

Términos de Referencia

1. Antecedentes.

En el marco del Programa de Fortalecimiento Institucional del INDEC y del Sistema Estadístico Nacional aprobado por la Dirección del Instituto para el período 2017-2020, en convenio con el Banco Interamericano de Desarrollo (BID), el INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS (INDEC) solicita la No Objeción del Banco para obtener expresiones de interés destinados para la posterior contratación de servicios de consultoría, con el fin de proceder a la realización del Censo Experimental 2020 (CE) y el Censo Nacional de Población, Hogares y Viviendas (CNPHyV).

Los oferentes serán seleccionados posteriormente por el método de Selección Basada en Calidad y Costo (SBCC) conforme a los procedimientos indicados en las Políticas para la Selección y Contratación de Consultores financiados por el Banco Interamericano de Desarrollo, GN 2350-9.

2. Objetivos del Trabajo.

El propósito de la contratación será fortalecer las capacidades de gestión de la ejecución del Censo Experimental 2020, cuyo objetivo será trazar la operatoria y los lineamientos del Censo Nacional de Población, Hogares y Vivienda que desarrollará el INDEC.

De acuerdo con las buenas prácticas y recomendaciones internacionales en la materia, como parte de las actividades de planeamiento y desarrollo metodológico que el INDEC diseñó para la ejecución del Censo Experimental 2020, se planificaron una serie de actividades con la finalidad de poner a prueba el contenido del cuestionario censal, el sistema de gestión informático, el relevamiento en campo, las tareas de los agentes que componen la estructura censal y la captura y recolección de los datos.

3. Alcance de los Servicios y Tareas Previstas.

Se espera recibir propuestas en formato de expresión de interés que promuevan una mejora del conocimiento técnico del Instituto en un proceso de captura de datos mediante lectura óptica de cuestionarios. Las firmas consultoras deberán contener las propuestas de mejora de los procesos operativos detallados en el presente.

Las propuestas deben ser adecuadas a objetivos trazados para el CE y por el CNPHyV y, por tanto, que permitan una implementación y puesta en marcha inmediata en el corto plazo. Deberán incluir la totalidad del proceso, desde la diagramación y diseño del cuestionario hasta la entrega de una base de datos lista para su validación.

Para ello, la adjudicataria de los trabajos deberá fortalecer a INDEC en la gestión del CE y por el CNPHyV, asegurando el desarrollo de capacidades de ejecución y recomendando mecanismos de coordinación entre

el Instituto, y los equipos provinciales e interadministrativos con responsabilidad directa en la realización de la tarea a efectuarse.

4.- Consideraciones generales.

Todas las tareas serán supervisadas en forma permanente por INDEC.

Antes del operativo censal, deberá realizarse una prueba piloto con todos los Cuestionarios Básicos y Ampliados y planillas, y entregue los mismos productos solicitados para el operativo censal.

Los Cuestionarios Básicos y Ampliados, así como las planillas anexas a los presentes Términos de Referencia son indicativas a efectos de la presentación de las expresiones de interés, y se encuentran sujetas a variaciones y modificaciones.

Todos los requerimientos deben estar probados por la empresa, y aprobados por INDEC previos al operativo censal.

Deberá realizarse un control de completitud y cobertura antes de comenzar con el escaneo de los documentos.

Las firmas consultoras a convocarse deberán contener en sus propuestas, la inclusión de un espacio físico y los recursos materiales (computadoras, impresoras, teléfono, etc.) dentro de su sede, para el personal del INDEC que acompañarán cada uno de los procesos (cantidad a definir).

