

Documento metodológico
para el aprovechamiento
estadístico de registros
administrativos económicos



NACIONES UNIDAS

CEPAL



años

Trabajando por
un futuro productivo,
inclusivo y sostenible



Conferencia
Estadística

de las
Américas
de la CEPAL

Gracias por su interés en esta publicación de la CEPAL



Si desea recibir información oportuna sobre nuestros productos editoriales y actividades, le invitamos a registrarse. Podrá definir sus áreas de interés y acceder a nuestros productos en otros formatos.

Deseo registrarme



NACIONES UNIDAS



www.cepal.org/es/publications



www.instagram.com/publicacionesdelacepal



www.facebook.com/publicacionesdelacepal



www.issuu.com/publicacionescepal/stacks



www.cepal.org/es/publicaciones/apps

Documento metodológico
para el aprovechamiento
estadístico de registros
administrativos económicos



NACIONES UNIDAS

CEPAL



**Conferencia
Estadística**

de las
Américas
de la CEPAL

Este documento fue desarrollado por el Grupo de Trabajo para la elaboración de un documento metodológico sobre aspectos conceptuales y aprovechamiento estadístico de registros administrativos económicos de la Conferencia Estadística de las Américas. El Grupo fue coordinado por Colombia (Departamento Administrativo Nacional de Estadística (DANE)) y México (Instituto Nacional de Estadística y Geografía (INEGI)) y tuvo como Secretaría Técnica a la División de Estadísticas de la Comisión Económica para América Latina y el Caribe (CEPAL). Los países miembros del Grupo son: Argentina (Instituto Nacional de Estadística y Censos (INDEC)), Brasil (Instituto Brasileño de Geografía y Estadística (IBGE)), Chile (Instituto Nacional de Estadísticas (INE)), Costa Rica (Instituto Nacional de Estadística y Censos (INEC)), Cuba (Oficina Nacional de Estadística e Información (ONEI)), Paraguay (Instituto Nacional de Estadística (INE)), Perú (Instituto Nacional de Estadística e Informática (INEI)) y República Dominicana (Oficina Nacional de Estadística (ONE)).

Publicación de las Naciones Unidas
LC/CEA.11/18/Rev.1
Distribución: L
Copyright © Naciones Unidas, 2023
Todos los derechos reservados
Impreso en Naciones Unidas, Santiago
S.23-00764

Esta publicación debe citarse como: Grupo de Trabajo para la elaboración de un documento metodológico sobre aspectos conceptuales y aprovechamiento estadístico de registros administrativos económicos de la Conferencia Estadística de las Américas, *Documento metodológico para el aprovechamiento estadístico de registros administrativos económicos* (LC/CEA.11/18/Rev.1), Santiago, Comisión Económica para América Latina y el Caribe (CEPAL), 2023.

La autorización para reproducir total o parcialmente esta obra debe solicitarse a la Comisión Económica para América Latina y el Caribe (CEPAL), División de Documentos y Publicaciones, publicaciones.cepal@un.org. Los Estados Miembros de las Naciones Unidas y sus instituciones gubernamentales pueden reproducir esta obra sin autorización previa. Solo se les solicita que mencionen la fuente e informen a la CEPAL de tal reproducción.

Reconocimientos

En el debate, la investigación y la redacción de este documento participaron:

- Argentina, INDEC: Javier Domench y Silvia Mario
- Brasil, IBGE: Vinicius Mendonça Fonseca, Synthia Kariny Silva de Santana, Cristiano de Almeida Martins y Raquel Rose Silva Correia
- Chile, INE: Pedro Ruz Zúñiga y Andrea Paola Marín
- Costa Rica, INEC: Odilia Bravo Cambroner y Rosa Cordero Peñaranda
- Cuba, ONEI: Evelyn Martínez Mendoza y Mariana Lobaina Sánchez
- Panamá, Instituto Nacional de Estadística y Censo (INEC): Roberto Arosemena
- Paraguay, INE: Mirta Leiva y Rafael Aquino
- Perú, INEI: Marleni Orrillo Huamán y Yessica Panuera Moreno
- República Dominicana, ONE: Juan Arias, Rene Guzmán y Orlando Hernández
- Colombia, DANE: José Richard Núñez Alejo, Julieth Alejandra Solano Villa, Sayda Patricia Morera Zarate, Sheila Isabel Centeno Martínez, Mónica Lorena Ortiz Medina y Juliana Benavides Pérez
- México, INEGI: Gerardo Durand Alcántara, Carmen Lizbet Corona Fuentes, Hugo Hernández Ramos, Elsa María Trueba Atienza, Juan Arturo Balderas Ávila, Adriana Herrera Canales y Adolfo Muñoz Torres
- CEPAL: Claudio Moris

Índice

Introducción.....	9
Capítulo I	
Importancia de los registros administrativos	11
Introducción.....	11
A. Fragmentación e integración.....	12
Capítulo II	
Visión de un sistema de registros administrativos económicos	15
Introducción.....	15
A. Sistema de registros base	15
B. Visión del registro base empresarial: el sistema de registros administrativos económicos.....	16
1. Objetivos del sistema de registros administrativos económicos.....	16
2. Objetivos del registro estadístico de negocios o registro estadístico empresarial como parte del sistema de registros administrativos económicos	17
3. Elementos que conforman el modelo del sistema de registros administrativos económicos.....	18
C. Registro estadístico de negocios o registro estadístico empresarial.....	19
1. Base de datos estructurada.....	19
2. Identificador único para cada unidad económica.....	19
3. Regularmente actualizado	20
4. Área territorial.....	20
5. Base de datos administrada por una oficina nacional de estadística	20
D. El Modelo Genérico de Procesos Institucionales Estadísticos: marco de trabajo para el sistema de registros administrativos económicos.....	20
Capítulo III	
Especificación de necesidades.....	23
Introducción.....	23
A. Subprocesos de la especificación de necesidades	23
1. Identificación de necesidades	24
2. Consulta y confirmación de necesidades	24
3. Definición de objetivos	25
4. Identificación de conceptos.....	25
5. Comprobación de la disponibilidad de datos.....	26
6. Elaboración del plan general y documentación.....	28
Capítulo IV	
Diseño	29
Introducción.....	29
A. Diseño conceptual	29
1. Definición del conjunto de objetos del registro.....	30
2. Número de identificación único para las unidades de observación que sea de uso público	31
3. Elaboración de un inventario de las fuentes	34
4. Criterios para la selección de fuentes.....	35
B. Diseño de las descripciones de las variables	37
1. Los resultados estadísticos que se espera generar con el proyecto	37
2. Identificación y selección de variables, siguiendo criterios estadísticos y conceptuales.....	37
C. Diseño de la recolección.....	39
D. Diseño del procesamiento y análisis.....	40
E. Diseño de los sistemas de producción y flujos de trabajo.....	43
1. Pasos para la selección y utilización de las fuentes	43
2. Pasos en el diseño de enlace entre los datos que aportan las fuentes y las bases de datos del registro estadístico	44
Capítulo V	
Construcción.....	45
Introducción.....	45
A. Definiciones a nivel de proceso.....	45

1. Construcción de instrumentos de recolección.....	45
2. Construcción o mejora de componentes	47
3. Construcción o mejora de componentes de difusión	50
4. Configuración de flujos de trabajo.....	50
5. Pruebas del sistema de producción	51
6. Prueba piloto del proceso estadístico.....	51
7. Finalización del sistema de producción	51
B. Módulos del sistema	51
1. Gestión con entidades proveedoras de datos	51
2. Acopio de registros administrativos	52
3. Transformación de los registros administrativos	52
4. Documentación de los datos.....	53
5. Administración	53
Capítulo VI	
Recolección.....	55
Introducción.....	55
A. Aspectos normativos o legales.....	56
1. Convenios interinstitucionales o memorandos de entendimiento	56
B. Aspectos metodológicos.....	58
1. Estructura e información necesaria.....	58
2. Periodicidad.....	58
3. Oportunidad.....	59
4. Medio o canal de intercambio	59
5. Confidencialidad	59
C. Mecanismo tecnológico.....	60
D. Capacitación	60
Capítulo VII	
Procesamiento	61
Introducción.....	61
A. Integración de datos	63
B. Clasificación y codificación	64
C. Revisión y validación	65
D. Edición e imputación.....	66
E. Derivación de nuevas variables y unidades	67
F. Finalización de los archivos de datos	67
Capítulo VIII	
Análisis	69
Introducción.....	69
A. Análisis del registro de empresa.....	69
B. Subprocesos de la fase de análisis	71
1. Preparación de borradores de resultados.....	71
2. Validación de los resultados.....	71
3. Interpretación y explicación de los resultados	72
4. Aplicación del control a la divulgación.....	72
5. Finalización de resultados.....	72
Capítulo IX	
Difusión.....	73
Introducción.....	73
A. Actualización de sistemas de salida	73
B. Generación de productos de difusión	74
C. Gestión de la publicación de productos de difusión	74
D. Promoción de productos de difusión	74
E. Gestión de soporte a las personas usuarias	74
Capítulo X	
Evaluación.....	75
Introducción.....	75

A. Características de la evaluación.....	75
B. Viabilidad de la evaluación.....	75
1. Dimensión institucional.....	76
2. Objetivos específicos.....	76
3. Dimensión técnico-metodológica.....	76
4. Objetivos específicos.....	76
5. Abordaje metodológico de la evaluación.....	77
6. Características a considerar en la evaluación de las dimensiones.....	78
7. Convenios o acuerdos.....	78
8. Necesidades.....	78
9. Diseño conceptual y analítico.....	81
10. Recepción de los registros administrativos.....	81
11. Integración de los registros administrativos.....	81
12. Análisis y difusión.....	81
13. Confidencialidad.....	81
14. Costos y beneficios.....	81
15. Asistencia técnica y registración.....	82
Capítulo XI	
Conclusiones y recomendaciones.....	83
Introducción.....	83
A. Hacia un sistema de registros administrativos económicos.....	83
B. Modelo del proceso estadístico (MGPIE).....	84
C. Detección y análisis de las necesidades.....	84
D. Diseño.....	85
E. Construcción.....	85
F. Recolección.....	86
G. Procesamiento.....	86
H. Análisis.....	86
I. Difusión.....	87
J. Evaluación.....	87
Bibliografía.....	89
Anexos	
Anexo A1.....	91
Anexo A2.....	99
Anexo A3.....	101
Anexo A4.....	104
Anexo A5.....	109
Anexo A6.....	121
Anexo A7.....	123
Anexo A8.....	124
Glosario.....	131
Cuadros	
Cuadro III.1 Ejemplo de ficha metodológica del registro administrativo.....	26
Cuadro III.2 Ejemplo de ficha metodológica del registro administrativo.....	26
Cuadro IV.1 Modelo de clasificación de las unidades estadísticas de subordinación empresarial.....	31
Cuadro X.1 Matriz de dimensiones institucionales.....	79
Cuadro X.2 Matriz de dimensiones técnico-metodológicas.....	80
Cuadro A1.1 Matriz de consulta para la identificación de necesidades de información.....	92
Cuadro A1.2 Matriz de necesidades de información y existencia de la información.....	92
Cuadro A1.3 Matriz de identificación preliminar de conceptos.....	93
Cuadro A1.4 Ejemplo para comprobar la disponibilidad de datos.....	95
Cuadro A1.5 Ejemplo del análisis de alternativas.....	96
Cuadro A1.6 Ejemplo para identificar el grupo de involucrados.....	97
Cuadro A1.7 Ejemplo de acta constitutiva.....	99
Cuadro A3.1 Formulario para el levantamiento del inventario de registros administrativos.....	102
Cuadro A3.2 Ficha técnica del registro administrativo.....	103
Cuadro A4.1 Ítems del canal de intercambio.....	106

