretaría de Salud emitió el Acuerdo por el que se reanudan los Censos y Encuestas que involucran la movilización de personas y su interacción física, sujeto a un protocolo de actuación elaborado por el INEGI y aprobado por la Secretaría de Salud. Con base en esas directrices, las actividades de la fase de verificación se reiniciaron, pero se agregaron al instrumento de captación preguntas para identificar a los residentes de la vivienda en marzo de 2020.

La encuesta de Posenumeración se canceló debido a la distancia temporal de su aplicación, respecto de la etapa de Enumeración, sin embargo, fue sustituida por la Encuesta de Muestreo para la Estimación Focalizada de Cobertura (ENCOBER).

Por su parte, el calendario de difusión de resultados se modificó, por lo que los resultados básicos se publicarán este mismo año dependiendo de la evolución de la pandemia.

b. ¿Cuántas semanas se extendió la recolección después del período oficial?: La Enumeración no se extendió ningún día, sin embargo, la etapa de Verificación que había iniciado el 23 de marzo, y que fue suspendida el 31 de marzo en su etapa de levantamiento de campo, se reanudó a partir del 8 de junio de manera escalonada en cada coordinación estatal, con base en el semáforo de riesgo epidemiológico Covid-19, y de manera generalizada en todo el territorio nacional, a partir del 17 de julio, de conformidad con el acuerdo emitido por la Secretaría de Salud.

2.4. Indique los métodos de recolección o captura utilizados en el último censo en su país con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas cubiertas mediante ese método.

		% de Viviendas	% de Población
<input checked="" type="checkbox"/>	Papel/Digitación	2.00	2.00
<input type="checkbox"/>	Papel/Reconocimiento Óptico		
<input checked="" type="checkbox"/>	DMC / CAPI	97.7	97.7
<input checked="" type="checkbox"/>	Web / CAWI	0.30	0.30

2.5. Describa de forma general la estrategia operativa del censo. Incluya si los métodos de recolección o captura se aplicaron de forma secuencial o en paralelo; qué métodos de recolección o captura fueron utilizados para operativos especiales (viviendas colectivas, situación de calle, población en tránsito u otro); indicar si la captura de datos fue centralizada o descentralizada, o cualquier otra información relevante para el contexto de su experiencia.

La recolección de la información del Censo cubrió el conjunto de actividades para obtener los datos de cada unidad de observación, de acuerdo con el programa y los procedimientos de trabajo establecidos, incluyendo una estructura operativa y controles que promovieron la eficacia en cada una de las acciones. Esto implicó ejecutar los esquemas para la captación de datos diseñados previamente, como son: la preparación y distribución de materiales de apoyo, la integración de recursos humanos, la comunicación y concertación, bajo una detallada programación de actividades, una estructura orgánica y controles que se monitorearon continuamente desde la fase de planeación hasta el cierre del levantamiento.

La estrategia operativa contempló lo siguiente:

1. El **Censo**: cuyo periodo de levantamiento fue del 2 al 27 de marzo de 2020, integró la enumeración exhaustiva de viviendas y los operativos especiales para captar la información de la población residente en alojamientos colectivos, en el Servicio Exterior Mexicano y de la población sin vivienda (en situación de calle).
2. La **Verificación de cobertura**: con un periodo del 23 de marzo al 10 de abril de 2020, mismo que fue postergado debido la pandemia del Covid-19, y retomado hasta después del 17 de julio, fecha en que la Secretaría de Salud emitió el acuerdo mediante el que se permitió la reanudación de Censos y Encuestas.
3. La encuesta de **Posenumeración**: que se realizaría del 27 de abril al 8 de mayo de 2020, y que fue cancelada debido a la pandemia por Covid-19, fue sustituida por la Encuesta de Muestreo para la Estimación Focalizada de Cobertura (ENCOBER).

El Censo 2020 incluyó tres métodos de recolección de datos:

1. La *entrevista directa (DCM y papel)*: fue el principal método utilizado, se realizaba utilizando dispositivos de cómputo móvil en la mayoría de los casos, los cuestionarios en papel se utilizaron en las áreas de alto riesgo o debido a alguna incidencia con el dispositivo.
2. La *Autoenumeración*: se realizó mediante la implementación de una página web para que el ciudadano, a partir de una carta invitación entregada en su domicilio por el entrevistador, proporcionara los datos de las características de su vivienda y sus residentes.
3. La entrevista asistida por teléfono: el ciudadano marcaba al teléfono señalado en la carta invitación para proporcionar los datos de su vivienda y los habitantes.

Para la **Autoenumeración** y la **Entrevista asistida por teléfono** se generaron cartas invitación que contenían las credenciales únicas para responder el cuestionario básico vía Internet, y fue entregada por el entrevistador, bajo demanda de los informantes, o si en la tercera visita a la vivienda no se logró realizar la entrevista cara a cara; durante el recorrido de levantamiento el entrevistador iba llenando el *Listado de inmuebles*, procedimiento que incluyó la lectura en campo del código QR de las etiquetas que pegaba en los inmuebles, así como el de las invitaciones, de esta manera la vivienda quedaba asociada a estos datos, lo que permitió el control de la captación de información en el sitio Web.

El sistema *Administrador Censal* permitió aplicar los **cuestionarios** durante la entrevista directa a los informantes e incluía las siguientes funciones:

- Visualizar y administrar las cargas de trabajo de cada figura operativa.
- El Módulo de Captura para el registro de cada inmueble (Listado de Inmuebles).
- Los cuestionarios básico y ampliado para las viviendas particulares habitadas.
- Los cuestionarios para el registro de información de Entorno Urbano y Localidad.

Para censar a la población residente en viviendas colectivas (cárceles, asilos, centros de acogida, etcétera), la estructura de operativos especiales utilizó la aplicación para la captura de ésta información en un dispositivo móvil tipo *Netbook*.

En el caso de la población del Servicio Exterior Mexicano, se reutilizó la herramienta web de Autoenumeración, omitiendo la sección de vivienda y realizando ajustes a las preguntas de *Residencia hace 5 años* y *Lugar de nacimiento*.

La población sin vivienda (en situación de calle) se captó integrando previamente un directorio de puntos de encuentro e implementando, en el dispositivo de cómputo móvil, un módulo para el registro de los códigos QR de los cuestionarios aplicados y el total de personas (hombres y mujeres) captados en el punto de encuentro. El levantamiento de esta población se realizó en un solo día, el 11 de marzo.

Al igual que la población en situación de calle, para la captura de las personalidades, también se elaboró previamente un directorio, a dicha población se aplicó el cuestionario básico utilizando el DCM, y se adecuó el módulo de captura del *Administrador Censal*.

Se contó con instrumentos de captación impresos, para ser utilizados en caso de que la aplicación llegara a tener problemas durante la entrevista o por falta de energía eléctrica, así como en aquellas áreas de trabajo inseguras debido a la delincuencia, donde los entrevistadores corrieran el riesgo de ser asaltados o agredidos. Cada instrumento tuvo un código de respuesta rápida (QR) único, el cual los hizo irrepetibles, esto sirvió para controlar su distribución entre las figuras responsables del levantamiento de información.

Todos los cuestionarios levantados en papel fueron capturados de manera descentralizada, para ello, se implementaron 69 centros de captura, los cuales fueron coordinados por un responsable de captura estatal.

Para la Autoenumeración, la captación de información se realizó a partir de la entrega de una carta invitación por parte del entrevistador a los informantes, la cual incluía una clave y una contraseña, requeridos para ingresar a la aplicación Web implementada para esta estrategia. Con la captura de estos datos, iniciaba un proceso de registro, en que se le solicitó un correo electrónico personal, medio por el cual se le proporcionaron nuevas credenciales, indispensables para ingresar y proporcionar los datos de su vivienda, así como la lista y datos de sus habitantes.