Asimismo, deberá presentarse la arquitectura informática y el detalle de equipamiento que va a utilizar para las tareas objeto de la presente, de acuerdo al siguiente detalle:

- a) Mapa de arquitectura informática de toda la propuesta detallando que equipamiento provee la empresa y cual INDEC.
- b) Cantidad de equipamiento para la impresión detallando marca, modelo y especificaciones técnicas asociadas por cada tipo de impresora o equipo asociado a la tarea de impresión.
- c) Equipamiento a utilizar para la digitalización incluyendo cantidad, marca, modelo y especificaciones técnicas asociadas.
- d) Procedimiento para la captura de los cuestionarios detallando equipamiento asociado para dicho fin, controles, etc. Detallar equipos informáticos asociados, marca, modelo (Ídem puntos b y c).
- e) Arquitectura para el almacenamiento de la información capturada, espacios, tipo, lugar de almacenamiento, metodología de resguardo y seguridad de la información.
- f) Detalle de la arquitectura del sistema de control / gestión y que equipamiento utiliza.
- g) Esquema de alta disponibilidad, mitigación de riesgos, procedimiento de recupero de datos, tiempos de disponibilidad ante un evento.

5. Procesos Operativos

Los procesos operativos principales a aplicarse en el operativo Censo Experimental 2020 son:

I. Impresión de formularios.

- II. Digitalización.
- III. Toma de datos
- IV. Visualización de datos.

Requerimientos:

1. Requerimientos de Impresión.
2. Requerimientos Sistema digitalización imagen.
3. Requerimientos Sistema Captura de datos.
4. Requerimientos Sistema "Visualizador".

El momento de preparación y despacho quedará sujeto a la entrega por parte de INDEC de los datos finales de "destinos" y "volúmenes".

Será necesario que INDEC brinde un listado completo según los siguientes datos:

1. Provincia.
2. Localidad.
3. Departamento/Partido/Comuna.
4. Fracción N°.
5. Radio N°.
6. Segmento N°.
7. Escuela.
8. Cantidad Viviendas "A" y "B".

I.- IMPRESIÓN DE FORMULARIOS.

- Todos los documentos deben tener un código de barras único e irrepetible.
- Los códigos de barras no se deben repetir ni para un mismo tipo de documento ni entre documentos.
- La tinta utilizada debe permitir la realización del drop out.
- Debe incluirse el gramaje de papel.
- El papel no debe permitir que el grafito marque el reverso de la hoja.
- El tamaño de las hojas debe ser igual para todos los cuestionarios.
- El guillotinado de los cuestionarios debe realizarse en el mismo lugar en todos los cuestionarios.

II.- DIGITALIZACIÓN

Los formularios serán digitalizados por "TIPO" de acuerdo a la prioridad establecida por el INDEC.

El proceso de digitalización consiste en:

1. Recepción.
2. Acondicionamiento.
3. Digitalización.
 - a. Escaneo.
 - b. Informe de recepción de Formulario.

4. Publicación de imagen y meta-data.
5. Verificación y Control.
6. Transmisión imágenes controladas.
7. Expedición.

Recepción

Se realizarán lecturas de las cajas recibidas, igual que en cada uno de los puntos anteriores de Correo donde han circulado las cajas.

Acondicionamiento

Para los formularios C1A, C1B, C2A y C2B, las cajas pueden pasar directamente al sector de digitalización ya que se trata de formularios de una sola hoja.

Para el caso de los formularios Ampliados y Básicos, se deben separar sus hojas, si bien se espera que los formularios estén troquelados, se contará con guillotinas a los efectos de acelerar el proceso.

Digitalización

Se procederá a digitalizar por “caja”, tomando la caja como un “lote de escaneo”.

La base de datos estará conectada al “perfil de escaneo”, esto permitirá mayor integridad en la digitalización y facilitará el control al operador.

Estos datos existen en la BD, por asociación del número de caja y CB del Formulario.

Si bien, posteriormente, en el proceso de “Toma de datos” se realizará ICR sobre estos campos del formulario, en esta instancia, sirve al operador de digitalización para verificar su trabajo.