Cuadro A4.2	Ítems de la estructura de datos para caracterizar archivos (tablas)	108
Cuadro A4.3	Ítems de la estructura de datos para caracterizar el contenido de cada archivo	108
Cuadro A5.1	Descripción de los softwares utilizados en la actualización	120
Cuadro A5.2	Determinación del segmento empresarial	120
Cuadro A5.3	Determinación de la organización jurídica	121
Recuadros		
Recuadro VII.1	Dimensiones clave en la calidad de las estadísticas	62
Recuadro A8.1	Cuestionario referido a convenios o acuerdos	125
Recuadro A8.2	Cuestionario referido a necesidades	125
Recuadro A8.3	Cuestionario referido a diseño conceptual y analítico	126
Recuadro A8.4	Cuestionario referido a la captura de los registros administrativos	126
Recuadro A8.5	Cuestionario referido a integración de los registros administrativos	127
Recuadro A8.6	Cuestionario referido a confidencialidad	127
Recuadro A8.7	Cuestionario referido a costos y beneficios	128
Recuadro A8.8	Cuestionario referido a análisis y difusión	128
Recuadro A8.9	Cuestionario referido a asistencia técnica y registración	129
Diagramas		
Diagrama II.1	Los cuatro registros base propuestos por Wallgren y Wallgren (2012).....	16
Diagrama II.2	Visión del sistema de registros administrativos económicos	18
Diagrama II.3	Ejemplos de identificadores únicos	19
Diagrama II.4	Fases y subprocesos del Modelo Genérico de Procesos Institucionales Estadísticos (MGPIE)	21
Diagrama III.1	Subprocesos de la fase de especificación de necesidades	24
Diagrama IV.1	Subprocesos incluidos en la fase de diseño	29
Diagrama IV.2	Esquema de la base de datos identificadora para la unidad de observación.....	33
Diagrama IV.3	Esquema del vínculo del registro estadístico de negocios con los registros estadísticos empresariales .	41
Diagrama IV.4	Ejemplo de modelo de base de datos del sistema de registros administrativos económicos	42
Diagrama V.1	Subprocesos incluidos en la fase de construcción	45
Diagrama V.2	Diferentes tipos de metadatos y herramientas en la documentación de un registro.....	50
Diagrama VI.1	Subprocesos incluidos en la fase de recolección	55
Diagrama VII.1	Subprocesos incluidos en la fase de procesamiento	61
Diagrama VII.2	Registros estadísticos a partir de registros administrativos	63
Diagrama VIII.1	Subprocesos incluidos en la fase de análisis	71
Diagrama IX.1	Subprocesos incluidos en la fase de difusión	73
Diagrama X.1	Subprocesos incluidos en la fase de evaluación	75
Diagrama XI.1	Fases y subprocesos del Modelo Genérico de Procesos Institucionales Estadísticos (MGPIE)	84
Diagrama A1.1	Flujograma de especificación de necesidades	91
Diagrama A1.2	Ejemplo de esquema para el análisis de problemas o necesidades	97
Diagrama A1.3	Ejemplo de esquema para determinar objetivos	98
Diagrama A4.1	Esquema de transferencia de archivos	106
Diagrama A4.2	Esquema de modelado de almacén de datos con datos de la Superintendencia de Pensiones	107
Diagrama A5.1	Proceso para la actualización del directorio central de empresas y establecimientos	111
Diagrama A5.2	Recepción y procesamiento de la información	111
Diagrama A5.3	Actualización de la base de datos	112
Diagrama A5.4	Validación de la base de datos	113
Diagrama A5.5	Georreferenciación del directorio	113
Diagrama A5.6	Georreferenciación a través de campo	114
Diagrama A5.7	Información cartográfica	115
Imágenes		
Imagen A5.1	Información cartográfica con código de zona y manzana	115
Imagen A5.2	Información cartográfica de los establecimientos por coordenadas X e Y	116
Imagen A5.3	Información cartográfica por código de frente censal, zona censal, manzana censal y coordenadas X e Y	116
Imagen A5.4	Información cartográfica de establecimientos con empleo de la base cartográfica de ciudades del Sistema de Focalización de Hogares (SISFOH)	117
Imagen A5.5	Empleo de Mapcity para direcciones con número de lote y manzana	117
Imagen A5.6	Registro de supervisión de campo	118
Imagen A5.7	Registro de digitación	119
Imagen A5.8	Resultado de estado y visita	119

Introducción

En los últimos años, diversos países del mundo han centrado sus esfuerzos en utilizar los registros administrativos como una fuente de información para la obtención de estadísticas. En el caso de los países desarrollados, estos poseen una infraestructura estadística que contiene, por lo menos, datos de identificación y ubicación de las unidades de observación (personas, viviendas, establecimientos, empresas, predios, entre otras), que actualizan mediante los registros administrativos. A partir de esa infraestructura, generan marcos muestrales para la realización de encuestas. Esta actualización es posible mediante la utilización de un identificador único y público para cada unidad de observación que adoptan los organismos gubernamentales, como las oficinas nacionales de estadística (ONE).

En los países de América Latina y el Caribe, la situación es diferente. Existen aún diversos problemas que impiden conformar una infraestructura similar a la de los países desarrollados, mediante la cual puedan crearse marcos de muestreo efectivos para generar estadísticas económicas precisas y confiables.

Entre los problemas que enfrentan los países de América Latina y el Caribe se encuentran:

- Marcos legales rígidos que impiden a los organismos gubernamentales compartir los registros administrativos con las oficinas nacionales de estadística.
- Los registros administrativos tienen objetivos específicos que no necesariamente persiguen un propósito estadístico. Además, los registros administrativos están basados en marcos conceptuales propios del objetivo que persiguen, y no siempre se corresponden con los que desarrollan las oficinas nacionales de estadística.
- Los organismos gubernamentales y las entidades privadas que poseen registros administrativos utilizan clasificadores que satisfacen las necesidades propias, pero que son diferentes a los utilizados por las oficinas nacionales de estadística.
- Las oficinas nacionales de estadística construyen de manera aislada los marcos de muestreo para el diseño estadístico de sus encuestas.
- Si bien existen diversas recomendaciones internacionales en materia de aprovechamiento estadístico de los registros administrativos, no existe una metodología ni herramientas estadísticas e informáticas que permitan a los países de la región construir la infraestructura a partir de registros administrativos para su potencial uso estadístico.

Bajo este planteamiento, se desarrolla el presente documento con el propósito de proporcionar a los países de la región un conjunto de lineamientos y recomendaciones que permitan generar estadísticas económicas novedosas y oportunas, y brindar mayores beneficios mediante el aprovechamiento estadístico de los registros administrativos. En el presente documento metodológico, al mencionar los registros administrativos se entenderá que se refiere a registros administrativos de temas económicos.

Los cuatro insumos fundamentales para la elaboración de este documento son:

- i) El libro *Register-based Statistics. Administrative Data for Statistical Purposes*, de Anders Wallgren y Brit Wallgren. Contiene una presentación exhaustiva de los registros administrativos que alimentan los cuatro registros base: registro de población, registro de empresas, registro inmobiliario y registro de actividades que, a su vez, soportan el sistema estadístico nacional (SEN) de cada país. En este documento solamente se abordará el registro de empresas.

- ii) El documento de la Comisión Económica para Europa (CEPE) *Guidelines on Statistical Business Registers*. Como se verá en el capítulo II, el registro estadístico de negocios o el registro estadístico empresarial es la columna vertebral de los procesos de generación de estadísticas económicas. Este documento está alineado a estas recomendaciones internacionales, ya que el registro estadístico de negocios o registro estadístico empresarial ofrece enormes ventajas en la generación de estadísticas económicas. En el presente documento, al mencionar un registro estadístico de negocios, se entenderá que se refiere al registro estadístico de negocios o registro estadístico empresarial.
- iii) El “Modelo genérico del proceso estadístico GSBPM (versión 1.0 en español, septiembre 2016)” de la CEPE. Este modelo es el marco de trabajo sobre el cual se desarrollan los lineamientos y recomendaciones para un mejor aprovechamiento estadístico de los registros administrativos económicos.
- iv) La experiencia de los países de la región.
Uno de los insumos más importantes para la elaboración de este documento es la experiencia de quienes representan a las oficinas nacionales de estadística de los países que conforman la región de América Latina y el Caribe. Se incorporan buenas prácticas realizadas por estos países y se evidencia cómo han resuelto problemas que son comunes en la región.

Este documento está conformado por 11 capítulos. En el primer capítulo se muestra la importancia de los registros administrativos. En el capítulo II se presenta la visión que se tiene del sistema de registros administrativos económicos. Entre los capítulos III y X, se desarrollan recomendaciones y lineamientos para cada una de las ocho fases del modelo: especificación de necesidades, diseño, construcción, recolección, procesamiento, análisis, difusión y evaluación. Finalmente, el capítulo XI contiene las conclusiones y recomendaciones. Para una mejor comprensión del contenido, al final del documento se incluye un glosario de términos utilizados.

Capítulo I

Importancia de los registros administrativos

Introducción

La información recabada en los registros administrativos puede ser un insumo para la generación de estadística básica, alternativa o complementaria a los censos y encuestas por muestreo, ya que se aprovecha información disponible en organismos públicos y privados¹.

Estos organismos se conciben como las fuentes de información estadística, fuentes informantes, oficinas de registro o fuentes administrativas. Sin embargo, para su utilización con fines estadísticos es necesario evaluar cuidadosamente sus características de operación, sus definiciones conceptuales, su cobertura de eventos y su desglose territorial y temporal. Igualmente, es pertinente revisar los procedimientos, las tecnologías para recabar la información y la evaluación de la calidad de los datos reportados para la generación de las estadísticas, conforme a las regulaciones administrativas de su competencia.

Así, por ejemplo, las oficinas de recaudación de impuestos en los países tienen como función sustantiva recibir los impuestos de las personas y los negocios sujetos a obligación fiscal, y también aplicar la legislación en la materia. Esta información, luego de evaluarse, puede integrarse a un proceso estadístico que permita su uso para el estudio de la estructura, características, comportamiento y tendencia de aspectos económicos, tales como: mantener un padrón de los negocios que pagan impuestos, conocer los nacimientos y muertes de negocios, el monto de los impuestos pagados y de los ingresos de los negocios, las variaciones en el tiempo, entre otros aspectos.

En este contexto, el interés por aprovechar estadísticamente los registros administrativos radica principalmente en las siguientes cualidades:

- Cobertura y continuidad de la captación de datos sobre eventos que pueden ser de interés general.
- Cobertura y desagregación geográfica de los registros, sobre todo cuando son de alcance nacional.
- Caracterización factible del fenómeno dadas las variables captadas, de acuerdo con la cobertura temática que se circunscribe al ámbito de competencia legal de los organismos o dependencias que elaboran el registro.
- Posibilidades de ampliar la cobertura conceptual de captación.
- Costos marginales del aprovechamiento estadístico ya en operación.

No obstante, estas ventajas, en el aprovechamiento estadístico de los registros administrativos inciden con relativa frecuencia los siguientes tipos de problemas:

- Complejidad y diversidad de funciones sustantivas cubiertas por la institución responsable del registro, entre las cuales no necesariamente se atienden objetivos de tipo estadístico.
- Insuficiente soporte presupuestal para las funciones estadísticas.
- Capacidad técnica limitada para mantener la continuidad de las funciones estadísticas, bajo requisitos de calidad de la información.

¹ La mayoría de los registros administrativos provienen de organismos públicos. Se originan como resultado de la actividad regulatoria que llevan a cabo. Sin embargo, en algunos casos se dispone de información por parte de organismos privados que también es susceptible de aprovecharse con fines estadísticos.

- Desinterés de las autoridades responsables del registro para su aprovechamiento estadístico.
- Dificultades de captación propias de las unidades objeto de registro y de las circunstancias en que se realiza la captación de los datos.
- Escaso control operativo y limitaciones técnicas del personal involucrado en el proceso estadístico.
- Ausencia de documentación sobre la conformación de los registros y la metodología de captación empleada.
- Modificación o alteración de la estructura de los registros a partir de la omisión y adición de variables.

Los registros administrativos constituyen un insumo en la producción estadística con aplicaciones diversas en los SEN, como:

- Generación de estadísticas. Dada la estructura del sector público en sus distintos niveles y la información con que cuentan los diferentes organismos, las necesidades de generación, integración y divulgación de información estadística oficial de interés para el SEN, se pueden clasificar en campos temáticos y asociarse con los sectores de la actividad pública para así integrar estadísticas por sectores temáticos o regiones.
- Elaboración de directorios o padrones que faciliten la generación de estadísticas por otros métodos. Con la integración de listados de organismos que realizan actividades en determinado campo de responsabilidad, estos datos pueden ser utilizados para la construcción de listados y directorios de organismos, como marcos de muestreo o como referencia para su combinación con la realización de censos.

Así es como el aprovechamiento de los registros administrativos para la producción estadística oficial es importante por la diversidad de ámbitos en los que participa el sector público. Bajo un esquema temático convencional, puede aprovecharse la información para integrar sistemas sectoriales de estadísticas oficiales, producidas por el organismo encargado de la estadística en el país o por la participación compartida entre el organismo de estadística oficial y los demás organismos públicos, según lineamientos acordados sobre la coordinación y la normativa, que faciliten la participación.

Los registros administrativos son fundamentales para conocer las brechas de género en dimensiones como demografía, educación, salud, trabajo, violencia basada en género, participación política y toma de decisiones, acceso a propiedades, migración, impacto de fenómenos climáticos, entre otras. Es relevante el uso de registros administrativos para el análisis de género mediante la producción de datos que reflejan el estado, las necesidades, las oportunidades y las contribuciones específicas de las mujeres y las niñas a la sociedad.

Por ello, a partir de los diferentes tipos de registros administrativos pueden recopilarse estadísticas de diversas temáticas, como subproducto de los procesos administrativos. Entre otros ejemplos al respecto, pueden mencionarse las estadísticas de salud, que se recopilan a partir de los registros de los hospitales; las estadísticas de empleo, que se obtienen de las oficinas de empleo; las estadísticas demográficas, que se compilan de los sistemas de registro civil, y las estadísticas de educación, que se obtienen de los informes de matrícula del ministerio de educación.

Además de las dependencias gubernamentales, es muy común que organismos privados posean información sobre las empresas que realizan algún tipo de actividad económica. Por ejemplo, puede existir una cámara de comercio que agrupe a todas las empresas que realizan algún tipo de actividad comercial, y por tanto cuenten con información, como son: directorios, datos de contacto, teléfonos, correos electrónicos, páginas web, entre otros. Esto las convierte en una fuente de actualización del SEN.

A. Fragmentación e integración

Cuando se mencionan las ventajas de los países que han desarrollado su sistema estadístico a partir de la integración de los diferentes registros administrativos, una de ellas tiene que ver con la existencia de un identificador público y único para las unidades de observación, cualesquiera sean estas.

Esa integración a partir del identificador se debe a que las diferentes dependencias que tiene cada una de las administraciones, tanto públicas como privadas, nacional y subnacionales, fragmentan las unidades de observación de acuerdo con cuál sea su objeto de registro.

En el caso de las unidades económicas existen registros administrativos sobre impuestos en los organismos impositivos nacionales y subnacionales; registros administrativos laborales en organismos relacionados con el trabajo; registros administrativos inmobiliarios o catastrales; registros administrativos sobre consumos de electricidad, gas, agua, telefonía e Internet, entre otras cosas; todos ellos en organismos, públicos o privados, nacionales y subnacionales.