En cuanto a la Entrevista asistida por teléfono, en la carta-invitación venían los números a los que el informante podía llamar, y un operador del Call Center del Instituto lo atendía. Para esta estrategia se hicieron adecuaciones a la aplicación de Autoenumeración. Es importante mencionar que, en caso de que el informante no contara con la clave y contraseña de la carta-invitación, podía proporcionar el dato del QR de la etiqueta pegada por el entrevistador en la fachada de su vivienda.

Para estas estrategias, la captura de información se realizó en paralelo, ya que gran cantidad de informantes pudieron acceder simultáneamente. Es importante mencionar que las respuestas a los cuestionarios se almacenaban de forma dinámica, evitando la pérdida de información, y permitiendo al informante proporcionar sus datos en varias sesiones.

3. Características del método: **Autoempadronamiento Web.**

3.2. Si usted tuviera que defender el uso de autoempadronamiento web frente a sus pares de otros países, ¿podría sintetizar el proceso de decisión para la elección de este método como parte del proceso de recolección censal?

El uso de dispositivos de cómputo móvil (DCM) para la captación de los datos censales, forma parte de las mejores prácticas internacionales y permitió que el Censo de Población y Vivienda 2020 estuviera a la vanguardia al determinar su uso como estrategia principal para la etapa de levantamiento. La Autoenumeración fue una alternativa más para la captación de información y con su implementación, se mantiene la tendencia de las mejores prácticas internacionales que promueven que los informantes se autoenumeren mediante el uso del Internet.

Un reto importante es garantizar que la información proporcionada por el informante pueda relacionarse con la vivienda correcta, para ello, es indispensable contar con un directorio o, como en el caso del Censo 2020, con un listado de inmuebles previamente generado y validado, el cual contó con un elemento como el QR de la carta invitación o el de la etiqueta pegada en el domicilio, que permitía vincular la información con el domicilio correcto en campo. Esto implicó que se realizara un recorrido previo en las áreas de trabajo o el envío de una carta por medio del servicio postal a cada domicilio.

Entre las ventajas de la Autoenumeración se pueden mencionar:

- Permite al informante proporcionar él mismo los datos de su vivienda y sus habitantes, sobre todo, cuando prefieren no abrirle la puerta a un entrevistador o aquellos que residen en zonas de estratos económicos altos en los que no se permite el acceso a los entrevistadores. Esto promueve además su participación consciente en el proyecto.
- Los mecanismos de seguridad son más robustos, ya que toda la capa de protección de datos se controla de manera centralizada, dependiendo exclusivamente de la institución o la compañía que proporcione el servicio de hospedaje de aplicaciones y almacenamiento de datos.
- Se evita el almacenamiento de materiales impresos o equipos de cómputo móvil y de oficina, lo que conlleva a la renta de espacios físicos.
- Si bien, se puede hablar de ahorros en tiempo, recursos humanos y monetarios, como ya se mencionó, es importante considerar las implicaciones de tener un directorio o un listado de inmuebles previo al operativo de captación por Internet.

3.3. Indique las pruebas realizadas durante la preparación de la recolección mediante autoempadronamiento web, incluir todas las pruebas tecnológicas, de cuestionario, operativas y censo experimental, donde se haya evaluado este método. Identificar los principales hallazgos en esas pruebas.

Para el Censo 2020, dos de las iniciativas más importantes para el levantamiento fueron la recolección de la información mediante el uso de dispositivos de cómputo móvil (DCM) para la *Entrevista cara a cara*, en tanto que, con la Autoenumeración, se impulsó la captación de información a través del uso del internet.

El primer ejercicio planteado para este proyecto fue la **Prueba de Estrategia operativa y dispositivos de cómputo móvil realizada en 2017**, en la cual, para probar la Autoenumeración, se seleccionaron áreas geográficas que, de acuerdo con otros proyectos, tuvieran un porcentaje considerable de viviendas con servicios de conexión a Internet. En esta primera versión, la captura de la información se realizaba de manera secuencial, y una vez que se iniciaba el cuestionario el informante tenía que responder todas preguntas. Como resultado de este ejercicio, para la siguiente versión se implementó un mecanismo para que los informantes pudieran responder el cuestionario en varias sesiones, sin perder la información proporcionada previamente, y se personalizó el acceso al cuestionario mediante el proceso de generación de credenciales a partir del registro del usuario.

El segundo ejercicio fue la **Prueba Piloto del Censo de Población y Vivienda 2018**, en la que se amplió el número de lugares y se probaron las adecuaciones realizadas al sistema de Autoenumeración. Se consideraron dos aplicaciones web, una para el registro del informante y otra exclusivamente para el cuestionario. También se implementaron esquemas de seguridad para encriptar la información, y se probó el guardado de datos por secciones, lo que permitió recuperar los datos guardados previamente y editarlos en caso de que el informante lo requiriera.

Las principales aportaciones de esta prueba fueron: estructurar el proceso de llenado de cuestionario, mejorar la interfaz del usuario para hacer más intuitiva la navegación, se afinaron las validaciones y, en general se mejoró la imagen del aplicativo considerando además la respuesta del cuestionario en diferentes idiomas.

Finalmente, se realizó el **Ensayo Censal 2019**, en el que la aplicación de Autoenumeración contaba ya con la estructura definitiva de un menú de avance del llenado, y permitió la navegación entre las diferentes secciones y preguntas del cuestionario. También se implementaron validaciones por sección y se mejoró aún más la imagen del sistema.

Como resultado de la Prueba, se implementó la funcionalidad para que la herramienta fuera responsiva y, se mejoró la experiencia del usuario corrigiendo los mensajes presentados, la redacción de los correos electrónicos que se enviaban al informante, así como las ayudas. También se mejoró el módulo de acceso al sistema y se implementaron mecanismos de tolerancia a fallos ocasionados por las desconexiones atribuibles a los servicios e Internet de los usuarios.

Es importante mencionar que, durante el año 2019 y principios del 2020, en materia informática se realizaron **pruebas de volumen y concurrencia**, a fin de garantizar la disponibilidad de las aplicaciones WEB, así como el desempeño de las bases de datos, tanto la del levantamiento mediante el uso de dispositivos móviles, como la correspondiente a la Autoenumeración. Se contrató a una empresa externa para la aplicación de *ethical hacking* en las aplicaciones, con el objetivo de detectar y eliminar riesgos de ataque externo durante la etapa de producción, etapa en la cual la Coordinación General de Informática del INEGI fue responsable del **monitoreo de seguridad y desempeño de la capa de servicios** implementada para el proyecto censal.

3.4. Sintetice el desarrollo TI de la recolección con autoempadronamiento web. Indique si el software de recolección fue desarrollado in house o se utilizó un software externo. Se solicita describir procesos de respaldo, transferencia de datos, elementos de seguridad y protección de datos personales (encriptaciones, manejo de claves y accesos, etc.); Incluir consideraciones para la adquisición (licitaciones de compra, arriendo o provisión de servicios).

Tanto para la **Autoenumeración** como para la entrevista asistida por teléfono, se implementó un sitio web, se desarrollaron aplicaciones responsivas y se implementaron validaciones primarias de la información y controles para la navegación para contribuir a la calidad e integridad de los datos proporcionados por el informante.

Las aplicaciones informáticas fueron desarrolladas por personal técnico del INEGI (*in house*), utilizando la tecnología de Microsoft con el entorno de Visual Studio 2017, del lado del servidor se basaron en .NET web pages Framework 4.62, y del lado cliente se utilizó Bootstrap para hacer responsivas las aplicaciones. La compatibilidad para el despliegue del sitio en los diferentes navegadores de Internet se consiguió basándose en los estándares de HTML 5, CSS 3, JavaScript 5 y jQuery; y, para el almacenamiento de la información, se diseñó e implementó un esquema de base de datos en Oracle 12c.

Para garantizar la seguridad del sitio Web, así como de la información proporcionada por el informante, se adquirió un certificado SSL (acrónimo de Secure Sockets Layer o capa de sockets seguros), tecnología estándar para mantener segura una conexión a Internet y el intercambio de datos punto a punto. Para todos los sitios expuestos en Internet con un certificado SSL, la URL comienza con **https:** indicando a los usuarios que sus datos se encuentran protegidos.