UBIGACION GEOGRÁFICA									
Provincia:	<input type="text"/>	Localidad:	<input type="text"/>	Departamento/ Partido/Comuna:	<input type="text"/>	Fracción Nº:	<input type="text"/>	Radio Nº:	<input type="text"/>
Código de calle:	<input type="text"/>	Nombre de calle:	<input type="text"/>						
Puerta Nº:	<input type="text"/>	Piso: Nº:	<input type="text"/>	Depto/ Pieza:	<input type="text"/>	Edificio/Mono- block e tira:	<input type="text"/>	Entrada/ Escalera:	<input type="text"/>

El software de digitalización permitirá rescatar datos de la base, como a su vez cargar datos en la misma.

De esta forma a medida que se van digitalizando se puede ir informando la recepción de los CB de los formularios.

Debe existir una tarea previa realizada en forma conjunta con el INDEC de ordenamiento y control de completitud.

Los documentos deben ser escaneados respetando los anclajes indicados en cada uno de ellos.

Evaluar la realización correcta del drop out en los documentos.

Tener un sistema de control que evalúe el escaneo por lotes, para poder rechazar en caso de error y volver a escanear los lotes indicados.

El personal que realiza el escaneo debe ser capacitado previamente a través de una prueba piloto de escaneo.

Escanear de manera ordenada teniendo en cuenta la ubicación geográfica de los documentos.

Tener un sistema de control para evaluar la omisión o repetición de documentos escaneados.

No se deben escanear el material sobrante (cuestionarios y planillas que no haya sido utilizados).

Publicación de imágenes y meta-data.

Luego que se realiza la publicación de las imágenes en el share, se debe realizar una verificación contra la base, arrojando los posibles faltantes para que el operador realice una verificación sobre digitalizado.

Verificación y Control

Se realizarán dos tipos de controles en esta instancia:

- Calidad de la imagen.
- Cuantitativo respecto a la base.

Este último implica poder consultar contra la BD la cantidad máxima esperada (es decir todos los formularios despachados en la caja), y también que no falten secuencia de los formularios multi- páginas (ampliadas y básicas).

En caso que las cantidades no coincidan, con la cantidad máxima “esperada” el operador deberá realizar un conteo para asegurarse que no quedaron formularios sin digitalizar.

Transmisión imágenes controladas.

Las imágenes de las cajas que pasaron el proceso de verificación y control, en forma exitosa serán transmitidas a un share definitivo, de donde el sistema de “toma de datos” realizará su trabajo.

Expedición

Con las cajas que han sido digitalizadas en forma exitosa se procederá a prepararlas para su expedición en pallets y generando los remitos correspondientes, de forma similar a como fueron expedidas originalmente. El detalle de los remitos sobre cantidades y seriales que contiene cada caja, debe surgir de los datos.