En una operación estadística tradicional como es un censo económico, cada una de estas temáticas de los registros administrativos es una dimensión o variable de la unidad de observación relevada, en un formulario de captación estructurado.

La sistematización de los registros administrativos fragmentados consiste en integrarlos a partir de ese identificador único y público.

Ahora bien, esa fragmentación o dispersión se da tanto a nivel horizontal como vertical. Horizontal, cuando en una misma jurisdicción de gobierno (por ejemplo, la administración nacional) las diferentes dependencias registran cuestiones impositivas, laborales, catastrales y otras. Vertical, cuando los registros se dan en diferentes jurisdicciones o niveles de gobierno. Así existen impuestos nacionales y subnacionales, cada uno con sus respectivos organismos fiscalizadores.

Estas características de los registros administrativos deberán considerarse al momento de la integración ya que, según los países sean federales o unitarios, los estados subnacionales también poseen registros que deberán considerarse para la integración.

Capítulo II

Visión de un sistema de registros administrativos económicos

Introducción

Se presenta una visión general sobre el papel que tienen los registros administrativos y su relación con el registro estadístico de negocios, con el propósito de facilitar el desarrollo conceptual para incrementar el aprovechamiento estadístico de los registros administrativos en temas económicos, los cuales tienen importancia fundamental para:

- La construcción y actualización del registro estadístico de negocios.
- La generación de estadísticas económicas mediante:
 - La explotación tradicional del registro administrativo.
 - La generación de estadísticas mediante la vinculación de los registros administrativos con el registro estadístico de negocios.
 - La integración de un sistema de registros administrativos económicos compuesto por el registro estadístico de negocios y los registros administrativos económicos con el propósito de establecer interrelaciones para atender la Agenda 2030 para el Desarrollo Sostenible.

A. Sistema de registros base

Wallgren y Wallgren (2012) señalan que es necesario construir un sistema de registros administrativos mediante el que se asegure que todos los datos se puedan integrar de manera eficiente y se potencie su aprovechamiento estadístico. Para ello, identifican cuatro registros base:

- i) Registro de población
- ii) Registro de actividades
- iii) Registro inmobiliario
- iv) Registro empresarial

Los vínculos entre ellos constituyen la columna vertebral del sistema de registros.

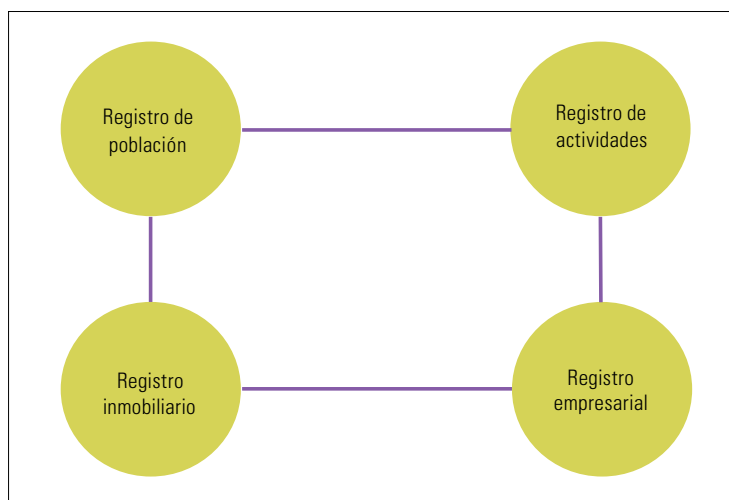
En tal sentido, señalan que cada registro base debe poseer las siguientes características:

- Tener definidos los tipos de unidades estadísticas.
- Definir conjuntos de unidades estadísticas o poblaciones de unidades estadísticas estandarizadas.
- Contener vínculos de otros registros base.
- Contener vínculos con otros registros relacionados con el mismo tipo de objeto.

Además, refieren que la información que soporta los registros base debe ser de alta calidad y estar bien documentada (véase el diagrama II.1).

Diagrama II.1

Los cuatro registros base propuestos por Wallgren y Wallgren (2012)



Fuente: A. Wallgren y B. Wallgren, *Estadísticas basadas en registros: aprovechamiento estadístico de datos administrativos*, Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

El presente documento se enfoca en la elaboración de lineamientos y recomendaciones que permitan a los países de América Latina y el Caribe la construcción o la consolidación del registro base empresarial.

B. Visión del registro base empresarial: el sistema de registros administrativos económicos

Se sugiere a las oficinas nacionales de estadística que cuenten con un sistema de registros administrativos económicos que contenga el registro estadístico de negocios, con información básica de las unidades económicas del país, que se actualice periódicamente y que permita reflejar la realidad del fenómeno de interés y extraer los marcos de muestreo requeridos en las encuestas económicas, además de generar indicadores sobre la demografía de los negocios y estudios longitudinales de las unidades económicas.

A partir de los registros administrativos es posible obtener información sobre puestos de trabajo, seguridad social, impuestos, inspección laboral, horas de trabajo y negocios registrados según tamaño e industria. Es posible generar productos estadísticos, por ejemplo: del registro de impuestos se puede tener un indicador sobre el pago mensual de impuestos para el total de negocios de un país, el cual contribuye al análisis sobre la evolución económica del país.

El análisis de esta información con enfoque de género es sumamente relevante para evidenciar los obstáculos que enfrentan las mujeres en distintas dimensiones de su autonomía económica. Por ello es importante hacer todos los esfuerzos por incrementar el número de variables disponibles en dicha temática.

Esta fuente de información permite además compilar estadísticas respecto de áreas geográficas y grupos de población más pequeños, e incluso proporcionar información longitudinal para períodos de referencia prolongados, si las unidades reciben identificaciones únicas en los registros. Ello es relevante para el estudio de trayectorias laborales de mujeres y el desempeño de negocios liderados por mujeres.

1. Objetivos del sistema de registros administrativos económicos

- Generar información estadística de forma permanente, con mayor calidad, reduciendo cargas excesivas a los informantes.
- Articular los registros administrativos con los registros estadísticos de cada país mediante un sistema de registros ágil y eficiente para la producción de estadísticas económicas combinadas.

- Establecer un marco común en el área de América Latina y el Caribe para la creación de los registros estadísticos de negocios utilizados con fines estadísticos.
- Crear un sistema homogéneo que facilite la vinculación de la información económica de todos los países y permita potenciar los análisis estadísticos económicos de la región.
- Utilizar los datos administrativos para la detección, actualización y clasificación de las unidades estadísticas y la conformación de los marcos de muestreo.
- Complementar la información estadística a partir de la imputación de datos y validación de la información, sin necesidad de emplear recursos en encuestas y formularios que encarecen los procesos estadísticos.

2. Objetivos del registro estadístico de negocios o registro estadístico empresarial como parte del sistema de registros administrativos económicos

- Detectar y construir las unidades estadísticas productivas de un país, y establecer su vínculo con el resto de los registros del sistema de registros administrativos económicos.
- Poseer una herramienta para la preparación y coordinación de encuestas económicas y la extrapolación de los resultados obtenidos a partir de ellas.
- Catalogarse como fuente de información para el estudio y análisis de la población de negocios y sus datos demográficos.
- Permitir la realización del análisis longitudinal de las unidades económicas en el tiempo.
- Tener una herramienta, estructurada por unidades estadísticas, para el traspaso de los datos administrativos a registros estadísticos; sobre todo a nivel geográfico u otro nivel de desagregación.
- Constituir una herramienta de integración y difusión de la actividad económica de un país y de la región de América Latina y el Caribe.
- Utilizarse como unidad de difusión de un directorio central de referencia que pueda proporcionar listas de empresas con ciertos datos, de acuerdo con las restricciones nacionales.

Para que la información contenida en el registro estadístico de negocios sea efectiva, es necesario que se actualice periódicamente a partir de registros administrativos que hagan referencia al mismo tipo de unidad económica (grupo de empresas, empresa o establecimiento) y que estén armonizados conceptual e informáticamente. Es decir, que además de conceptos y cuestiones informáticas, los registros administrativos deberán utilizar los mismos catálogos para las variables categóricas.

Los censos económicos y las encuestas económicas generan sus propios productos estadísticos (tabulados, sistemas de consulta, gráficos y otros). Dicho esto, en los países donde se llevan a cabo censos económicos pueden vincularse mediante el registro estadístico de negocios con registros administrativos para incrementar las variables que sobre una misma unidad económica se tienen, lo cual aumenta el potencial estadístico de la información. Dicha vinculación se realiza fácilmente mediante el identificador único que cada unidad económica tiene en el sistema de registros administrativos económicos.

La actualización del registro estadístico de negocios se hará de manera sincronizada, de acuerdo con el momento, priorizando las fuentes (censos, encuestas o los registros administrativos). En virtud de que el registro estadístico de negocios contiene información vigente e histórica de cada unidad económica, a partir de él pueden realizarse estudios longitudinales.

El rol de la oficina nacional de estadística apunta principalmente a la coordinación permanente con las entidades gubernamentales que poseen registros administrativos con el fin de cuidar que estos reúnan los criterios técnicos para su aprovechamiento estadístico, que deberá asentarse debidamente en los metadatos correspondientes.

En este contexto, cada unidad económica precisa de un identificador único que debe ser utilizado por todas las entidades gubernamentales o privadas que poseen registros administrativos para referenciar a una misma unidad económica. En consecuencia, por cada nueva unidad económica se asigna su identificador y se comparte con las diferentes entidades gubernamentales y privadas que mantengan o utilicen registros administrativos.

3. Elementos que conforman el modelo del sistema de registros administrativos económicos

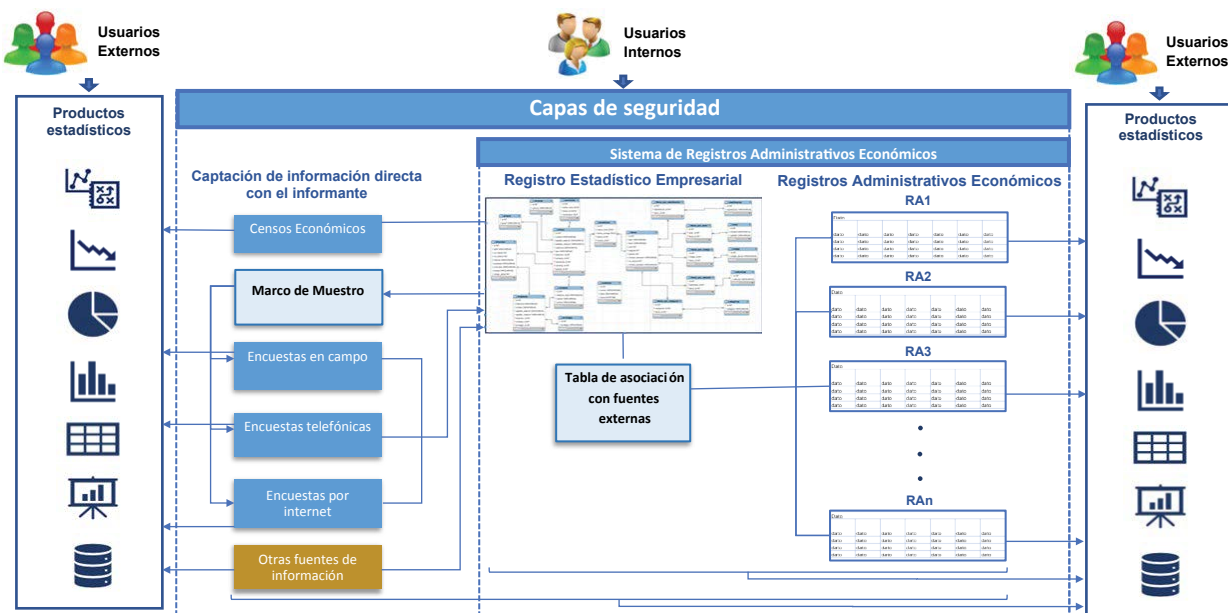
El modelo del sistema de registros administrativos económicos comprende los siguientes elementos:

- 1) Sistema de registros administrativos económicos
 - a) Registro estadístico de negocios
 - b) Registros administrativos económicos
 - c) Relación dentro de los componentes y entre ellos
- 2) Bases de datos construidas mediante la captación directa de la información
 - a) Censos económicos (si se realizan en el país)
 - b) Marcos de muestreo
 - c) Encuestas en campo
 - d) Encuestas telefónicas
 - e) Encuestas por Internet
 - f) Otras fuentes de información

Además, se establecen relaciones entre los diferentes componentes. Por ejemplo, los registros administrativos están vinculados con el registro estadístico de negocios mediante un identificador único para cada unidad económica. Los registros administrativos se encuentran vinculados entre sí mediante la tabla de asociación con fuentes externas. En el diagrama II.2 se muestra el modelo de datos del sistema de registros administrativos económicos, que incluye los registros administrativos y el registro estadístico de negocios.

Diagrama II.2

Visión del sistema de registros administrativos económicos



Fuente: Instituto Nacional de Estadística y Geografía (INEGI) de México.

C. Registro estadístico de negocios o registro estadístico empresarial

En el documento de la CEPE (2015), el registro estadístico de negocios es una base de datos estructurada de unidades económicas en un área territorial, administrada por una ONE. Esa base de datos se actualiza regularmente y se utiliza para fines estadísticos.

A continuación, se definen algunos conceptos necesarios para la mejor comprensión de lo que es un registro estadístico de negocios.

1. Base de datos estructurada

Es una base de datos en que se almacenan las características de cada unidad económica, así como las relaciones que existen entre los diferentes tipos de unidades económicas. Por ejemplo, los establecimientos que pertenecen a una misma empresa, o las empresas que son parte de un grupo empresarial.