También se implementó un CAPTCHA en el módulo de registro del informante, en el cual el usuario crea sus credenciales e ingresa al cuestionario básico del Censo. Este CAPTCHA es una imagen almacenada en una tabla de la base de datos y la técnica desarrollada impide *hackeos*. Además, se implementaron métodos de programación segura para evitar la inyección de instrucciones SQL.

3.5. Identifique elementos de logística operativa asociados al autoempadronamiento web, por ejemplo, distribución de claves de acceso a la población, manejo de sesiones y soporte de ayuda a usuarios y resolución de problemas y cualquier otro elemento relevante en este sentido.

La distribución de dispositivos móviles se realizó mediante las figuras de soporte informático y auxiliares administrativos, desde las oficinas estatales a las municipales, las cuales contaron con servicio de voz y datos, además de ser seguras para el resguardo de los dispositivos. Para la entrega y la recepción de los dispositivos, el área administrativa implementó formatos de control impresos para su resguardo personalizado.

En cada una de las oficinas municipales se contó con la figura operativa de soporte informático municipal (SIM), responsable de proporcionar el servicio de soporte técnico a los entrevistadores y supervisores durante el levantamiento de información, y de apoyar en el seguimiento a incidencias reportadas con los DCM.

También se contó con una mesa de ayuda dirigida centralmente, la cual recibió y atendió toda la problemática operativa durante el levantamiento de información.

Al presentarse una pérdida o robo del dispositivo, la figura operativa levantó una denuncia ante las autoridades judiciales para entregar en el área administrativa el acta de hechos y documentar el caso. En estos casos, el entrevistador o supervisor recibió otro dispositivo o, si el área era insegura, trabajó con cuestionarios impresos.

Cuando los dispositivos presentaron algún daño físico, como en la batería, los dispositivos se registraron en una aplicación que llevó el control de estos casos y fueron sustituidos con los de reserva.

El SIM se encargó de la configuración de los insumos informáticos, servicios de Internet y red privada virtual, así como la instalación y actualización de las aplicaciones de cómputo móvil en los DCM, tanto las versiones de capacitación como las productivas, requeridas para la captación de datos, actualización cartográfica e implementación del entorno empresarial. También fue responsable de monitorear y supervisar la infraestructura y control de fallas, accesorios y redes, a fin de reportar las incidencias ocurridas en la coordinación estatal de su adscripción.

Para apoyar las labores de soporte, se implementó un mecanismo de atención a usuarios, específicamente para la atención de fallas en las aplicaciones informáticas, en la cual las figuras operativas reportaban al SIM, quién resolvía la situación o la escalaba a su jefe inmediato hasta llegar a la coordinación estatal y de ser necesario hasta a oficinas centrales, en donde las áreas de desarrollo atendían directamente la situación.

Para la sistematización del mecanismo de atención a usuarios se desarrolló el módulo de **Soporte Informático y Gestión de DCM**, integrado en el sistema en línea **OPERA**, el cual funcionaba como un registro de tickets con la redacción de la situación presentada, este medio era el utilizado por los SIM para atender de manera ordenada, ágil y oportuna a sus usuarios.

Adicionalmente se hizo uso del **Sistema para la Consulta y Trazabilidad de los Equipos de Cómputo SICTEC**, herramienta web operada por la estructura de soporte informático, y que permitió consultar, actualizar y generar información de los Activos Informáticos del Censo 2020, permitiendo conocer la ubicación, estatus, proyecto, usuario asignado y vigencia de uso, así como cantidades totales de equipamiento en cada unidad administrativa.

El Censo contó con instrumentos de captación impresos, para ser utilizados en caso de que la aplicación llegara a tener problemas durante la entrevista o por falta de energía eléctrica, así como en aquellas áreas de trabajo inseguras debido a la delincuencia, donde los entrevistadores corrieran el riesgo de ser asaltados o agredidos.

Los códigos QR eran únicos para cada instrumento impreso y para controlar su distribución entre las figuras responsables del levantamiento de información, pues mediante el módulo de control de instrumentos del *Administrador Censal*, fueron asignados a los supervisores y entrevistadores a través de la lectura de dicho código, con el uso de la cámara digital de los dispositivos, procedimiento que también se realizaba al momento de vincular los cuestionarios impresos, a la vivienda en la que habían sido utilizados para la realización de la entrevista. Posteriormente fueron validados y empaquetados mediante el módulo mencionado, y junto con un módulo auxiliar implementado en el OPERA, se controló su flujo hasta ser recibidos en los centros de captura.

Para la **Autoenumeración** y la **Entrevista asistida por teléfono**, se generaron cartas invitación, que contenían las credenciales únicas para responder el cuestionario básico vía Internet, y fue entregada por el entrevistador, bajo demanda de los informantes, o si en la tercera visita a la vivienda no se logró realizar la entrevista cara a cara; durante el recorrido de levantamiento el entrevistador iba llenando el *Listado de inmuebles*, procedimiento que incluyó la lectura en campo del código QR de las etiquetas que pegaba en los inmuebles, así como el de las invitaciones, de esta manera la vivienda quedaba asociada a estos datos, lo que permitió el control de la captación de información en el sitio Web.

3.6. *Describa la forma de monitoreo de la recolección con autoempadronamiento web. Indique como se realizó el monitoreo y control de la transferencia de datos; la forma de control y monitoreo del avance de la recolección en términos de cobertura territorial y poblacional. Qué validaciones se realizaron para evitar duplicados (incluir algunos criterios de identificación), pérdida de datos durante la captura, o datos erróneos por mal funcionamiento del método.*

Para la **Autoenumeración** y la **Entrevista asistida por teléfono**, se generaron cartas invitación, que contenían las credenciales únicas para responder el cuestionario básico vía Internet, y fue entregada por el entrevistador, bajo demanda de los informantes, o si en la tercera visita a la vivienda no se logró realizar la entrevista cara a cara; durante el recorrido de levantamiento el entrevistador iba llenando el *Listado de inmuebles*, procedimiento que incluyó la lectura en campo del código QR de las etiquetas que pegaba en los inmuebles, así como el de las invitaciones, de esta manera la vivienda quedaba asociada a estos datos, lo que permitió dar seguimiento de las viviendas con respuesta y, en su caso, volver a visitar aquellas que no hubieran contestado.

Se implementó un reporte en el módulo de Seguimiento Operativo del OPERA, que informaba lo siguiente:

- Número de invitaciones entregadas a nivel nacional, estatal, y por estructura operativa hasta supervisor de entrevistadores
- Número de invitaciones con respuesta en el sistema de Autoenumeración, y
- Número de invitaciones sin respuesta

Para la eliminación de duplicados, se definió una estrategia de consolidación de la base de datos para tener datos homogéneos y únicos, respecto a la información censal que se captó en los diferentes operativos de campo del Censo 2020. Entre sus objetivos estuvieron, asegurar la conciliación histórica de mnemónicos e implementar reglas de integridad de la información, así como definir los campos requeridos para cada instrumento y operativo para que cada equipo de desarrollo se ajustara a las reglas del negocio definidas.

En la definición de campos, se determinó generar un identificador único para cada unidad de observación, así se tiene uno para cada inmueble, vivienda particular habitada y vivienda colectiva, residente de la vivienda y persona migrantes, así como para cada localidad y manzana.

Lo anterior permitió que en la etapa de integración de la base de datos operativa se pudieran identificar duplicados fácilmente, y en la Autoenumeración, tanto el identificador único del inmueble, como los códigos QR de la invitación y de la etiqueta, fueron los elementos que permitieron la identificación de entrevistas duplicadas, incluso si, a pesar de haber dejado una invitación en la vivienda, posteriormente el entrevistador pudo realizar la entrevista.

Síntesis del aprendizaje sobre el método de captura.

4.1. Si usted tuviera la oportunidad de entregar tres recomendaciones sobre este método de captura a otro país, ¿qué elementos a partir de su experiencia deberían ser críticos para que el método funcione de forma adecuada?