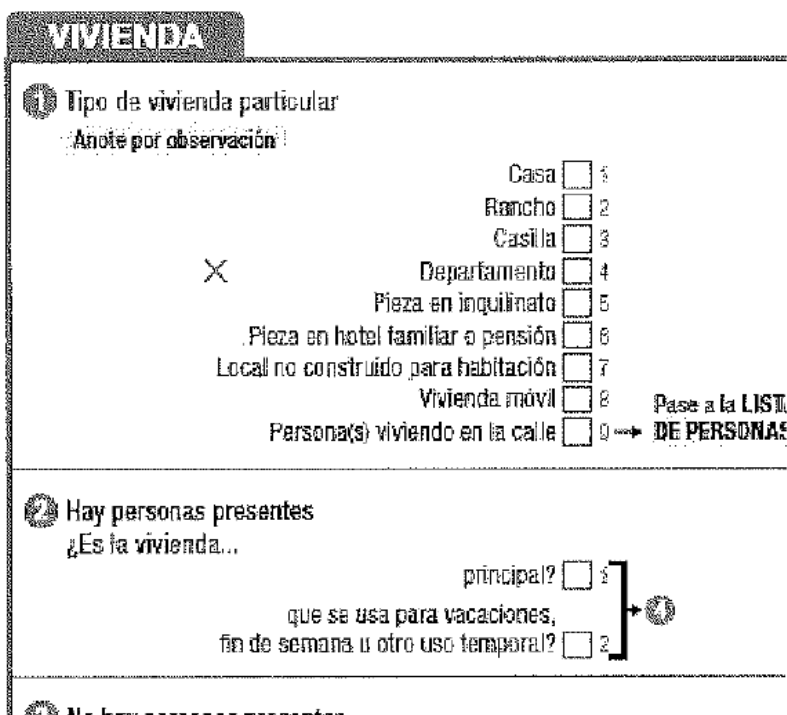
III.- TOMA DE DATOS

El proceso de toma de datos consiste en reconocer los datos del formulario en forma “automática” y trasladarlo a la BD, según la estructura solicitada por el cliente.

BCR: en el proceso de digitalización esto ya se realiza y las imágenes quedan indexadas de acuerdo al número de serie del formulario que se trate. Igualmente, en el proceso de “Toma de datos”, se valida la lectura y se verifica contra la indexación ya establecida, contemplando, establecer mecanismos de corrección en caso que se requiera.

OMR: el cliente pide específicamente reconocimiento de la marca “X” en los casilleros que así se deben completar.

Teniendo en cuenta que estos campos se completan con lápiz y que estos tipos de reconocimientos constituyen la mayor cantidad de campos de los formularios del Censo, se debe tratar de garantizar la mayor efectividad posible por el sistema de reconocimiento.



VIVIENDA

1 Tipo de vivienda particular

Anote por observación:

Casa 1

Rancho 2

Casilla 3

X Departamento 4

Pieza en inquilinato 5

Pieza en hotel familiar o pensión 6

Local no construido para habitación 7

Vivienda móvil 8

Persona(s) viviendo en la calle 9 → **Pase a la LISTA DE PERSONAS**

2 Hay personas presentes

¿Es la vivienda...

principal? 1

que se usa para vacaciones,
fin de semana u otro uso temporal? 2

3 No hay personas presentes

ICR

- Reconocimiento de datos que escribe el censista, pero ya existen los datos precargados y/o asignados: Como son algunos datos del Ubicación Geográfica.
- Reconocimiento de datos numéricos que calcula el censista, producto la cantidad de formularios completados: Como por ejemplo el “Total de Cuestionarios entregados” en los formularios.
- Reconocimiento de textos varios, producto de las respuestas recibidas de los censados.

Debe contemplarse:

- La calibración de todos los escáneres (constante y umbral) de acuerdo con el grafito utilizado.
- Asegurar que todos los escáneres tengan la misma calibración para cada tipo de documento.
- Ajustar el porcentaje de cobertura del check box para alcanzar la mejor captación de las marcas. Si no se garantiza la captación adecuada se deberá probar el reemplazo del check box por la burbuja.

- Garantizar el correcto reconocimiento de la marca (OCR) en todos los check box de los documentos.
- Garantizar el correcto reconocimiento de letras (ICR).
- Garantizar el correcto reconocimiento de los números (ICR).
- Realizar la interpretación respetando los instrumentos de control y consistencia brindados por INDEC (Ej. Diseño de registro, códigos geográficos, tablas de segmentación, y diccionarios para variables seleccionadas).
- Homogeneizar la intensidad del escaneo a través de la utilización del mismo lápiz y el que sea sugerido por la empresa.
- Incorporar la experiencia de la II Prueba Piloto para el entrenamiento del software en la interpretación de números y letras/palabras.
- Interpretación de la letra Ñ de manera correcta y no como N.
- Incorporar variantes de caligrafía para cada letra y número de forma de mejorar la interpretación las preguntas abiertas.
- Respetar los espacios de las palabras en las preguntas abiertas, es decir no permitir que se lean palabras sin espacio entre palabras, tanto en mismo renglón como en más de un renglón. Ejemplo: si dice MANEJA COLECTIVO DE PASAJEROS no puede decir MANEJACOLECTIVO DE PASAJEROS O MANEJA COLECTIVO DEPASAJEROS O MANEJACOLECTIVODEPASAJEROS.
- Asegurar que los errores de captura de las marcas, letras y números no superen los límites establecidos por el INDEC.