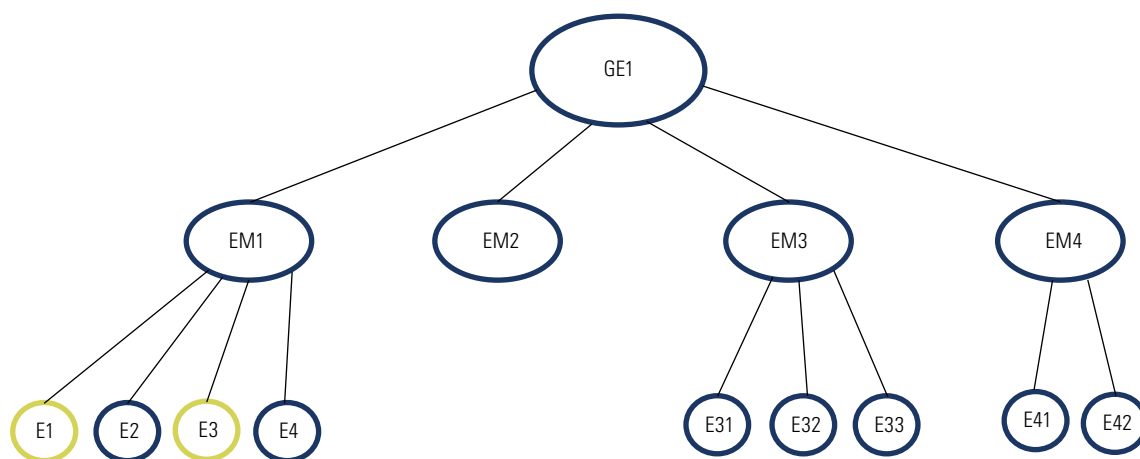
2. Identificador único para cada unidad económica

Es de gran importancia que, dentro del registro estadístico de negocios, cada unidad económica tenga un identificador único, y que se aplique para todos los tipos de unidad económica. Este permitirá hacer referencia a una unidad económica en específico, por lo que dichos identificadores no deben repetirse. Mediante el identificador único asignado a un grupo de empresas se hace referencia a todas las empresas que pertenecen a dicho grupo y, mediante el identificador único de una empresa, se hace referencia a todos los establecimientos que conforman dicha empresa.

En el diagrama II.3 se muestra un ejemplo de un grupo de empresas cuyo identificador único es GE1, que está conformado por las empresas con identificadores únicos EM1, EM2, EM3 y EM4. La empresa EM1 se forma por los establecimientos cuyos identificadores únicos son E1, E2, E3 y E4; la empresa EM3 está formada por E31, E32 y E33; mientras que los establecimientos E41 y E42 conforman la empresa EM4.

Diagrama II.3

Ejemplos de identificadores únicos



Fuente: Instituto Nacional de Estadística y Geografía (INEGI) de México.

Es importante señalar que los identificadores únicos no deben ser reutilizados para ninguna otra unidad económica, ya que harán referencia a la unidad económica aun cuando su ciclo de vida haya concluido.

En el capítulo IV se detalla la metodología para la utilización de un identificador único.

3. Regularmente actualizado

Un registro que no se actualiza con regularidad pierde rápidamente su valor, ya que las unidades económicas presentan continuos cambios tanto en sus variables económicas como en su estructura. Por ejemplo, pueden crearse nuevas unidades económicas, algunas unidades económicas pueden dejar de operar de manera definitiva, se dan fusiones o escisiones, se pueden presentar cambios en la ubicación o en la actividad económica, e incluso puede variar su valor de ventas o el número de personas ocupadas.

La frecuencia de actualización del registro estadístico de negocios depende de la disponibilidad de los datos, es decir de los registros administrativos.

4. Área territorial

Se conforma un registro estadístico de negocios para cubrir todas las unidades económicas que se ubican en un área territorial específica, la cual suele corresponder con el territorio de un país. Por lo general, una unidad económica afiliada a una empresa ubicada en el país y que se encuentra en otro país, no se encuentra en el registro estadístico de negocios.

Por otro lado, toda la información de las unidades económicas del registro estadístico de negocios debe estar concentrada en una misma base de datos. Es decir, no debe separarse por regiones, ya que no se cumpliría el objetivo del registro.

5. Base de datos administrada por una oficina nacional de estadística

Los registros administrativos son utilizados por autoridades administrativas para propósitos específicos relacionados con su función, por lo que la cobertura, los métodos de actualización y en general todo lo relacionado con el registro administrativo, es diseñado para atender asuntos administrativos. Por esta razón, la oficina nacional de estadística debe trabajar en conjunto con las entidades proveedoras de registros administrativos a fin de lograr la homologación conceptual e informática, con el propósito de que se facilite la conformación y actualización del registro estadístico de negocios, el cual necesariamente debe ser administrado por la ONE.

Uno de los objetivos del registro estadístico de negocios es proporcionar el marco de muestreo para las encuestas económicas, por lo que es fundamental que el registro se base en conceptos estadísticos, de preferencia incluidos en las recomendaciones internacionales.

D. El Modelo Genérico de Procesos Institucionales Estadísticos: marco de trabajo para el sistema de registros administrativos económicos

Con el propósito de que las oficinas nacionales de estadística cuenten con elementos suficientes para disponer de un sistema de registros administrativos económicos, en los siguientes capítulos se desarrollan lineamientos en que se considera el contexto y nivel de adelanto estadístico de los países de América Latina y el Caribe.

Para el desarrollo de los lineamientos se considera el Modelo Genérico de Procesos Institucionales Estadísticos (MGPIE) propuesto por la CEPE (2016). En este se describen y definen las fases y subprocesos necesarios para generar estadísticas oficiales.

El MGPIE provee un marco de trabajo estándar y terminología armonizada para ayudar a las organizaciones estadísticas a modernizar los procesos de producción estadística, así como compartir métodos y componentes entre los diferentes programas estadísticos. Además, puede ser usado para integrar estándares en datos y metadatos, como modelo para la documentación de procesos, la armonización de infraestructuras de cómputo estadístico y la facilitación de un marco de referencia para mejoras y evaluaciones del proceso de calidad.

El MGPIE debe ser aplicado e interpretado de manera flexible. No es un marco rígido en el que todos los pasos se deban seguir en un orden estricto, sino que se identifican los posibles pasos dentro del proceso estadístico y las interdependencias entre ellos.

Aunque la presentación del MGPIE sigue una secuencia lógica de pasos en la mayoría de los procesos estadísticos, los elementos del modelo pueden presentarse en diferente orden, de acuerdo con las diferentes circunstancias. Además, algunos subprocesos serán revisados en varias ocasiones, formando ciclos iterativos, sobre todo dentro de las fases de procesamiento y análisis.

El MGPIE debe ser visto, por lo tanto, más bien como una matriz en la cual hay muchos caminos posibles. De esta manera, el MGPIE pretende ser lo suficientemente genérico para poder ser aplicable en sentido amplio y para promover una vista estandarizada del proceso estadístico, sin llegar a ser demasiado restrictivo o abstracto y teórico.

El MGPIE considera la realización de ocho fases: especificación de necesidades, diseño, construcción, recolección, procesamiento, análisis, difusión y evaluación, tal y como se observa en el diagrama II.4. El MGPIE ha sido adaptado por algunas oficinas nacionales de estadística según sus propias necesidades de organización de la producción estadística que tienen a cargo.

Diagrama II.4

Fases y subprocesos del Modelo Genérico de Procesos Institucionales Estadísticos (MGPIE)



Fuente: Comisión Económica para Europa (CEPE), "Modelo genérico del proceso estadístico GSBPM (versión 1.0 en español, septiembre 2016)", Ginebra, 2016.

Cada fase se divide en subprocesos. En total existen 44 subprocesos. No necesariamente todos se aplican para el desarrollo de estos lineamientos, debido a que se orientan a aspectos muy particulares que abarcan otros métodos tradicionales para elaborar estadísticas como las encuestas por muestreo.

En los siguientes ocho capítulos, se desarrollan recomendaciones que las ONE pueden considerar para la integración de un sistema de registros administrativos económicos. A continuación, se enuncian los aspectos que se consideran en cada fase.

- Fase 1: Especificación de necesidades (capítulo III)
 - Identificación de necesidades
 - Consulta y confirmación de necesidades

- Definición de los objetivos
- Identificación de conceptos
- Comprobación de la disponibilidad de los datos
- Elaboración del plan general
- Fase 2: Diseño (capítulo IV)
 - Diseño conceptual
 - Diseño de las descripciones de las variables
 - Diseño de la recolección
 - Diseño del procesamiento y análisis
 - Diseño de los sistemas de producción y flujos de trabajo
- Fase 3: Construcción (capítulo V)
 - Construcción de instrumentos de recolección
 - Construcción o mejora de componentes
 - Construcción o mejora de componentes de difusión
 - Configuración de flujos de trabajo
 - Pruebas del sistema de producción
 - Prueba piloto del proceso estadístico
 - Finalización del sistema de producción
- Fase 4: Recolección (capítulo VI)
 - Preparación de la recolección
 - Ejecución de la recolección
 - Cierre de la recolección
- Fase 5: Procesamiento (capítulo VII)
 - Integración de datos
 - Clasificación y codificación
 - Revisión y validación
 - Edición e imputación
 - Derivación de nuevas variables y unidades
 - Cálculo de agregaciones
 - Finalización de los archivos de datos
- Fase 6: Análisis (capítulo VIII)
 - Preparación de borradores de resultados
 - Validación de los resultados
 - Interpretación y explicación de los resultados
 - Aplicación del control a la divulgación
 - Finalización de resultados
- Fase 7: Difusión (capítulo IX)
 - Actualización de sistemas de salida
 - Generación de productos de difusión
 - Gestión de la publicación de productos de difusión
 - Promoción de productos de difusión
 - Gestión de soporte a las personas usuarias
- Fase 8: Evaluación (capítulo X)
 - Concentración de insumos para la evaluación
 - Evaluación
 - Determinación de un plan de acción

Como puede observarse, en este documento no se consideran los subprocesos relacionados con el muestreo, pues no se aplican para el sistema de registros administrativos económicos.

Capítulo III

Especificación de necesidades

Introducción

Como punto de partida, se debe realizar un análisis de las necesidades de información de las personas usuarias, tanto internas como externas, que permitan la estructuración de una estrategia que satisfaga la demanda de información para la conformación del sistema de registros administrativos económicos, cuyo componente principal es el registro estadístico de negocios. Asimismo, esta estructura debe ser flexible para que, en el tiempo, pueda ajustarse a la evolución de necesidades y prioridades dado un entorno cambiante.

La fase de especificación de necesidades consiste en tener un panorama de cuáles son las necesidades de información que un sistema de registros administrativos económicos debería de satisfacer en cada país de América Latina y el Caribe. Ello permitiría mejorar la producción de estadísticas económicas, para así satisfacer la actual y creciente demanda de las personas usuarias. Se daría a las ONE la posibilidad de acoplarse a las disposiciones actuales y futuras de generación de información, tanto nacionales como internacionales.

El alcance de la fase de especificación de necesidades permitirá conocer por qué se decide crear un sistema de registros administrativos económicos. También permitirá determinar los antecedentes a nivel nacional e internacional que conllevan a la detección de la necesidad de crearlo, los requerimientos de información que existen actualmente en cuanto a registros empresariales y los alcances metodológicos que deben tomarse en cuenta en el diseño del sistema.

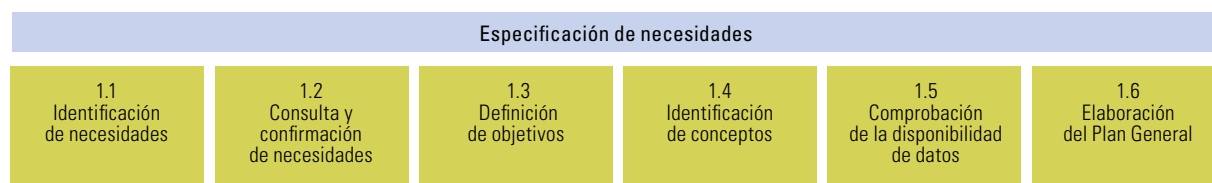
Es importante en esta fase analizar en particular las necesidades de datos para el desarrollo de políticas relacionadas con la igualdad de género y la autonomía de las mujeres e identificar un marco de indicadores de seguimiento con claridad de las fuentes de datos. Además, es importante la consulta no solo con los ministerios sectoriales, sino también con los mecanismos de promoción del avance de las mujeres y las organizaciones feministas. Ello permitiría identificar las brechas entre los datos requeridos y los datos disponibles, así como las acciones para cubrir estas brechas.

A. Subprocesos de la especificación de necesidades

La especificación de necesidades se divide en los seis subprocesos que se indican en el diagrama III.1, los cuales permiten: i) identificar la necesidad de la estadística; ii) confirmar con más detalle las necesidades estadísticas de las partes interesadas; iii) establecer los objetivos de alto nivel para los resultados estadísticos; iv) identificar los conceptos y variables relevantes para los que se requiere la información; v) comprobar el grado en que las fuentes de datos actuales pueden satisfacer estas necesidades, y vi) preparar la argumentación necesaria con miras a obtener la aprobación para producir estadísticas.

Diagrama III.1

Subprocesos de la fase de especificación de necesidades



Fuente: Comisión Económica para Europa (CEPE), "Modelo genérico del proceso estadístico GSBPM (versión 1.0 en español, septiembre 2016)", Ginebra, 2016.

1. Identificación de necesidades

Este subproceso consiste en realizar un análisis de las necesidades de información. Estas pueden ser producto de una necesidad a nivel de país o institucional, o bien una solicitud explícita de algún organismo internacional o de un ente gubernamental. La finalidad es obtener información que permita la toma de decisiones de política pública y privada.