Recomendación 1:

Definir una estrategia informática para sistematizar todo el proceso de captación por Internet (Autoenumeración). Dicha estrategia debe responder a las necesidades funcionales del proyecto censal y debe contemplar un esquema de seguridad robusto que contribuya a la calidad de los datos captados. Para ello, es importante contar con un requerimiento de desarrollo de software para cada sistema a desarrollar.

Recomendación 2:

Con el objetivo de verificar el correcto funcionamiento de toda la infraestructura de TIC, es decir el hardware, el software, los elementos de red, los diferentes sistemas operativos utilizados, y el almacenamiento de datos, se debe definir un plan de pruebas considerando.

Pruebas de sistemas. El objetivo principal es asegurar que cada aplicación informática satisfaga las necesidades funcionales incluidas en el(los) requerimiento(s) de desarrollo de software realizado(s) por las áreas responsables de cada uno de los procesos a sistematizar para el proyecto censal.

Pruebas de volumen. Hacen referencia a grandes cantidades de datos para determinar los límites en que se causa que un sistema falle. Sirven para identificar la carga máxima o volumen que el sistema puede manejar en un periodo determinado, principalmente el más crítico, así, se realizan para verificar que la aplicación informática funcione adecuadamente bajo los siguientes escenarios de volumen: máximo número de clientes conectados y tamaño máximo de base de datos.

Hackeo ético (pruebas de penetración). Su objetivo es identificar y reparar posibles vulnerabilidades en la capa de servicios implementada para el proyecto, a fin de fortalecer y mejorar la estrategia de seguridad implementada.

Recomendación 3:

En términos de tecnología, para el uso de dispositivos de cómputo móvil y la Autoenumeración para la captación de datos, los aspectos a considerar son:

Proveedores de servicios de Internet. Para evitar problemas de conectividad de los usuarios, se debe sugerir a los informantes que utilicen conexión por cable a sus computadoras en lugar de redes inalámbricas. En caso contrario, es imperante que el servicio proporcionado por el proveedor de internet sea estable, tenga una velocidad aceptable y no existan dispositivos conectados a la red que consuman altos niveles de datos como los que usan streaming (televisiones inteligentes, videojuegos o aplicaciones de video).

Desarrollo de aplicaciones de software. Es muy importante determinar la plataforma de desarrollo para dispositivos de cómputo móvil, ya que puede enfocarse para un determinado sistema operativo (Android, IOs o Windows). Esto aplica también para las herramientas para el desarrollo de aplicaciones en ambiente web.

Infraestructura de TI. Considera la infraestructura de cómputo requerida para el hospedaje de aplicaciones web (servidores de aplicaciones), integración de la información en una base de datos centralizada (Servidores de bases de datos); almacenamiento (NAS) y respaldo de archivos (dispositivos de respaldo como cintas magnéticas) así como el equipo de escritorio y de Cómputo móvil. Con respecto a redes y comunicaciones, se deben considerar servicios de voz y datos en oficinas del Instituto y de campo, el servicio de correo electrónico institucional, datos móviles para la telecomunicación en los DCM, entre otros.

Cobertura de los servicios de internet. Es importante tener en cuenta la capacidad de la infraestructura instalada en términos de tecnologías de la información y comunicación (TIC), así como la cobertura nacional ya que es determinante para la estrategia de Autoenumeración.

4. Planes para el próximo censo de población y vivienda.

5.1. Marque con una X el Tipo de Censo que espera realizar para la próxima ronda censal en su país:

De Hecho De Derecho

5.2. Indique la duración de la recolección censal que espera para su próximo censo en el país, en semanas.

semanas. Si su censo será de hecho anotar 0.

5.3. Indique los métodos de recolección o captura que espera utilizar en su país para el próximo censo con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas que espera cubrir con cada método.

		% de Viviendas	% de Población
<input checked="" type="checkbox"/>	Papel/Digitación	2.00	2.00
<input type="checkbox"/>	Papel/Reconocimiento Óptico		
<input checked="" type="checkbox"/>	DMC / CAPI	95.00	95.00
<input checked="" type="checkbox"/>	Web / CAWI	3.00	3.00

5.4. Identificar con una X todos los tipos de cuestionario censal que va a considerar en el próximo censo en su país.

<input checked="" type="checkbox"/>	Viviendas particulares
<input checked="" type="checkbox"/>	Viviendas Colectivas
<input checked="" type="checkbox"/>	Personas en Situación de Calle
<input type="checkbox"/>	Personas en Tránsito o Viajeros
<input checked="" type="checkbox"/>	Otros

Si marca "otros", especificar en este recuadro:

- Servicio exterior mexicano (embajadas y oficinas consulares);
- Alojamiento de asistencia social
- Entorno urbano
- Cuestionario de la localidad
- Personalidades

5.5. Describa de forma breve los métodos a utilizar en el próximo censo, explicité de forma general las razones de mantener o cambiar los métodos utilizados. Además, incorpore los principales desafíos y avances en la preparación del censo en esta materia.

Los procesos censales en México han tenido transformaciones importantes a lo largo de los años, sobre todo derivados del uso de las tecnologías, y el Censo de Población y Vivienda 2020 no es la excepción, ya que su contribución más importante fue el uso de dispositivos de cómputo móviles (DCM) para el levantamiento de información, lo cual representó grandes retos.

En términos informáticos el reto más importante fue definir una estrategia informática que considerara los **aspectos técnicos** de la sistematización de procesos del *Censo*, que garantizara la **infraestructura tecnológica** requerida para el proyecto censal, considerando aspectos como:

La adquisición de DCM y equipamiento de las oficinas de campo y personal.

- Definición de las características del equipo de cómputo móvil y de escritorio
- Elaboración de dictámenes técnicos
- Definición y ejecución del plan de pruebas técnicas a los DCM.

Infraestructura de TIC

- Servidores de aplicaciones y base de datos, ancho de banda, red y correo institucionales, certificado de seguridad (SSL), almacenamiento NAS, firewall, voz y datos en las oficinas de campo y datos móviles en los DCM.

Servicios de TI /Seguridad /Redundancia

- Administración de servidores de aplicaciones y bases de datos.
- *Testing* de aplicaciones web. Plataforma OPERA, sistema de Autoenumeración, herramienta para posenumeración-cotejo, y el sistema de seguimiento y control del procesamiento SISEG.
- *Testing* de aplicaciones locales para la captación de información.
- Implementación de esquemas robustos de seguridad de la información.
- Esquematización de pruebas de volumen, concurrencia, conectividad e integración de información a base datos, entre otros.
- Monitoreo para el correcto desempeño de aplicaciones web, los servidores de aplicaciones y base de datos, servicios de comunicación, etc.

Desarrollo de software

- Desarrollo, mantenimiento evolutivo y correctivo de aplicaciones móviles y web.

Recursos humanos

- Personal capacitado para el desarrollo de aplicaciones móviles y web.
- Estructura de soporte informático en las oficinas de campo.
- Personal especializado en la gestión de infraestructura de TIC.

Es importante considerar que el uso de la tecnología en este proyecto, proporcionó grandes experiencias y conocimiento a todas las áreas involucradas, fortaleciendo principalmente a las áreas informáticas del INEGI, y significó un gran esfuerzo a nivel institucional. En este sentido, es importante mencionar que las áreas de desarrollo y definición de las estrategias en materia informática dependieron directamente del área responsable del proyecto, lo que favoreció la consolidación de aplicaciones robustas y funcionalmente alineadas a las necesidades.

Por otro lado, la captación por Internet también conocida como Autoenumeración, representa otro desafío: la falta de participación de la población en general. El INEGI ha sido históricamente identificado con las entrevistas cara a cara y la gente difícilmente identifica el Censo sin esta interacción humana; es así que para fortalecer la estrategia de Autoenumeración, será necesaria una campaña más agresiva de difusión y comunicación sobre la importancia de la participación activa de la población. Se debe buscar que los informantes estén lo suficientemente sensibilizados para que, llegado el momento, capturen ellos mismos su información sin la presencia física del personal del Instituto. La Autoenumeración toma unos pocos minutos y se evita el contacto cara a cara con el entrevistador, lo que en tiempos de pandemia o cuando se trata de zonas inseguras es deseable para todos.