IV.- SISTEMA “VISUALIZADOR”

El sistema de visualización dispondrá de acceso web de:

La posibilidad de filtrar a través de distintos criterios (campos).

La opción de visualización de “lista” de resultados posibles.

La opción de poder exportar los datos sobre la lista.

La opción de poder ingresar y visualizar una selección específica de la lista.

La opción de poder exportar los datos sobre la imagen visualizada.

La opción de poder exportar la imagen visualizada.

6. Productos esperados.

1. Una base de datos por cada tipo de documento digitalizado.

Consideraciones respecto de la base de datos:

- Respetar el diseño de registro indicado por el INDEC:
 - o Respetar las posiciones indicadas en el diseño de registro de las variables. Por ejemplo, si una variable cuenta con dos posiciones y el cuestionario dice “1” que la lectora grabe “01”.
 - o Respetar el formato indicado para cada variable: no permitir símbolos en las variables numéricas. Por ejemplo, en edad “∩31”, en ese caso debería decir “31”.
- Integridad de la base de datos:
 - o Cada registro de la base de datos debe ser único y debe identificarse con los siguientes campos: tipo de documento, código de barra y carilla.
 - o No pueden existir registros duplicados en la base de datos
 - o Todos los registros deben tener un ID. No pueden quedar registros con ID en blanco.

- o No puede haber dos registros diferentes con el mismo código de barra.
 - Ubicación Geográfica
 - o Validar la Ubicación geográfica a través de la utilización de diccionarios geográficos y la tabla de segmentación para asegurar que estos campos estén completos cuando corresponde.
 - o Los cuestionarios correspondientes al mismo segmento deben contar con exactamente los mismos datos en los campos: PROV – DPTO – LOC – FRAC – RADIO – SEG
 - o Validar los nombres de las calles a partir la utilización de un diccionario con el código y nombres de las calles.
2. Una imagen por cada documento digitalizado.
- En el envío la de las imágenes escaneadas se debe incluir un resumen que indique la cantidad de imágenes por tipo de documento. Este resumen debe ser consistente con la base de datos enviada.
 - La organización y los nombres de los archivos de las imágenes debe ser clara (debe contener el tipo de documento, el ID y un código geográfico)
 - No puede existir más de una imagen para el mismo documento.
 - No puede haber archivos mal nombrados.
 - No puede haber nombres de carpetas poco intuitivas, es decir que los nombres de las carpetas y archivos deben ser indicativas de su contenido.
 - Se debe entregar un documento que especifique la organización implementada para poder encontrar los formularios fácilmente.
3. Software visualizador:
- Que contenga todas las imágenes.
 - Que sea de fácil manejo y ágil.
 - Que permita el acceso por tipo de documento, por ID, por ubicación geográfica.
 - Que permita la descarga individual y simultánea de imágenes (por ejemplo, todos los cuestionarios de una misma ubicación geográfica).
- Controles que debe realizar el proveedor antes de entregar los productos.
- El total de documentos debe ser el mismo para la base de datos, las imágenes escaneadas y el visualizador. Estos tres productos deben ser parte de la misma entrega, es decir, no pueden ser entregados de forma independiente y en fechas diferentes.
 - Antes de la entrega de los productos se solicita que la empresa realice una evaluación entre lo registrado en la base de datos y las imágenes de todas las preguntas. Evaluar marcas, letras y números por separado. Definir márgenes de error para cada caso.



República Argentina - Poder Ejecutivo Nacional
2020 - Año del General Manuel Belgrano

Hoja Adicional de Firmas
Informe gráfico

Número:

Referencia: TDR CR2020-CNPHyV

El documento fue importado por el sistema GEDO con un total de 8 pagina/s.