Esta identificación de necesidades debe contar con una visión de país que permita la conformación de un único sistema de registros administrativos económicos basado en un registro estadístico de negocios y en registros administrativos que sea mantenido por la oficina nacional de estadística de cada país. Su objetivo sería evitar la duplicación de esfuerzos, a costos menores, con información más oportuna y eficiente. El sistema de registros administrativos económicos se utilizaría para obtener un marco estadístico para las encuestas o estudios económicos, de manera que no exista la necesidad de realizar consultas adicionales a la población empresarial, sino que se puedan satisfacer mediante los registros administrativos ya existentes en el país.

Por esa razón deben incluirse las necesidades de información de todas las personas usuarias (externas e internas) con el fin de evitar la duplicación de esfuerzos y, además, no realizar las mismas solicitudes de registros administrativos empresariales a una determinada institución por parte de la ONE. De este modo la institución podrá trabajar de forma coordinada y bajo una misma visión, teniendo claro los criterios de homogenización de la información y las clasificaciones, entre otros.

Por otro lado, el sistema debe asegurar la comparabilidad internacional. Para ello, en este proceso de identificación de necesidades debe incluirse la investigación de la producción de estadísticas similares o prácticas implementadas por la ONE y los organismos nacionales e internacionales. Esta información podría servir de referencia en la propuesta de generación de la estadística requerida (por ejemplo, la revisión bibliográfica de organismos internacionales sobre los requisitos mínimos de información que debe cumplir un registro estadístico de negocios basado en registros administrativos).

2. Consulta y confirmación de necesidades

Resulta vital conocer cuáles son las necesidades de información de las diversas personas usuarias potenciales que se han identificado previamente, para así tener una hoja de ruta de los aspectos más importantes a considerar en la fase de diseño del sistema de registros administrativos económicos. En la medida de lo posible, deberá incluirse como parte de las necesidades la desagregación de las variables por sexo, lo que permitirá atender necesidades internas y externas.

Un error que suele cometerse es considerar solamente las necesidades externas (ya sea a nivel de país o a nivel internacional). Las necesidades internas de la ONE son tan importantes como estas, ya que puede haber varios procesos o departamentos que sean las personas usuarias o entidades proveedoras del sistema de registros administrativos económicos.

Al realizar la consulta internamente (en primera instancia) se podrá evitar la duplicación de esfuerzos, costos, entre otros. Además, ello permitirá una mejor coordinación y, por ende, potenciar el uso de los registros administrativos².

Con respecto a la consulta externa, se refiere a instituciones que conforman el SEN, y que en algunos países puede ocurrir que estén establecidas por ley. Por ejemplo, en el caso de Costa Rica, en la reciente Ley núm. 9694 (aprobada el 30 de mayo de 2019), se establece en la sección V, Acceso de los usuarios a la información y consulta de sus necesidades, en su artículo 24, “[l]as instituciones del SEN deberán identificar claramente a los usuarios principales de las estadísticas a su cargo y consultarlos regularmente sobre la pertinencia de los resultados publicados, la precisión y calidad observada, las nuevas necesidades de información, el nivel de acceso a las informaciones estadísticas y considerar sus opiniones en las mejoras del servicio estadístico” (Asamblea Legislativa de Costa Rica, 2019).

Además, en el artículo 35 se establece que “[e]l INEC consultará a los usuarios, por los medios que considere pertinentes, sobre las estadísticas actuales y las nuevas necesidades de información. Para desarrollar nuevas estadísticas oficiales, se analizará la viabilidad y el costo de satisfacer tales requerimientos, y su elaboración quedará sujeta a la obtención del financiamiento adicional para el INEC o al aporte de los interesados, las que deberán ser incorporadas al PEN” (Asamblea Legislativa de Costa Rica, 2019).

Las necesidades de información deben ser actualizadas con las personas usuarias de la información. Por ejemplo, conviene detectar aquellas que pudieron captarse en el pasado y no se captaron, o aquellas que en algún momento fueron captadas y se dejaron de consultar. Por otro lado, es necesario analizar los requerimientos de las personas usuarias y la factibilidad de atenderlas, confirmar aquellos que pueden atenderse y en qué fecha, priorizar dichos requerimientos y los que no pueden atenderse, y documentar la justificación del por qué (un ejemplo de cómo documentar se presenta más adelante, en el punto 5 del anexo A1).

Es importante tener presente que este proceso de consulta (interna o externa) puede requerir más de un contacto de la ONE con los actores identificados (personas usuarias o entidades proveedoras de información), por lo que el acercamiento inicial con cada parte interesada es clave para el proceso.

3. Definición de objetivos

Una vez realizada la consulta y confirmación de las necesidades de información de los diversos actores, este subproceso requiere convertir los resultados que se espera alcanzar en objetivos específicos que permitan satisfacer las necesidades identificadas en la consulta.

Se trata de fijar un objetivo que permita la conformación de un sistema de registros administrativos económicos, de manera que se puedan potenciar las ventajas de su uso. En su caso, estas pueden ser un complemento para otras fuentes de información, de modo que se logre una producción de más y mejores resultados estadísticos de calidad, oportunidad y eficiencia.

4. Identificación de conceptos

Consiste en una primera identificación de los conceptos que se utilizarán como parte de la temática del sistema de registros administrativos económicos, en el caso de los registros administrativos económicos que sean de interés y respecto de los cuales se disponga de información suficiente y de calidad. Dichos conceptos serán insumos importantes para la normalización de definiciones empresariales con fines estadísticos, entre otros.

En el caso del registro estadístico de negocios, es importante disponer de conceptos que hagan referencia a temáticas como identificación, ubicación, tamaño y actividad económica. Además, pueden incluirse los de demografía empresarial, exportaciones e importaciones, temas ambientales, entre otros. Para abarcar esos temas,

² Lo ideal es que el uso, acceso, resguardo, seguridad, entre otros, de los registros administrativos esté normado y estandarizado en la oficina de estadística, de manera que se pueda asegurar un uso eficiente y un compromiso con las entidades proveedoras de la información.

es necesario disponer de diversos registros administrativos e identificar los conceptos temáticos. Así, por ejemplo, si se está trabajando con el registro administrativo de la Oficina de Recaudación de Impuestos, algunos de los conceptos temáticos podrían ser: ingresos, impuestos, régimen fiscal, persona natural (física) o persona moral (jurídica), entre otros. La identificación de estos conceptos será de gran utilidad para determinar las variables que se pueden utilizar del registro administrativo.

5. Comprobación de la disponibilidad de datos

Se debe investigar cuáles son los registros administrativos existentes que podrían ser los insumos clave para obtener la información requerida para la generación de las estadísticas económicas.

Para dicha investigación se puede realizar un inventario de los registros administrativos económicos existentes, como se describe en el inciso 4.1.3. El SEN de Costa Rica, entre otros, cuenta con un inventario de operaciones estadísticas que podría ser un ejemplo. En los cuadros III.1 y III.2 se presentan dos ejemplos de ficha metodológica del registro administrativo.

Cuadro III.1

Ejemplo de ficha metodológica del registro administrativo

Costa Rica: ficha metodológica del registro administrativo	
Nombre del registro administrativo	
Nombre de la institución responsable	
Nombre de la unidad administrativa que elabora el registro administrativo	
Objetivo general del registro administrativo	
Sector o tema	
Subsector o subtema	
Cobertura geográfica	
Desagregación geográfica	
Unidades de análisis	
Principales variables	
Periodicidad de la captura de datos	
Periodicidad de la difusión	
Medios de difusión	
Marco legal	
Nombre de la persona encargada del registro administrativo	
Puesto	
Correo electrónico	
Teléfono	
Sitio web	

Fuente: Instituto Nacional de Estadística y Censos (INEC) de Costa Rica, *Informe de resultados del Inventario Nacional de Operaciones Estadísticas*, San José, 2018.

Cuadro III.2

Ejemplo de ficha metodológica del registro administrativo

México: esquema de temas, categorías y variables a captar en el inventario de registros administrativos	
Registros administrativos identificados	
Código de identificación	
Nombre oficial	
Otras denominaciones	
Fecha de incorporación al inventario de registros administrativos	
Tipo de registro	
Registros complementarios	
Condición de operación	

Cuadro III.2 (conclusión)

México: esquema de temas, categorías y variables a captar en el inventario de registros administrativos	
Registros administrativos con almacenamiento de la información en medios electrónicos	
Formato de la base de datos	
Registros administrativos con consulta de su información por Internet	
Dirección URL	
Registros administrativos aprovechados para generar estadísticas	
Nombre de la institución que genera las estadísticas	
Tipo de las personas usuarias de la información estadística	
Nombre de los proyectos estadísticos relacionados	
Medio de presentación de las estadísticas	
Medio físico de almacenamiento	
Registros administrativos vinculados con proyectos estadísticos	
Condición de inclusión en el registro estadístico de negocios	
Registros administrativos que presentan estadísticas en Internet	
Dirección URL	
Registros administrativos con otras instituciones participantes	
Tipo de participación	
Registros administrativos en operación	
Nombre de la institución responsable	
Siglas de la institución	
Ámbito de gobierno	
Unidad administrativa responsable	
Funcionario responsable	
Cargo del responsable	
Domicilio oficial	
Teléfono	
Correo electrónico	
Otras instituciones participantes	
Nombre de las bases normativas de operación	
Fecha de la última publicación en el Diario Oficial de la Federación	
Nombre de los trámites o servicios relacionados	
Temática principal, subtema, apartado	
Objetivo	
Sujetos de registro	
Variables que caracterizan a los sujetos de registro	
Cobertura de los sujetos de registro	
Cobertura geográfica	
Desglose geográfico	
Disponibilidad de la información	
Corte temporal	
Medio físico de registro	
Nombre del formato o expediente de registro	
Frecuencia de captación	
Medio físico de almacenamiento	
Instancia donde se procesan los datos	
Nombre de los documentos conceptuales disponibles	
Nombre de los glosarios disponibles	
Nombre de los manuales disponibles	
Nombre de los instructivos disponibles	
Tipo de las personas usuarias de la información del registro	
Medio de difusión o consulta de la información del registro	
Aprovechamiento de la información para generar estadísticas	

Fuente: Instituto Nacional de Estadística y Geografía (INEGI) de México.

Una vez que se hayan identificado los registros administrativos económicos disponibles, se debe realizar un análisis del marco legal de cada fuente, acceso y uso de la información, calidad, periodicidad y otras características. El objetivo es identificar si son los más adecuados, o priorizarlos ante otras fuentes. Por ello se recomienda, a partir de los insumos citados, elaborar un inventario de registros administrativos económicos y sugerir las variables que debería contener.

Se debe trabajar de igual manera en acuerdos o documentos formales, como convenios o cartas de entendimiento (en caso de no existir una ley que establezca la obligatoriedad de la entrega del registro), que respalden la entrega de los datos (incluye formatos, instrumentos de entrega, periodicidad, entre otros).

6. Elaboración del plan general y documentación

La documentación de un proceso es importante, dado que evidencia las acciones realizadas durante esta fase, de manera que se pueda utilizar para la aprobación por las autoridades correspondientes y, por ende, para la implementación del proceso u operación estadística.

Este documento debe incluir una descripción del estado actual del sistema de registros administrativos económicos (si ya existe), con información sobre cómo se producen las estadísticas actuales. Se debe destacar cualquier área de oportunidad que sea necesario abordar. Se ha de incluir una propuesta de solución que detalle cómo se desarrollará el proceso para producir las estadísticas nuevas, o las ya existentes pero que requieren mejoras o ajustes y, por último, una evaluación de los costos y beneficios.

Capítulo IV

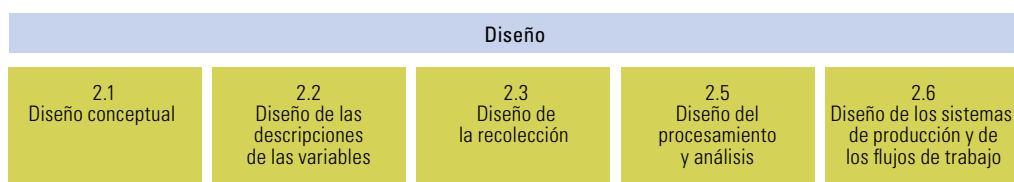
Diseño

Introducción

Esta fase tiene por objeto diseñar los productos de información estadística que atenderán las necesidades estructuradas de información, determinadas de acuerdo con los elementos documentales recabados en la fase anterior. Es decir, se pretende definir las salidas, conceptos, metodologías, instrumentos de recolección de la información, protocolos y canales de intercambio, así como las estrategias generales para las siguientes fases del modelo MGPIE (véase el diagrama IV.1).

Diagrama IV.1

Subprocesos incluidos en la fase de diseño



Fuente: Comisión Económica para Europa (CEPE), "Modelo genérico del proceso estadístico GSBPM (versión 1.0 en español, septiembre 2016)", Ginebra, 2016.

Nota: Se omite el subproceso 2.4, "Determinación del marco muestral y tipo de muestreo", que no se aplica a registros administrativos, debido a que se orienta a aspectos que abarcan otros métodos tradicionales para elaborar estadísticas, como el caso de las encuestas por muestreo.

A. Diseño conceptual

En la fase de diseño, los propietarios del sistema de registros administrativos económicos definirán las salidas, conceptos, metodologías, instrumentos de captación, protocolos y canales de intercambio. También definirán las estrategias generales para el desarrollo de las fases de construcción, recolección, procesamiento, análisis y difusión, la modalidad metodológica de ejecución y otros aspectos que se consideren relevantes dentro del proceso de producción de información.