Asimismo, es importante que el informante tenga claridad de lo valioso que es su información para la toma de decisiones en el país, se debe insistir en la utilidad social de la misma. Hace falta mayor sensibilización sobre la relevancia de la oportunidad del registro y transformar a los participantes en agentes de cambio activos. La experiencia del Censo 2020 dejó en claro que aún hay un largo trecho por recorrer en este sentido, ello debido a que un sector muy acotado de la población fue el que participó en la Autoenumeración.

Otro tema importante a considerar es que se debe garantizar que la información proporcionada por el informante pueda relacionarse con la vivienda correcta, para ello, es indispensable contar con un directorio o listado de inmuebles previamente generado y validado, en el cual se tenga un elemento, que puede ser el QR de una carta invitación o el de una etiqueta pegada en la vivienda, que permita vincular la información que proporcione el informante con el domicilio en campo. Esto implica que se realice un recorrido previo en las áreas de trabajo o el envío de una carta por medio del servicio postal a cada domicilio.

Contemplar mecanismos de Autoenumeración es fundamental para no detener la generación de información estadística en casos de fuerza mayor como las pandemias, terremotos, inundaciones, entre otros. En el caso de la pandemia por COVID-19, se observó que algunos sectores de la población fueron reacios a atender a los entrevistadores incluso con las medidas de distanciamiento social y uso de cubrebocas y gel. En este sentido, el uso de servicios de comunicación y colaboración por videoconferencia, con plataformas como Microsoft Teams, Zoom (Zoom Video Communications, Inc.), entre otros, deberán ser evaluados, ya que su uso se ha visto incrementado.

Otro tema a impulsar es el uso de registros administrativos en el Censo, pues son una fuente de información confiable y de menor costo, sin embargo, esto deberá ser a largo plazo pues hace falta robustecer los mecanismos de captación y procesamiento de las instituciones que los generan, así como robustecer el marco legal que propicie el intercambio de información interinstitucional.

Finalmente, hace falta transmitir a la población los beneficios del uso de la tecnología en los Censos, y concientizarla sobre cómo los datos captados por el INEGI contribuyen a la generación de información que permita el establecimiento de políticas públicas y mejora la toma de decisiones en beneficio de la población.

5.6. Si tuviera la oportunidad de realizar tres preguntas a otros países de América Latina y el Caribe que hayan implementado el método captura, o los métodos de captura, que usted proyecta para el próximo censo, ¿cuáles serían?

Pregunta 1:

Con respecto del desarrollo de aplicaciones para la Autoenumeración, ¿cuáles fueron sus características funcionales? y ¿Qué requisitos mínimos de infraestructura y conectividad se plantean para que los informantes pudieran responder el cuestionario?

Pregunta 2:

¿Qué mecanismo implementaron para garantizar la correcta georreferenciación de las viviendas captadas por Internet? Incluya los procedimientos operativos implícitos, los mecanismos de seguridad instalados tanto del lado cliente como en el servidor, así como en la transferencia e integración de la información a la base datos.

Pregunta 3:

Para determinar la captación de información por Internet (Autoenumeración) como la estrategia principal para un proyecto Censal ¿qué consideraciones se deben tener tanto en la estrategia operativa como en la informática?

Ficha de Sistematización de Experiencias para Países de América Latina y el Caribe
Grupo de Trabajo Censos de la Conferencia Estadística de las Américas – CEPAL
Documento Lineamientos generales para la captura de datos censales

En el marco del plan de trabajo del **Grupo de censos de la CEA CEPAL** para el **bienio 2020-2021** se comprometió la realización de un documento que abordara lineamientos generales para la captura de datos censales para los países de América Latina y el Caribe. Uno de los objetivos específicos es la recopilación de experiencias por parte de los países respecto de la utilización de algunos de los métodos de recolección y captura de datos.

La presente ficha tiene por objetivo sistematizar la experiencia nacional respecto del método de captura de datos censales que ha sido solicitada a usted como representante de su país por parte del equipo de coordinación del Grupo de Trabajo de Censos de CEA Cepal.

La información aquí recopilada será utilizada como insumo para la redacción del documento, durante febrero del 2021 le enviaremos la versión preliminar del documento que será generado con su aporte para que sea previamente aprobada por la dirección de la oficina nacional de estadísticas de su país.

La ficha que se presenta consta de **5 partes**: Información de contacto; Características generales del último censo realizado en su país; Información del Método de Captura; Síntesis de aprendizajes; y Planes para el próximo censo.

Le pedimos enviar la ficha completa a más tardar el día **16 de octubre del 2020**. Su contenido será revisado por el equipo de coordinación del Grupo de Trabajo de Censos y de ser necesario le solicitará una reunión para profundizar en alguno de sus aspectos.

Agradecemos desde ya su colaboración en este proceso y ante cualquier duda o consulta por favor contactarse con Olga Barquero del INE-Chile: olga.barquero@ine.cl y Alejandra Silva de CELADE: alejandra.silva@cepal.org

1. Información de Contacto.

Por favor llenar con su información de contacto como responsable del llenado de la ficha por parte de su país.

Nombre	Victor A. Gonzalez R.
Correo electrónico	vgonzalez@ine.gob.ve / victor.alfredo.gonzalez@gmail.com
Cargo	Gerente General de Estadísticas Demográficas
Institución	Instituto Nacional de Estadística
País	Venezuela

2. Características generales del **último censo** realizado en el país.

2.1. Marque con una X el Tipo de Censo para el último censo realizado en su país:

De Hecho De Derecho

2.2. Indique la fecha de inicio de la recolección del último censo en su país:

Día Mes Año

2.3. Indique la duración de la recolección del último censo en su país en semanas, se debe incluir el período oficial además de las semanas que se extendió la recolección más allá de ese período hasta el cierre del operativo censal.

a. Duración del período oficial: semanas. Si su censo fue de hecho anotar 0.

En caso de que quiera especificar condiciones especiales para el período oficial de recolección, como, por ejemplo, diferencia de duración entre urbano y rural, especifique en el cuadro siguiente:

De manera pública y oficial el XIV Censo de Población y Vivienda, 2011, se levantó en 12 semanas, el resto de las semanas fueron consideradas para los operativos de casos específicos principalmente de la población en áreas de difícil acceso, rurales y/o indígenas.
--

b. ¿Cuántas semanas se extendió la recolección después del período oficial?: semanas.

2.4. Indique los métodos de recolección o captura utilizados en el último censo en su país con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas cubiertas mediante ese método.

		% de Viviendas	% de Población
<input checked="" type="checkbox"/>	Papel/Digitación	0,01%	0,02%
<input type="checkbox"/>	Papel/Reconocimiento Óptico		
<input checked="" type="checkbox"/>	DMC / CAPI	93%	95%
<input type="checkbox"/>	Web / CAWI		

2.5. *Describa de forma general la estrategia operativa del censo. Incluya si los métodos de recolección o captura se aplicaron de forma secuencial o en paralelo; qué métodos de recolección o captura fueron utilizados para operativos especiales (viviendas colectivas, situación de calle, población en tránsito u otro); indicar si la captura de datos fue centralizada o descentralizada, o cualquier otra información relevante para el contexto de su experiencia.*

Para el XIV Censo de población y vivienda, fue aplicado solamente el método tradicional de levantamiento de datos en campo tocando puerta por puerta en cada una de las unidades inmobiliarias con fines residenciales. La etapa pre-censal fue hecha según el método PAPI y el levantamiento de los datos censales según el método CAPI.