No se debe pretender desde un inicio que el sistema de registros administrativos económicos brinde toda la información requerida del conjunto de términos y variables necesarios para el análisis estadístico de la economía. Es mejor considerar metas alcanzables en un corto período de tiempo, sobre la base de la información de los registros administrativos y estadísticos accesibles, que reúnan los requisitos de calidad esperados, y posteriormente definir un plan de mejoras sobre la base de objetivos y metas para alcanzar a largo plazo.

Los requisitos para un registro estadístico de negocios pueden realmente ser diferentes a los de un registro administrativo de negocios. Aun si el registro estadístico de negocios se alimenta de fuentes de información que constituyen registros administrativos, los conceptos, las características, metodologías, mantenimiento de las bases de datos, entre otras cosas, del registro estadístico de negocios necesitan basarse en conceptos estadísticos, preferentemente en recomendaciones internacionales.

Una característica importante del sistema de registros administrativos económicos es que solo se crea para ser usado con propósitos estadísticos. No puede servir para otros propósitos, en especial para propósitos

administrativos. Esto se debe a la necesidad de mantener la confidencialidad de la información individual recabada por los sistemas informativos. Se debe garantizar que los datos relacionados con personas físicas individuales o con negocios recabados por los sistemas informativos para su recopilación estadística sean estrictamente confidenciales y sean usados exclusivamente con propósitos estadísticos.

1. Definición del conjunto de objetos del registro

El marco teórico principal sobre el que se debe basar el diseño para las unidades de observación estadísticas del sistema de registros administrativos económicos y el registro estadístico de negocios como centro del sistema, es el Sistema de Cuentas Nacionales (SCN) 2008 (SNA 2008, por sus siglas en inglés). En este pueden encontrarse las referencias a términos y definiciones sobre los que se debe consolidar el registro estadístico.

Tanto las unidades de observación básicas del sistema de registros administrativos económicos como las del registro estadístico de negocios como parte de ese sistema responden a la misma clasificación. Pueden ser:

- Unidades institucionales, que son las unidades económicas capaces de realizar la gama completa de transacciones, de poseer activos y de contraer pasivos en su propio nombre.
- Las unidades institucionales residentes en la economía se agrupan en cinco sectores mutuamente excluyentes, compuestos por los siguientes tipos de unidades:
 - i) Sociedades no financieras: son unidades institucionales que se dedican principalmente a la producción de bienes y de servicios no financieros de mercado.
 - ii) Sociedades financieras: son unidades institucionales que se dedican principalmente a la prestación de servicios financieros, incluida la intermediación financiera.
 - iii) Gobierno general: incluidos los fondos de seguridad social, consiste en unidades institucionales que, además de cumplir con sus responsabilidades políticas y con su papel en la regulación económica, producen servicios (y posiblemente bienes) no de mercado para el consumo individual o colectivo, y redistribuyen el ingreso y la riqueza.
 - iv) Instituciones sin fines de lucro que sirven a los hogares: son entidades jurídicas dedicadas principalmente a la producción de servicios no de mercado para los hogares o para la comunidad, y cuyos recursos principales son contribuciones voluntarias.
 - v) Hogares: son unidades institucionales formadas por un individuo o grupo de individuos. Todas las personas físicas de la economía deben pertenecer a un único hogar. Las funciones principales de los hogares son proporcionar mano de obra, llevar a cabo el consumo final y, en cuanto empresarios, producir bienes y servicios no financieros (y posiblemente financieros) de mercado. Las actividades empresariales de un hogar son las realizadas por las empresas no constituidas en sociedad que permanecen dentro de dicho hogar, excepto en determinadas circunstancias.

El sistema de registros administrativos económicos solo va a tener en cuenta como unidades de observación de las estadísticas básicas de los registros administrativos aquellas que tienen producciones de bienes o servicios, las que se corresponden con:

- sociedades no financieras;
- sociedades financieras;
- gobierno general, e
- instituciones sin fines de lucro que tengan producción de bienes y servicios.

No forman parte del sistema de registros administrativos económicos las unidades administrativas clasificadas como instituciones sin fines de lucro que no tengan ningún tipo de producción, ni los hogares.

Las unidades institucionales que realizan actividades económicas de producción de bienes y servicios son comúnmente clasificadas como empresas. Las empresas pueden estar vinculadas, formando grupos empresariales, y a su vez una empresa puede estar dividida en otras unidades subordinadas a ellas. Estas pueden ser constituidas en unidades estadísticas dentro del registro estadístico de negocios. Se pueden observar diferentes clasificaciones de las unidades estadísticas referidas a las empresas atendiendo a la forma de división:

- i) Unidades institucionales
- ii) Grupos empresariales
- iii) Empresas
- iv) Establecimientos (unidades locales por actividad)
- v) Unidades locales
- vi) Unidades por tipo de actividad
- vii) Otras (especificar)

En el inciso vii (otras) se pueden encontrar las instituciones sin fines de lucro que producen bienes o servicios u otras unidades no constituidas en empresas.

Para la clasificación de las unidades estadísticas de subordinación empresarial, puede utilizarse el modelo del cuadro IV.1.

Cuadro IV.1

Modelo de clasificación de las unidades estadísticas de subordinación empresarial

		Divididas por localidad	
		No	Sí
Divididas por actividad	No	Empresa	Unidades por localidad
	Sí	Unidades por actividad	Establecimientos (unidades por tipo de actividad y por localidad)

Fuente: Comisión Europea y otros, *Sistema de Cuentas Nacionales 2008*, Nueva York, 2016; Naciones Unidas, *Clasificación Industrial Internacional Uniforme de todas las actividades económicas (CIIU): Revisión 4 (ST/ESA/STAT/SER.M/4/Rev.4)*, Nueva York, 2009.

2. Número de identificación único para las unidades de observación que sea de uso público

El número identificador de la unidad de observación tiene la utilidad de establecer un vínculo con la totalidad de la información que tributen todas las fuentes de información del sistema de registros administrativos económicos, ya sean administrativas o estadísticas. Ello permite realizar los análisis económicos propios y garantizar el traspaso de información desde los registros administrativos hasta registros estadísticos. Por ejemplo, se pueden lograr análisis estructurales de la actividad económica vinculando estadísticas y datos procedentes de registros administrativos de producción, aduanas, tributos, finanzas, trabajo, salarios, entre otros.

El registro estadístico de negocios contiene todas las unidades que realizan actividades económicas, relacionadas con un número de identificación único, también denominado número clave.

El diseño y uso del número de identificación único debe definirse particularmente en cada país, cumpliendo siempre con la premisa de que sea un código único para cada unidad de observación del registro estadístico de negocios, aun cuando la unidad desaparezca físicamente. Se recomienda que sea un código alfanumérico, con uno o dos dígitos que identifiquen el tipo de unidad registrada de acuerdo con una de sus clasificaciones más importantes. Los criterios de clasificación pueden seleccionarse entre los más significativos, como tipo de actividad, sector de propiedad, producto principal, entre otros, y añadir a continuación cuatro o cinco dígitos que respondan a un número consecutivo, aunque la estructura no es relevante.

El registro estadístico de negocios debe ser capaz de contener atributos que definan la unidad de observación, en particular aquellos aspectos de su caracterización que no sean objeto de cambios frecuentes. Puede tener componentes que la identifiquen en sus características fundamentales, como los que se enumeran a continuación:

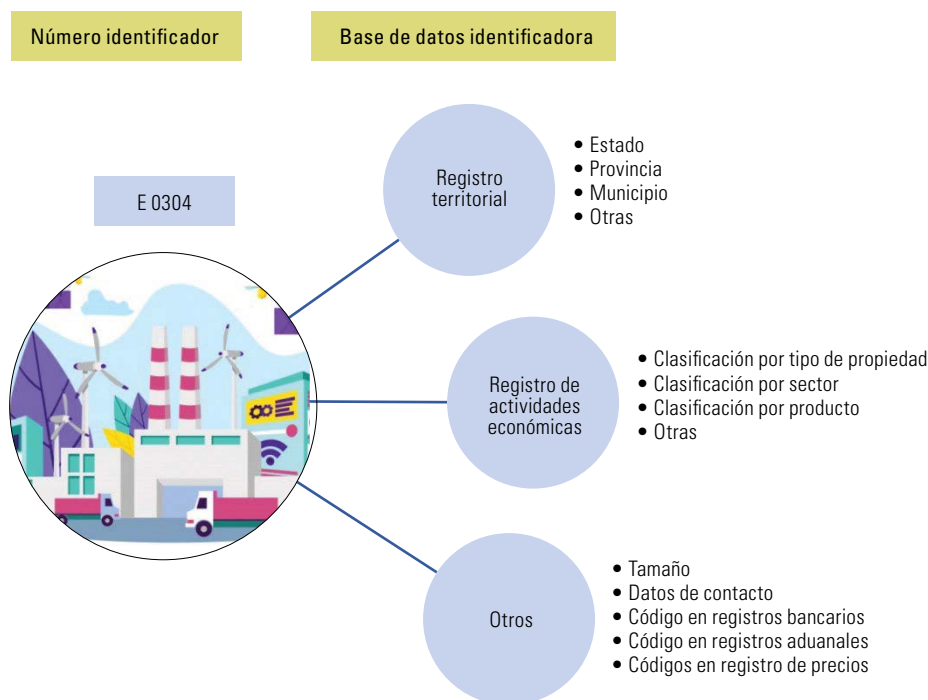
- Nombre comercial: nombre con el que se conoce la unidad para realizar sus relaciones económicas.
- Dirección de residencia: la residencia de cada unidad de observación es el territorio económico con el que tiene la relación más fuerte, en otras palabras, su centro predominante de interés económico. Cada unidad de observación es residente de un territorio económico único. Cada empresa miembro de un grupo de empresas es residente de la economía en que se localiza, y no de la economía donde se localiza la oficina central del grupo empresarial.
- La dirección de residencia debe proveer de forma independiente los datos por cada división territorial, de acuerdo con el marco geográfico legal establecido por el país y, si es posible, la georreferenciación de cada unidad. De esta forma se pueden obtener análisis a diferentes niveles territoriales.
- Datos de contacto: incluidos teléfono, correo electrónico, fax y dirección postal.
- Sector institucional clasificado según el SCN 2008.
- Clasificación del sector empresarial según la propiedad (estatal, privado, mixto, o también, nacional, extranjero, mixto).
- Tipo de producción (productores de mercado, productores por cuenta propia o productores no de mercado) según el SCN 2008.
- Productos fundamentales (bienes o servicios) (mínimo 1) (Clasificación Central de Productos (CPC) versión 2).
- Clasificación de las actividades económicas que desempeña.
- Categorización de acuerdo con el tamaño (micro, pequeña y mediana, grande).
 - Fecha de surgimiento o creación de la unidad.
 - Fecha de inicio de operaciones.
 - Fecha de cese definitivo de las actividades.
 - Estado operativo (por ejemplo, alta, baja, paro, quiebra) que caracteriza el nivel de funcionalidad en el momento actual.
 - Unidad a que pertenece y las subordinadas (vinculadas con el número identificador único correspondiente, de forma que se puedan vincular en la base de datos).
 - Categoría por su aporte al PIB.
 - Número identificador único que se utilizó para la constitución de la empresa.
 - Números identificadores únicos de la unidad otorgada durante la vida de la empresa debido a modificaciones en la inscripción original respecto de la cual la empresa puede ser considerada una continuidad de otras inscriptas previamente.
 - Número de empleados por sexo y categoría salarial.

Al abrir cada unidad de observación en el registro estadístico de negocios, desde la clave identificadora se podrá conocer, además de la información del número identificador único y el nombre de la unidad, la información necesaria y disponible al respecto. En el diagrama VI.2 se muestra un ejemplo de esta caracterización.

Estos atributos pueden ser parte del propio registro estadístico de negocios o estar contenidos en otra base de datos. Es decir, el registro estadístico de negocios puede contener solo un listado de las unidades estadísticas con su número de identificación. Quedan documentadas en otra base de datos las variables asociadas al registro y se mantienen accesibles en reportes de consulta por su número identificador. Estas variables identificadoras son de utilidad para hacer análisis estadísticos a partir de la combinación de las informaciones económicas que se reciban, asociadas al número identificador de la unidad básica registrada.

Diagrama IV.2

Esquema de la base de datos identificadora para la unidad de observación



Fuente: Oficina Nacional de Estadística e Información (ONEI) de Cuba.

En desarrollos futuros del sistema de registros administrativos económicos, posterior a la creación de los registros estadísticos de negocios en todos los países de la región, la base de datos identificadora de las unidades de observación de cada país podría ser de utilidad para considerar establecer un diseño de número identificador único a nivel regional, que utilice las variables identificadoras como fuentes de diseño para los dígitos del identificador único. Lo anterior sería de utilidad para obtener más información sobre las unidades registradas por su número identificador, sin tener necesidad de acceder a los datos de la identificación de la unidad para cada país mediante su registro identificador. No obstante, tiene la desventaja de que algunas de estas variables pueden sufrir cambios durante la vida activa de la unidad, lo que complicaría los análisis estadísticos seriados.

El código único de identificación debería asignarse desde la creación de la entidad, lo que varía en dependencia de la legislación de cada país. Algunos países utilizan el código de identificación tributaria o se asigna uno por la ONE una vez incorporado al registro estadístico de negocios. También pueden utilizarse correlaciones entre los diferentes códigos de identificación, por ejemplo, entre el código del registro estadístico de negocios y el número de identificación tributaria.

El reconocimiento legal de este código de identificación única o la sinergia entre las oficinas nacionales de estadística y las diferentes instituciones donde de forma obligatoria deben registrarse las unidades institucionales como las oficinas tributarias, los bancos y las aduanas, es lo que garantiza el uso generalizado de este código de identificación y contribuye a su aplicación y generalización dentro del sistema de registros administrativos económicos del país.

La clave o número identificador único asociado a cada unidad de observación, además de tener la función de vincular la información asociada, contribuye al análisis estadístico económico con los otros registros administrativos, sean o no objetivo económico. También puede ser utilizada por estos registros administrativos para vincular la información correspondiente a sus funciones con la unidad económica básica, con lo que se crea un flujo informativo en ambos sentidos.

3. Elaboración de un inventario de las fuentes

En general, los registros administrativos no son creados para responder a propósitos estadísticos. De ahí la importancia de evaluar su potencial como una eventual fuente de información para generar estadísticas dentro del sistema de registros administrativos económicos o para servir como información complementaria o de contraste en otras operaciones estadísticas del sistema de registros administrativos económicos.

Cada país debe realizar un estudio para seleccionar las fuentes de información idóneas para obtener la información estadística requerida por el sistema de registros administrativos económicos. El enfoque de selección, en su primera fase, consiste en determinar el valor de las fuentes administrativas para suministrar información en la elaboración del sistema de registros administrativos económicos en cuanto a las necesidades planteadas. En otras palabras, se trata de determinar cuán efectivamente se pueden usar los datos. Entre otros aspectos que pueden definirse previamente al análisis de calidad de la fuente administrativa cabe mencionar:

- i) Legalidad e institucionalidad: el registro administrativo está amparado por leyes y resoluciones que regulan su uso con carácter obligatorio. Tiene procedimientos que definen los procesos de actualización y mantenimiento de los datos, así como otros descriptivos. Estos procedimientos son evaluados periódicamente con vistas a mejorar la calidad de registro administrativo.
- ii) Cobertura geográfica: se define hasta qué punto de la división territorial (establecida en el país por ley) se van a agrupar las unidades geográficamente (municipalidad, condado, provincia, estado, entre otras) y se comprueba que la fuente administrativa cumpla con la cobertura deseada.
- iii) Cobertura empresarial: lo mejor sería que las fuentes administrativas seleccionadas puedan cubrir todos los niveles y clasificaciones de las unidades productoras del país, basadas en las clasificaciones reguladas por los institutos estadísticos nacionales, como parte de su función: clasificaciones internacionales y nacionales, tipo de actividad, sectores, tamaño y otras clasificaciones económicas y territoriales. Sería aceptable, sin embargo, escoger aquellas que los registros administrativos puedan abarcar si sus variables cumplen con los requisitos de calidad establecidos para los datos estadísticos (relevancia, exactitud, oportunidad y puntualidad, accesibilidad y claridad, comparabilidad, coherencia y completitud) (véase la sección IV. B, sobre diseño de las descripciones de las variables).
- iv) Rango: se pueden elegir las fuentes administrativas que brinden información total de la población de unidades de observación recopiladas en el registro estadístico de negocios, o información parcial siempre que brinden el porcentaje representativo del total de unidades de observación, que se pueda asumir como válida en el país o que se pueda completar con la información de otros registros.
- v) Integridad referencial: un número de identificación único para cada unidad, que pueda ser homologado con uno propio del registro estadístico de negocios. Se utilizan clasificadores estandarizados que pueden ser homologados con los utilizados internacionalmente. Las bases de datos están documentadas por metadatos estandarizados.
- vi) Tecnologías informáticas: se evalúa, entre los registros administrativos que aporten o intercambien información con el sistema de registros administrativos económicos, cuáles operan con tecnologías informáticas que son aceptadas como válidas en el país en cuanto a formas de acceso y actualización de los datos del registro administrativo y que tengan, además, la capacidad de vincularse con la tecnología informática con la que opera el sistema de registros administrativos económicos.
- vii) Uso de registros administrativos para análisis, entrega y difusión de la información: evaluar si los registros administrativos como parte de sus funciones contiene el análisis, la entrega y la difusión de datos, así como la frecuencia y calidad con que se realizan estos procesos.

En un inicio se pueden aceptar las coberturas, rangos y tecnologías que sean capaces de brindar las diferentes fuentes consultadas que se hayan reconocido como válidas por el cumplimiento de los requisitos de calidad de sus datos, según se verá a continuación. Posteriormente, mediante acuerdos de desarrollo o modificando las leyes que limiten el uso, se puede ir profundizando más en los límites de aceptabilidad y accesibilidad. Las oficinas nacionales de estadística pueden trabajar con dueños de registros administrativos realizando preparación

metodológica y de convencimiento, de modo que se vayan ajustando las características del registro administrativo a las necesidades de su aprovechamiento estadístico.

En el anexo A3 se puede estudiar un ejemplo de cuestionario para la realización de un inventario de registros administrativos con fines estadísticos aplicado en Cuba.

En el caso del Perú, para el inventario de registros administrativos se ha desarrollado un sistema informático que incluye cuestionarios que determinan las características de las entidades proveedoras y del contenido de los registros administrativos. El Sistema de Inventario de Registros Administrativos (SIRA) ha permitido disponer de información de municipalidades y de entidades gubernamentales.

En paralelo a las fuentes administrativas seleccionadas pueden aceptarse otras fuentes que aportarían información estadística de referencia. Con ellas se puede evaluar que los datos sean lo más confiables posibles y se ajusten a los requerimientos. El marco legal para el uso de estas fuentes puede corresponderse con acuerdos institucionales que regulen los datos y la frecuencia de aportación. Ejemplo de estas son:

- Los datos aportados por los ministerios que regulan las actividades económicas de las unidades del registro estadístico de negocios, diferentes a los contenidos en los registros administrativos, incluso vinculados a estos, cumpliendo los mismos requisitos de calidad previamente convenidos.
- Los registros estadísticos que ya están definidos para los estudios estadísticos económicos y demográficos por otras instituciones estadísticas que operen en el país, y que pueden aportar los datos de identificación y las variables de análisis de las unidades de observación registradas en el registro estadístico de negocios, de acuerdo con las diferentes clasificaciones que tengan reguladas para sus informes. Se debe comprobar que tengan la capacidad de vincularse con los registros administrativos seleccionados como fuentes.
- Los registros de las oficinas nacionales de estadística que dan respuesta a solicitudes de información que vengan dadas por acuerdos o compromisos internacionales (por ejemplo, la información necesaria para cumplir con los Objetivos de Desarrollo Sostenible).
- Los registros de las ONE que poseen la información que brindan departamentos gubernamentales de las divisiones territoriales al amparo de un marco legal de cumplimiento obligatorio, con los datos estadísticos requeridos para conformar las bases de datos estadísticas económicas desde la unidad básica territorial hasta el nivel nacional.
- Las encuestas o reportes directos diseñados para obtener información económica en la página web de las ONE con acceso regulado.
- Los datos que las ONE de cada país, desde sus dependencias territoriales, obtengan por las vías tradicionales de comunicación con el nivel nacional.
- Los datos económicos desde los puestos gerenciales de las unidades del registro estadístico de negocios de las localidades a partir de encuestas convenidas previamente con carácter obligatorio.

4. Criterios para la selección de fuentes

Para la selección de los registros administrativos que constituyan fuentes del sistema de registros administrativos económicos, estos deben cumplir atributos de calidad. Estas exigencias se pueden resumir desde los principios de calidad de la información estadística de la siguiente manera:

- Contenido: el registro debe incluir los datos que permiten preparar y hacer las operaciones estadísticas necesarias para su uso.
- Exactitud (o validez): garantía de que los datos que aportan no sean erróneos.
- Antigüedad: los datos de aportación deben referirse a una actualización con fecha reciente.
- Coherencia: las definiciones de las unidades y de las variables no deben cambiar en el tiempo.

- Claridad y transparencia: las definiciones de las unidades, de las variables y otros datos deben ser de fácil comprensión por parte de las personas usuarias.
- Disponibilidad y accesibilidad: los datos del registro deben ser fáciles de obtener.

Antes de optar por la utilización de un registro administrativo es necesario garantizar algunas premisas básicas institucionales y operacionales de esos registros (registros administrativos y estudios estadísticos y prospectivos de la CEPAL):

- i) Acceso regular al registro administrativo establecido en algún tipo de legislación o acuerdo institucional que disponga cuestiones relacionadas con la frecuencia y los plazos en que los datos deben estar disponibles.

La limitación fundamental que se puede encontrar en la primera selección es el marco legal que garantice la accesibilidad de los registros administrativos con fines estadísticos.

Las leyes vigentes pueden frenar el uso de estos por otras instituciones o el intercambio de información, con fines diferentes a los que fueron concebidos. Debido a esto es necesario también realizar la búsqueda y estudio de las leyes que amparan los registros administrativos y en caso necesario, promover los cambios legislativos que permitan flexibilizar el uso de aquellos que tengan valor como fuente en la creación del registro estadístico. En todo caso se ha de garantizar que se cumplan los aspectos de confidencialidad establecidos.

- ii) Nivel mínimo de calidad de las variables importantes para los directorios, además de incluir alguna variable indicativa de tamaño (facturación, número de empleados u otras) con fines de muestreo, para lo cual es necesario conocer los procedimientos usados en el levantamiento de los datos administrativos y en los procesos de depuración.
- iii) Grado estimado de cobertura, error asociado a cada una de las principales variables y tipo de tratamiento adoptado para identificar creaciones y disoluciones de unidades.
- iv) Procedimientos utilizados para ajustar los datos administrativos a los requisitos del directorio estadístico en cuanto a definición estadística de las unidades básicas de información, sobre todo en empresas con más de un establecimiento que requieren ajustar el concepto de unidad institucional al de unidad estadística.
- v) Idealmente, los órganos estadístico y administrativo deben adoptar un sistema de clasificación de actividades uniforme.
- vi) Identificación de un vínculo entre las unidades del registro administrativo y las correspondientes unidades del directorio estadístico; por ejemplo, el código de identificación administrativo de tributación.
- vii) Trabajo conjunto entre los órganos administrativo y estadístico en discusiones técnicas para introducir mejoras en los procesos de levantamiento y tratamiento de los datos.

En general la calidad del registro administrativo y las bases de datos que van a ser usadas como fuentes del sistema de registros administrativos económicos deben poder cumplir requisitos con indicadores fijados, tales como:

- Porcentaje de unidades controladas con relación al total de la población.
- Porcentaje de unidades con campos de datos vacíos.
- Porcentaje de unidades registradas más de una vez en el registro o base de datos.
- Porcentaje de unidades con campos falsos en relación con lo establecido por los procedimientos y metadatos del sistema de registros administrativos económicos.
- Porcentaje de campos con última verificación de más de dos años.

En cuanto al nivel de calidad, las normas para cada indicador pueden ser las siguientes: menos del 5% se considera buena, menos del 10% y más del 5% se considera media, y más del 10% se considera mala.

Ya seleccionadas las fuentes que cumplen los requisitos básicos, puede existir más de una que, cumpliendo estos requisitos, sea capaz de brindar la misma información. Para seleccionar entre ellas se opta por la que más y mejor cumpla los requisitos que se tuvieron en cuenta al evaluar la selección.

En todos los puntos, durante el estudio del registro administrativo, debe quedar especificado cómo este cumple los requisitos relacionados con los nombres y especificaciones regulados a nivel de país, o a nivel internacional según sea el caso, así como sus ventajas y limitaciones, con vistas a poder hacer evaluaciones futuras más exhaustivas y definitorias.

B. Diseño de las descripciones de las variables

1. Los resultados estadísticos que se espera generar con el proyecto

Los resultados estadísticos que se obtienen del sistema de registros administrativos económicos deben corresponder con las necesidades de información que resultaron en el análisis del capítulo anterior (sobre la especificación de necesidades). Luego las variables estadísticas que recoge el catálogo deben contener como mínimo los datos necesarios para la obtención de estas.

Se hace un estudio de las fuentes administrativas y estadísticas que sean capaces de aportar las variables necesarias con la calidad requerida (sección IV.A.4, sobre criterios para la selección de fuentes). Estas variables quedan registradas en los metadatos del sistema de registros administrativos económicos y serán las que se deberán captar para estudios estadísticos hasta que haya una revisión programada de toda la metodología empleada, sobre la base de un nuevo análisis de las necesidades de información que requiera modificar el diseño del sistema de registros administrativos económicos.

Con determinada frecuencia especificada en la normativa se hace una revisión y evaluación de las nuevas necesidades a partir de las cuales las variables e indicadores resultantes pueden ser cambiados.

Primero se deben definir las variables que permitan realizar el cálculo de estos indicadores económicos para análisis cualitativos y cuantitativos, y seleccionar las fuentes y el método estadístico que se va a emplear para la obtención de las variables. Una vez hecho esto, se tienen las premisas para diseñar el funcionamiento del sistema de registros administrativos económicos en cuanto a cuáles serán las bases de datos y cómo estarán vinculadas desde la obtención de datos de las fuentes administrativas y estadísticas seleccionadas, hasta brindar toda la información pertinente en la obtención del resultado estadístico.

2. Identificación y selección de variables, siguiendo criterios estadísticos y conceptuales

Los datos han de procesarse para generar las variables de interés del proyecto o programa en cuestión. Las variables se eligen y diseñan partiendo de dos premisas:

- i) Las que sean necesarias para cumplir las necesidades establecidas en el capítulo anterior.
- ii) Las que tengan los datos suficientes para los análisis que se puedan obtener de las fuentes seleccionadas.