Actualización Cartográfica (AC) y Registro de Unidades Inmobiliarias (RUI) 2010 (Método PAPI)

En el año 2010, inició la etapa pre-censal con el levantamiento en campo en papel de la AC y el RUI, a partir del kit que tenían los actualizadores –mapa impreso del segmento, hojas formateadas para hacer los croquis y un cuadernillo para el registro del RUI–. La operación duró un año. El proceso de transcripción del RUI inició a mediados del 2010 en tres horarios –de 7am a 2pm, de 2pm a 7pm y de 7pm a 7am–, donde se transcribían de forma centralizada los cuadernillos del RUI levantados en campo. El desarrollo del DataEntry del RUI fue interno en ambiente Web (JavaScript de FrontEnd y Oracle de BackEnd). Posteriormente inició el proceso de digitación de las Áreas Operativas Censales, dándole prioridad a las Áreas Urbanas, las cuales se logró digitalizar en más de un 99% y de las Áreas Rurales posteriormente, de las cuales no se ha logrado digitalizar aún por lo pesada y amplias respecto a la capacidad tecnológica disponible en el INE.

Levantamiento de los datos censales 2011 (Método CAPI)

El levantamiento del censo se realizó desde el 1 de septiembre de 2011 hasta el 30 de noviembre de manera descentralizada, el arranque fue realizado de manera simultáneo en todo el país, según la secuencia establecida con base a las incidencias y condiciones de campo que eran manejadas por las Gerencias Estadales de Estadísticas del INE. Fueron implementados 946 Centros de Recolección y Transmisión Censal (CRTC) con 18 funcionarios por CRTC –2 supervisores, cada supervisor tenía 9 empadronadores a su cargo–, a cada empadronador se le asignaba un Área Operativa Estadística (AOE) –segmento en las áreas urbanas y sectores en las áreas rurales–. Se le asignaba en el DMC la carga de trabajo a cada Empadronador –Una AOE urbana o rural–, quién luego iba a campo a levantar los datos. Era guiado por el supervisor quien garantizaba la ubicación e inicio de cada empadronador. Los datos censales levantados eran consolidados en el CRTC a través de un cable conectada por USB a la laptop del CRTC, quien tenía conexión Web y transmitía diariamente los datos levantados en campo a nivel central.

En el último mes, se realizó un barrido para el levantamiento en zonas especiales o de difícil acceso, donde en algunos casos se asignaba el levantamiento en campo en lugar de ser por AOE, se hacía por manzanas censales para agilizar el levantamiento en campo con mayor número de empadronadores. Las colectividades o refugios eran levantadas en operativos especiales con un cuestionario ajustado a las condiciones de estas estructuras –información de la unidad inmobiliaria y de las personas que residen en ella, fuertes militares, cárceles, hospitales y otros–

3. Características del método: **Dispositivo Móvil de Captura.**

3.2. Si usted tuviera que defender el uso de dispositivos móviles de captura frente a sus pares de otros países, ¿podría sintetizar el proceso de decisión para la elección de este método como parte del proceso de recolección censal?

El método CAPI (Computer Assisted Personal Interview), es un método de recolección que implica principalmente la presencia del entrevistador en las unidades de observación y el uso de un dispositivo móvil (tablet, laptops, smartphone) con un cuestionario electrónico para realizar la entrevista. Esto significa que todas las ventajas de la presencia del empadronador son factibles; sin embargo, la mayoría de los errores que ocurren en el método PAPI (Paper And Pencil Interviewing), ya no ocurren al usar este método, gracias a las preguntas de cribado y a la programación de saltos y validaciones, que permiten la gestión de entrevistas más complejas, ayuda a evitar errores de captura de datos y hace que los datos estén disponibles en poco tiempo según las siguientes ventajas:

1. Admite la automatización de las secuencias de control y salto considerados en los cuestionarios de acuerdo a las respuestas ingresadas, lo que exime conocer a cabalidad todo el cuestionario, ya que las variables son presentadas o seccionadas para su captura.
2. Permite la segmentación y personalización de las preguntas, así como incluir textos de ayuda contextuales adicionales para aclaraciones de las preguntas o sus categorías.
3. Permite el control y la disminución considerable de los tiempos de captura, al permitir la codificación y/o precarga asistida de listas predefinidas de opciones o categorías de variables textuales –divisiones políticas administrativas, actividad económica, ocupación, países, pueblos indígenas, entre otros–, que se puede realizar durante la entrevista.
4. Es factible complementar las funcionalidades de los dispositivos de capturas de datos, con las aplicaciones móviles de geo-posicionamiento y el uso de mapas digitales o mosaicos de imágenes que permitan la ubicación y visualización de las áreas operativas censales de manera digital.
5. Mejora la gestión de los datos levantados en campo al unir el levantamiento de los datos con la captura o ingreso de los datos; lo que permite simultáneamente a nivel local o en campo, el monitoreo en tiempo real de las actividades vinculadas a la operación de campo, así como, la reducción de los tiempos y costos de la captura de los datos al eliminar la separación de los distintos procesos de obtención, captura y procesamiento de los datos, permitiendo optimizar el aseguramiento de la calidad de los datos (cobertura y contenido).
6. Puede ser utilizada como complemento del método CATI, en las condiciones de requerir levantar datos vía telefónica sin necesidad de estar en un espacio acorde para transcripción en línea, ya que permite el registro de las encuestas offline que se puedan levantar telefónicamente desde los hogares de los entrevistadores y posteriormente ser enviadas o centralizadas vía Bluetooth, FTP o a través de un servicio web.
7. Pueden reutilizarse el desarrollo y los dispositivos en otras operaciones estadísticas.

3.3. Indique las pruebas realizadas durante la preparación de la recolección con dispositivos móviles de captura, incluir todas las pruebas tecnológicas, de cuestionario, operativas y censo experimental, donde se haya evaluado este método. Identificar los principales hallazgos en esas pruebas.

En general fueron realizadas tres pruebas donde fue usado el dispositivo móvil de captura.

En la primera prueba sólo se consideró el manejo del dispositivo en paralelo con el llenado del cuestionario en papel, a fin de identificar las diferencias o incidencias respecto a los tiempos de ingreso de los datos en el DMC. Se identificaron debilidades en la lógica programación de saltos y validaciones y debilidades de la carga de las listas de ocupación y/o profesión y en la selección de la categoría o código idóneo, ya que las listas utilizaban nombres o descripción muy técnicas o estandarizadas distintas a la jerga declarada por los informantes, por lo que se optó captar las listas de pueblos indígenas, actividad económica y ocupación con campos abiertos. Se identificó debilidad en el manejo o gestión de los saltos o desplazamiento hacia atrás lo cual también generó la posibilidad de ingresar datos inconsistentes. Esta prueba fue realizada en paralelo por cinco encuestadores que aplicaron el método PAPI y cinco que aplicaron el método CAPI.

En la segunda prueba piloto se realizó exclusivamente con el uso de DMC, logrando mejorar los tiempos de respuestas y las aplicaciones correctas de los saltos y las validaciones generales de los datos censales. Se consideró evaluar el uso del dispositivo por los empadronadores, la duración de la batería y la consistencia de los datos ingresados. Como resultado se identificaron aún errores en lógica programación de la validación y en un algún caso de saltos. Se corroboró la visualización, personalización y fácil lectura de las variables del cuestionario en los DMC.

Las pruebas de gestión y asignación de cargas de trabajo se realizaron en el Censo Experimental realizado en la Parroquia Macuto y el Municipio Atures un semestre previo al Censo, con la supervisión de veedores internacionales, donde se realizaron pruebas de transmisión de los datos, desde los DMC a los CRTC a través de un cable USB disponible por CRTC donde se asignaba la carga de trabajo y descargaba los datos censales de los DMC; y desde los CRTC a nivel central vía internet. En la gestión de la carga de trabajo se identificaron incidencias que luego se mejoraron organizando los archivos de trabajo y los archivos levantados en campo en carpetas digitales donde se organizaron y controlaron mejor los archivos.

Se identificaron algunas incidencias en campo por lo que se optó construir un protocolo de respaldo de los CRTC en carpetas digitales, así como en los dispositivos móviles a fin de poder corregir los datos inconsistentes y los archivos corruptos omitidos o rechazados por las reglas de validación.

El mayor reto fue manejar las incidencias con los caracteres especiales vocales acentuadas, la ñe, entre otras y el manejo de los formatos de archivos de texto ASCII y UTF-8, los cuales generaban desplazamientos y errores en la carga a la base de datos. Lo que generó la instalación de dos salas de seguimiento y control, una para el procesamiento, revisión y recuperación de los datos censales acorde a las estructuras de entrada y otra de revisión de contenidos a nivel central.

3.4. Sintetice el desarrollo TI de la recolección con dispositivos móviles de captura. Indique si el software de recolección fue desarrollado in house o se utilizó un software externo. Se solicita describir procesos de respaldo, transferencia de datos, elementos de seguridad y protección de datos personales (encriptaciones, manejo de claves y accesos, etc.); Incluir consideraciones para la adquisición (licitaciones de compra, arriendo o provisión de servicios).

Para la adquisición de los DMC, se consideró y ejecutó una licitación con el apoyo y gestión del UNFPA para mejorar las condiciones de diligenciamiento, gestión y/o adquisición de los DMC.

Cada CRTC se instalaba en una parte acorde a las condiciones mínimas de espacio, dentro de un establecimiento de la administración con fines públicos –alcaldías, establecimientos de salud, establecimientos educativos, entre otros–, por lo cual no fue necesario el alquiler de los espacios utilizados para los CRTC.

Para el caso de levantamiento de los datos censales en campo, fueron utilizados los vehículos institucionales para trasladar los funcionarios desde los CRTC a los lugares donde se llevaba a cabo el levantamiento de los datos censales, fue necesario en variados casos contratar servicios de transporte aéreos, fluviales o terrestres tercerizados para llegar hasta las AOE asignadas a las cuadrillas de campo.

El desarrollo del sistema de captura CAPI/DataEntry fue realizado entre el equipo de desarrollo de la Oficina de Tecnología de Información (OTI) del INE y el equipo técnico de la Gerencia General de Estadísticas Demográficas, desarrollando sobre la herramienta de procesamiento y captura de datos censales CPro 4.2. Tanto en los dispositivos móviles como los equipos de desarrollo utilizaron Windows como sistema operativo. Adicionalmente se realizó una App para la transformación y gestión de los archivos de trabajo y los archivos levantados, desarrollado en Visual .NET como FrontEnd y SQLite como BackEnd. Esta aplicación copiaba los archivos completos y/o válidos a las carpetas de gestión según los estatus de las encuestas, generaba respaldos y creaba los archivos XML para la transmisión de los datos a nivel local (CRTC).

La transmisión de los datos se realizó a través de un cable USB disponible por CRTC donde se descargaba los datos XML de los DMC y los archivos de respaldo (comprimidos con una clave única). El CRTC gestionó los datos para la asignación y control de carga de trabajo, gestión del personal y los DMC, el seguimiento y control a través de indicadores que destacaban las incidencias de cobertura o contenido a través de semáforos, transferencia y carga de los archivos XML DMC-CRTC y luego a nivel central. Esta aplicación de gestión en los CRTC fue desarrollada por el equipo de la OTI en Visual .NET como FrontEnd y SQLite como BackEnd. El acceso a los sistemas era manejado por claves individuales para cada empadronador y supervisor en los DMC y Coordinador en los CRTC.

Las últimas dos semanas de la etapa censal se realizó un respaldo de todos los y de los DMC CRTC en carpetas, los cuales fueron posterior a la centralización de los equipos y de los CRTC revisados y consolidados a nivel central por la sala de procesamiento, revisión y recuperación de los datos censales en la que se recuperó aproximadamente un 0,5% de los casos que tuvieron algún tipo de incidencias en la transmisión de los datos.

3.5. Identifique elementos de logística operativa asociados a los dispositivos móviles de captura, por ejemplo, distribución de material para la recolección, retorno de material de recolección, manejo de contingencias en caso de pérdidas, destrucción o fallas en los materiales y cualquier otro elemento relevante en este sentido.

La gestión de los materiales y equipos utilizados en el XIV Censo de Población y vivienda, fueron gestionados a nivel central por un equipo paralelo complementario contratado para el control de la gestión y trasladado de los insumos (equipos, materiales de oficina, papeles de trabajo, etc.) coordinado por la Gerencia de Administración y la Coordinación de Bienes del INE.

El traslado de los equipos se realizó con las camionetas asignadas a cada Gerencia de Estadísticas Estadales del INE para el levantamiento del Censo y con el camión 350 del INE a nivel Central.

Todos los insumos fueron trasladados a las Gerencias de Estadísticas Estadales, quienes a su vez fueron trasladando de manera particular los insumos y requerimientos para las instalaciones de los CRTC en los espacios habilitados.

Las fallas u omisiones de los insumos –omisión de las carpetas de mapas o croquis de las AOE, problemas con los equipos CRTC, DMC o los cables USB para la trasmisión de los datos– para el levantamiento en campo por AOE eran notificadas por los CRTC a las Gerencias de Estadísticas Estadales y este a su vez a nivel central quienes gestionaban los requerimientos y luego enviaban a los CRTC, para poder reprogramar el levantamiento de los casos con incidencias.

El retorno de los equipos se realizó de igual forma en dos niveles, primero a nivel estatal y luego a nivel central.

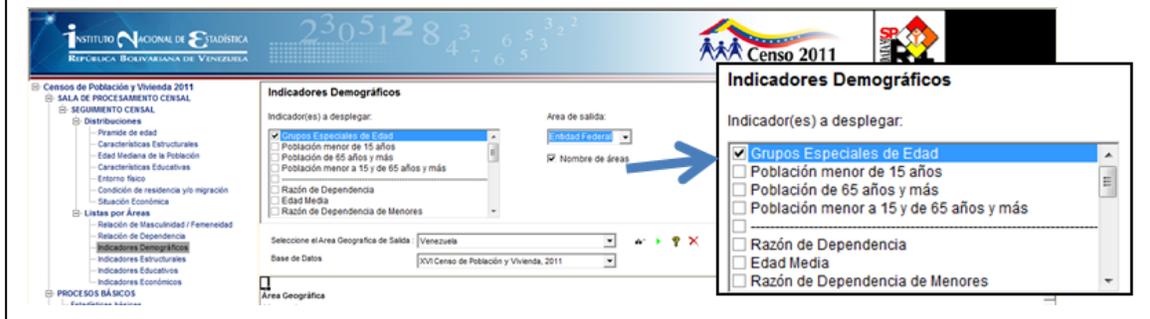
De los 15.000 sólo 170 dispositivos fueron perdidos por hurto o robo, de los cuales la mayoría fue por desatención o descuido de los empadronadores. De igual forma menos de 100 dispositivos tuvieron problemas de pila, pantalla y/o su funcionalidad, para lo cual el coordinador del CRTC estaba capacitado para la configuración o corrección de la funcionalidad de los dispositivos, en los que era factible su corrección –principalmente la reconfiguración del software–.

3.6. Describa la forma de monitoreo de la recolección con dispositivos móviles de captura. Indique como se realizó el monitoreo y control de la transferencia de datos; la forma de control y monitoreo del avance de la recolección en términos de cobertura territorial y poblacional. Qué validaciones se realizaron para evitar duplicados (incluir algunos criterios de identificación), pérdida de datos durante la captura, o datos erróneos por mal funcionamiento del método.

El monitoreo y control en el Censo 2011 se realizó en dos niveles, en el nivel local en los CRTC, y en el nivel central en la sala situacional o de seguimiento censal o en la sala de procesamiento, revisión y recuperación de los datos censales. A nivel local se realizó un seguimiento por semáforos de un conjunto de indicadores predefinidos que se adjuntan en la siguiente imagen, donde se le hacía el seguimiento cobertura y contenido del levantamiento de datos censales a partir de consultas o informes que procesaban los datos consolidados y mostraban las distribuciones o relaciones propuestas para el seguimiento, donde los valores resultantes eran comparados con valores esperados rangos inferiores o superiores. El semáforo se colocaba en rojo si dos o más criterios estaban por debajo de lo esperado, en amarillo si sólo uno estaba por debajo y en verde si todos estaban en el rango esperado.

INSTITUTO NACIONAL DE ESTADÍSTICA REPUBLICA BOLIVARIANA DE VENEZUELA		XIV CENSO GENERAL DE POBLACION Y VIVIENDA INDICADORES: CUARTILES (percentil 5 - percentil 95) POR CENTROS DE RECOLECCION		IQ1 XIV Censo Nacional de Población y Vivienda 2011					
Fecha de Impresión: 15/09/2011 11:57:58		Total de CRTC : 78		Page 1 of 30					
Entidad: DISTRITO CAPITAL									
CRTC	Pers/Hogar	Hogar/vivienda	Niños/Mef	Mujh/Mef	Ind.masc	Actividad	Cant. Viviendas Acumuladas	Empad. EN CRTC	
01001	3,76	1,21	31,51	58,04	96,23	62,96	826	14	■
01002	3,61	1,08	22,75	49,10	85,59	64,34	905	16	■
01003	3,91	1,02	37,81	62,75	100,90	52,83	1464	19	■
01004	3,63	1,03	37,05	65,10	101,58	56,27	929	18	■
01005	3,79	1,06	31,95	58,03	97,42	53,48	944	16	■

La información llega desde los Centros de Recolección y Transmisión Censal (CRTC) al Sistema Central en dos formatos (SQL y *.dat). La primera, SQL era presentada en la página web para ver la calidad del contenido de los datos. Luego con CSPro se verificaba y corregían los datos erróneos u omitidos y se generaba en Redatam para actualizar una aplicación X-Plan de Redatam para el análisis complementario. Se verificaban los distintos niveles de desagregación geográfica (entidad, municipio y parroquia) con informes por cortes semanales, donde se compara con el estimado en base a las proyecciones de población oficiales de 2001 y con la tendencia histórica censal y se determina si se encuentran dentro de los intervalos de confianza. Las validaciones de saltos y consistencias estuvieron programadas en el DataEntry por registro luego en la etapa post-censal se realizó entre registros y la limpieza final.



4. Síntesis del aprendizaje sobre el método de captura.

4.1. Si usted tuviera la oportunidad de entregar tres recomendaciones sobre este método de captura a otro país, ¿qué elementos a partir de su experiencia deberían ser críticos para que el método funcione de forma adecuada?

Recomendación 1:

Implementar el uso de aplicaciones tecnológicas a través de dispositivos móviles para la actualización cartográfica y el registro del padrón inmobiliaria, tales como CPro, Survey Solution, OruxMap, entre otros, a fin de unir el levantamiento y registro o captura de los atributos de estas unidades de observación, y a su vez, automatizar la data pre-censal insumo para la etapa censal, de tal manera que se puedan establecer controles en la operacionalización realizada en esta etapa.

Recomendación 2:

Una sugerencia para el aprovechamiento del método CAPI, es implementar la transferencia y sincronización de los datos desde el supervisor al empadronador para poder establecer controles de cobertura y/o contenido en campo a fin de aprovechar y optimizar al máximo los tiempos en campo para el levantamiento y diligenciamiento de los datos censales en las AOE.

Recomendación 3:

Considerar el uso de herramientas o soluciones tecnológicas de cuarta generación aplicadas para el desarrollo de encuestas en dispositivos móviles a fin de minimizar los tiempos de desarrollo e investigación necesarias en la etapa de programación de las aplicaciones.

5. Planes para el próximo censo de población y vivienda.

5.1. Marque con una X el Tipo de Censo que espera realizar para la próxima ronda censal en su país:

De Hecho De Derecho

5.2. Indique la duración de la recolección censal que espera para su próximo censo en el país, en semanas.

semanas. Si su censo será de hecho anotar 0.

5.3. Indique los métodos de recolección o captura que espera utilizar en su país para el próximo censo con una X, en caso de utilizar más de uno indicar el porcentaje de viviendas y el porcentaje de personas que espera cubrir con cada método.

		% de Viviendas	% de Población
<input type="checkbox"/>	Papel/Digitación		
<input type="checkbox"/>	Papel/Reconocimiento Óptico		
<input checked="" type="checkbox"/>	DMC / CAPI	92	93
<input checked="" type="checkbox"/>	Web / CAWI	8	7

5.4. Identificar con una X todos los tipos de cuestionario censal que va a considerar en el próximo censo en su país.

<input checked="" type="checkbox"/>	Viviendas particulares
<input checked="" type="checkbox"/>	Viviendas Colectivas
<input type="checkbox"/>	Personas en Situación de Calle
<input type="checkbox"/>	Personas en Tránsito o Viajeros
<input checked="" type="checkbox"/>	Otros

Si marca "otros", especificar en este recuadro:

Comunidades indígenas

5.5. Describa de forma breve los métodos a utilizar en el próximo censo, explicita de forma general las razones de mantener o cambiar los métodos utilizados. Además, incorpore los principales desafíos y avances en la preparación del censo en esta materia.

Los métodos a utilizar en el XV Censo de Población y Vivienda de Venezuela serán el método CAPI y el método alternativo de Auto-empadronamiento previo a la etapa censal.

Dado a las condiciones actuales de acceso a internet en los hogares o en los lugares de trabajo, estudio o particulares, se consideró usar el método complementario de levantamiento de datos censales por auto-empadronamiento a fin de permitir llegar a los hogares en primera instancia de la Administración Pública Nacional y luego para la población general. Esta opción se consideró abrir a toda la población como desafío para poder llegar a zonas que anteriormente no se ha llegado y en consideración de la pandemia del COVID-19.

Para el caso del levantamiento de datos en campo será utilizado CSPro/CAPI nuevamente por la capacidad de desarrollo en la institución, sin embargo, será utilizado tanto para la actualización cartográfica, el registro de unidades inmobiliarias, como para el levantamiento en campo del cuestionario censal en sus tres versiones (básico, ampliado y en colectividades).

El desafío del uso de tecnología en la etapa pre-censal obligó a normalizar y estandarizar la cartografía digital disponible, y de igual forma, la construcción de la metodología de elaboración de mosaico de imágenes de las AOE que no están disponibles, para que puedan ser precargados en los dispositivos móviles, y así puedan ofrecer la oportunidad de utilizar la cartografía digital y las imágenes satelitales vinculadas a los mapas o polígonos de las áreas operativas censales para la ubicación y actualización cartográfica.

De igual forma se utilizarán las bondades de geo-posicionamiento para la identificación de las coordenadas geográficas de las edificaciones y/o unidades inmobiliarias, así como la vinculación con la identificación del código QR impresas en las etiquetas de censado y de identificación del código compuesto de identificación de los registros de las viviendas con fines residenciales con el UBIGEO.

Por último, se desarrolló el manejo de la asignación de carga, el registro de la actualización cartográfica, el registro del cuestionario censal, la transferencia y sincronización de los datos y el aseguramiento de calidad con CSWeb, CSPro y CSEntry. Offline en campo y en línea entre los CRTC y el nivel Central.

5.6. Si tuviera la oportunidad de realizar tres preguntas a otros países de América Latina y el Caribe que hayan implementado el método captura, o los métodos de captura, que usted proyecta para el próximo censo, ¿cuáles serían?

Pregunta 1:

¿Cuál considera usted es el mayor reto que tiene las Oficinas Nacionales de Estadísticas para de implementar un Censo de cuarta generación en su país?

Pregunta 2:

De utilizar el método CAPI para el levantamiento de datos en campo de forma digital, ¿Qué opción consideraría para obtener la solución tecnológica y por qué?

- 1.- Realizaría un desarrollado in-house donde desarrollaría la estructura, funcionalidad, interfaces y control completo a partir de lenguajes de programación
- 2.- Realizaría un desarrollado in-house a partir de herramientas de automatización de encuestas, como por ejemplo CSPro, Survey Solutions, Survey Monkey o KoBoToolBox
- 3.- Un desarrollo llave en mano por terceros.

Pregunta 3:

¿Identifique tres recomendaciones para que las Oficinas Nacionales de Estadísticas puedan implementar una actualización cartográfica digital a través de dispositivos móviles?