A la hora de conformar el sistema de registros administrativos económicos es posible que sea necesario adiestrar a un registrador o equipo de registradores que sea capaz de homogenizar los datos que se reporten desde las fuentes que contienen los datos requeridos para el registro estadístico. Debe tenerse en cuenta que las fuentes pueden exceder o estar limitadas respecto a la información que se necesite obtener. Deben definirse previamente todas las características de las variables que se quieren obtener a partir de los metadatos del sistema de registros administrativos económicos.

Las oficinas nacionales de estadística deben tener regulado como normativa que los catálogos de metadatos del sistema de registros administrativos económicos contengan y describan las variables que el registro contiene y el uso estadístico de esas variables.

Dicho catálogo debe tener un diseño que se regule en un documento respaldado y regulado jurídicamente con carácter obligatorio. Los metadatos deben tener normativas que permitan determinar que cumplen con el rigor de calidad que se espera de la información que brinda el registro. Dichas normativas deben estar referenciadas en este documento.

El catálogo de metadatos del sistema de registros administrativos económicos debe contener la siguiente información:

- Identificación del registro administrativo
- Entidad responsable
- Datos de contacto
- Objetivos
- Marco normativo
- Unidad de observación
- Cobertura geográfica
- Conceptos básicos
- Uso de nomenclaturas y clasificaciones
- Variables de la base de datos
- Metodología para el acopio de los datos
- Instructivos, manuales o guías para el desarrollo del registro
- Control y calidad del registro
- Almacenamiento
- Uso del registro
- Accesibilidad

Una guía de los requisitos para la revisión de las variables sería definir en los metadatos del sistema de registros administrativos económicos las siguientes características:

- Técnica a emplear para la transferencia de los datos: los datos pueden ser introducidos manualmente desde una encuesta impresa, o actualizados desde una encuesta automática diseñada para captar la información desde la web o por transferencia desde una base de datos conformada previamente.
- Unidad de medida de captación: se deben seleccionar unidades de medida homogéneas para los análisis que se requieran. Con base en lo anterior hay dos variantes:
 - i) las unidades de medidas capturadas de los datos de las fuentes es la misma que se utiliza para los análisis estadísticos, incluido el caso de captura de los datos desde un *software* de transferencia, y
 - ii) la otra variante sería que se pudiera tener programada su captura junto con la elección de las unidades del dato del reporte, ya que el *software* de captura realiza las conversiones a las unidades de medida adecuadas para que el resultado de los análisis sea homogéneo. Se recomienda que lo más práctico sería establecer una sola unidad de medida de captura para los datos, teniendo en cuenta la posibilidad de que el registrador o el reportador no sean conocedores de las unidades de medidas apropiadas, lo que podría traer confusiones y errores.
- Número de decimales permitidos: se debe definir el número de decimales permitidos en cada variable.
- Números negativos: se debe definir si la variable admite datos negativos, o valor cero ("0"), y regular los pasos que se siguen en caso de que no sean admitidos.
- Cobertura de la variable esperada por cada fuente: se deben definir las limitaciones de uso de los datos en cuanto a la coincidencia o diferencia que se pueda encontrar cuando es necesario acudir a más de una fuente para la obtención de una variable. También se debe especificar por cada fuente cuáles son o no son los datos deseados.
- Valores de referencia: los datos que se suministren como valores de texto o rangos numéricos deben ser bien conocidos por el reportador y el registrador, deben estar homologados en un archivo asequible y se debe especificar cómo actuar en caso de que no coincidan los de la fuente de información con los del registro estadístico que se está conformando.
- Validación: patrones de referencia con los que se puede evaluar si el dato obtenido cumple con los requisitos esperados.

Los datos de entrada de las fuentes deben coincidir con la observación real en un tiempo dado de una muestra de unidades seleccionadas como representativas, o de todas las unidades con información disponible. Se debe reunir suficientes datos de las unidades objeto de análisis, que permitan realizar un análisis estadístico válido, que pueda ofrecer una tendencia o patrón de acuerdo con la muestra analizada.

Deben obtenerse por tanto los datos de variables que permitan conseguir indicadores que sean capaces de ofrecer tendencias en función del tiempo y el grado de afectación que tienen los diferentes componentes de las variables que lo conforman en la tendencia temporal.

A su vez, la forma de obtención de los resultados atemporales debe poder identificarse como una característica constante del análisis.

C. Diseño de la recolección

Las oficinas nacionales de estadística trabajan de conjunto con las instituciones en establecer el diseño de recolección de la información que poseen en sus registros administrativos, cuyo contenido tiene información y se pueden vincular al registro estadístico de negocios.

Como parte del diseño de la recolección, se determinan los métodos e instrumentos más apropiados de acuerdo con las características definidas para el registro. También se incluye el diseño de cualquier acuerdo formal relacionado con el suministro de datos, tales como convenios interinstitucionales o memorandos de entendimiento, así como la confirmación de las bases legales para recolectar los datos.

Los pasos importantes en el diseño son:

- i) Revisión de los marcos legales de las diferentes instituciones que poseen registros administrativos con valor estadístico.
- ii) Definición de acuerdos de confidencialidad sobre el buen uso de la información de los registros administrativos, en particular cuando incluyen información sensible.
- iii) Definición de esquemas de comunicación de las oficinas nacionales de estadística con las instituciones gubernamentales y privadas, de manera que se garantice el intercambio de información.
- iv) Diseño de protocolos de almacenamiento y custodia de los datos en un repositorio centralizado para uso interno por los distintos departamentos de las ONE.
- v) Uso de plataforma de *software* mediante la cual se transmita la información de manera eficiente y segura.
- vi) Mecanismos de comunicación dentro de las ONE para compartir información proporcionada por instituciones públicas o privadas, garantizando la confidencialidad de la información.

Se recomienda revisar cuáles son las regulaciones que amparan el registro administrativo, con el objeto de confirmar las bases legales para la recolección de los datos y si incluye sus metadatos, frecuencia de actualización y si es auditable, entre otras cosas.

A partir de ello, se deben establecer formatos de acuerdo que ayuden a estandarizar los términos y condiciones para el intercambio de la información entre las instituciones y las ONE. En los convenios interinstitucionales o memorando de entendimiento deben quedar recogidos aspectos legales y metodológicos, como el marco legal que ampara el intercambio, las responsabilidades por cada institución en cuanto a las fechas de entrega, las periodicidades, la tecnología y procedimientos para la transferencia de datos y los procesos de conciliación y control de calidad y las cláusulas de confidencialidad, tratamiento de la información y difusión.

En el contenido de este acuerdo legal (sea convenio o memorando) deben aparecer todas las disposiciones acordadas en cuanto a la entrega de la información, las personas responsables y las fechas y periodicidad de la entrega y recepción de la información. Además, se deben detallar las variables y datos requeridos, los mecanismos de conciliación de la información, acuerdos de asesoramiento, responsabilidad de la institución respecto de la oportunidad y veracidad de la información, compromisos de auditorías de la información por parte de la ONE, tratamiento de confidencialidad, condiciones para la difusión de los datos y cláusulas de acciones en los casos en que se viole alguno de los términos establecidos en el convenio.

El formato en que se acuerde la entrega depende de las características del registro administrativo. Se puede utilizar el mismo formato o solo una parte de la información que contiene. El registro administrativo puede ser entregado como bases de datos o mediante cuestionarios, en formato digital o impreso. También deben indicarse las vías de transmisión o envío, o por vínculos autorizados en la web institucional, en dependencia del grado de digitalización y automatización que posean las instituciones. Previo al acuerdo del intercambio de información deben definirse claramente las características y plazos de este.

En caso de que tanto la institución que entrega la información de su registro administrativo como la ONE, posean una plataforma de *software* que se pueda alinear en red y que admita la consulta de bases de datos previamente consensuadas, este puede ser un mecanismo que permita la captura y comprobación de la información en ambas direcciones. Si es posible en el momento de la actualización del registro administrativo, conviene permitir que se consulten o validen los datos previamente enlazados en los registros de ambas partes, de forma automática, para evitar duplicaciones o enlaces falsos entre los registros. Esto puede ser un objetivo de trabajo entre ambas partes a largo plazo.

La ONE debe tener una infraestructura que permita la consulta de los registros estadísticos oficiales en una base de datos común a todas las áreas que trabajen con la información de los registros. Las actualizaciones realizadas a los registros estadísticos de negocios desde las fuentes que aportan los registros institucionales con cierta periodicidad pasan por un período de revisión y validación antes de considerarse información útil para su procesamiento estadístico. Estas deben ser publicadas después de su aprobación por parte de la ONE.

D. Diseño del procesamiento y análisis

Los datos que ayuden a la conformación de registros base para unidades u objetos de observación de interés estratégico, entre otros.

En el diseño del sistema de registros administrativos económicos se puede obtener información a tres niveles:

- i) Registros administrativos seleccionados como fuentes de información divididas en:
 - a) Registros administrativos de negocios
 - b) Registros administrativos económicos
 - c) Otros registros administrativos de utilidad en la clasificación y vinculación de información entre los anteriores
- ii) Registro estadístico de negocios caracterizado por:
 - a) Clave identificadora única
 - b) Nombre de la unidad
 - c) Bases de datos de las variables identificadoras (base de datos identificadora (BDI)1) (BDI2) (BDI3) ... (BDIn)
- iii) Bases de datos construidas mediante la captación directa de la información:
 - a) Censos económicos (si se realizan en el país)
 - b) Marcos de muestreo
 - c) Encuestas en campo
 - d) Encuestas telefónicas
 - e) Encuestas por Internet
 - f) Marcos estadísticos internacionales
 - g) Otras fuentes de información

El registro estadístico de negocios se nutre principalmente del registro administrativo de negocios y se complementa o valida mediante otras fuentes, que pueden ser censos y encuestas diseñados para configurar un registro estadístico de negocios que garantice la cobertura y rangos preestablecidos para abarcar todas las unidades de observación estadísticas del país.

Para conformar el registro administrativo con valor estadístico, se enlazan mediante su clave identificadora las unidades del registro de negocios obtenido con los números identificadores de las bases de datos de clasificadores estadísticos que forman parte de los marcos estadísticos internacionales, con las que operan las ONE.

Ya creadas estas bases informativas, con la actualización periódica de los registros y bases de datos se puede obtener en una consulta la caracterización actual y vigente de las unidades de negocio que tiene el país. Además, mediante la realización de varias consultas con determinada periodicidad (por ejemplo, anual) se puede hacer el estudio de cómo se modifica el ambiente económico empresarial durante mediciones periódicas sucesivas.

Este sistema se debe vincular e interactuar intercambiando informaciones y datos con los registros estadísticos que contienen las variables e indicadores de los análisis estadísticos económicos que ya están definidos en la institución, como los sistemas de cuentas, pagos tributarios, plantillas de empleados, reporte de producción, comercio, aduanas, entre otras cosas.

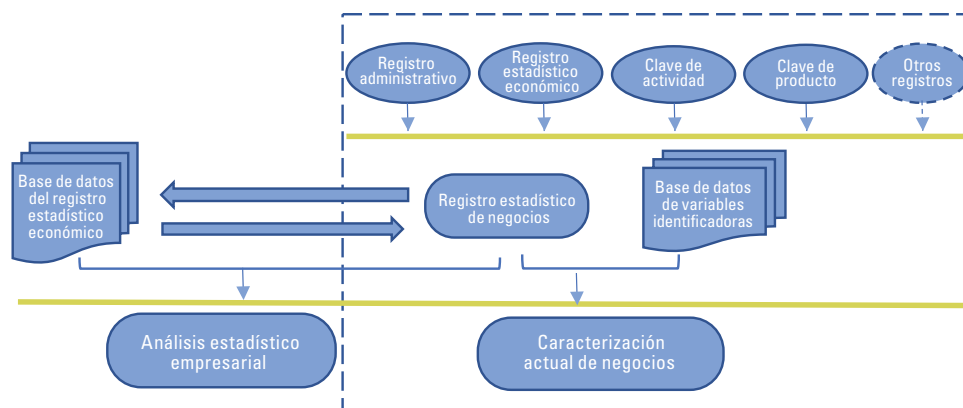
Es posible vincularlos partiendo de cuatro premisas a considerar:

- Se puede vincular la entrada de los datos de estos registros estadísticos con el número único que identifica a la unidad registrada en el registro estadístico de negocios, sin hacer modificaciones a las bases de datos de los registros actuales.
- Se vincula la entrada de los datos de estos registros estadísticos con otro número de identificación que puede homologarse con el número único que identifica la unidad registrada en el registro estadístico de negocios.
- Los registros estadísticos económicos no están diseñados para vincularse con el número único que identifica la unidad registrada en el registro estadístico de negocios, pero se puede establecer un campo que los enlace con pequeñas modificaciones de las bases de datos.
- No es posible enlazar los datos estadísticos económicos sin una modificación previa considerable de los sistemas informáticos actuales.

Cada una de estas premisas aporta un sistema de trabajo diferente a ser evaluado y corregido, para cumplir el objetivo de establecer un vínculo entre el registro estadístico de negocios y los registros estadísticos económicos empleados actualmente. La estructura debe quedar como el esquema del diagrama IV.3.

Diagrama IV.3

Esquema del vínculo del registro estadístico de negocios con los registros estadísticos económicos



Fuente: Oficina Nacional de Estadística e Información (ONEI) de Cuba.

Los datos de entrada proporcionados para conformar el registro estadístico de negocios pueden sufrir cambios significativos en un período de un año, debido al dinamismo de la economía propia de los negocios, así como la forma de afectación del ambiente económico que con relativa frecuencia hace que los negocios cambien de objeto, tecnología, estructura y hasta de ubicación territorial. Por ese motivo la actualización del registro estadístico de negocios partiendo de las fuentes elegidas debe tener un carácter continuo.

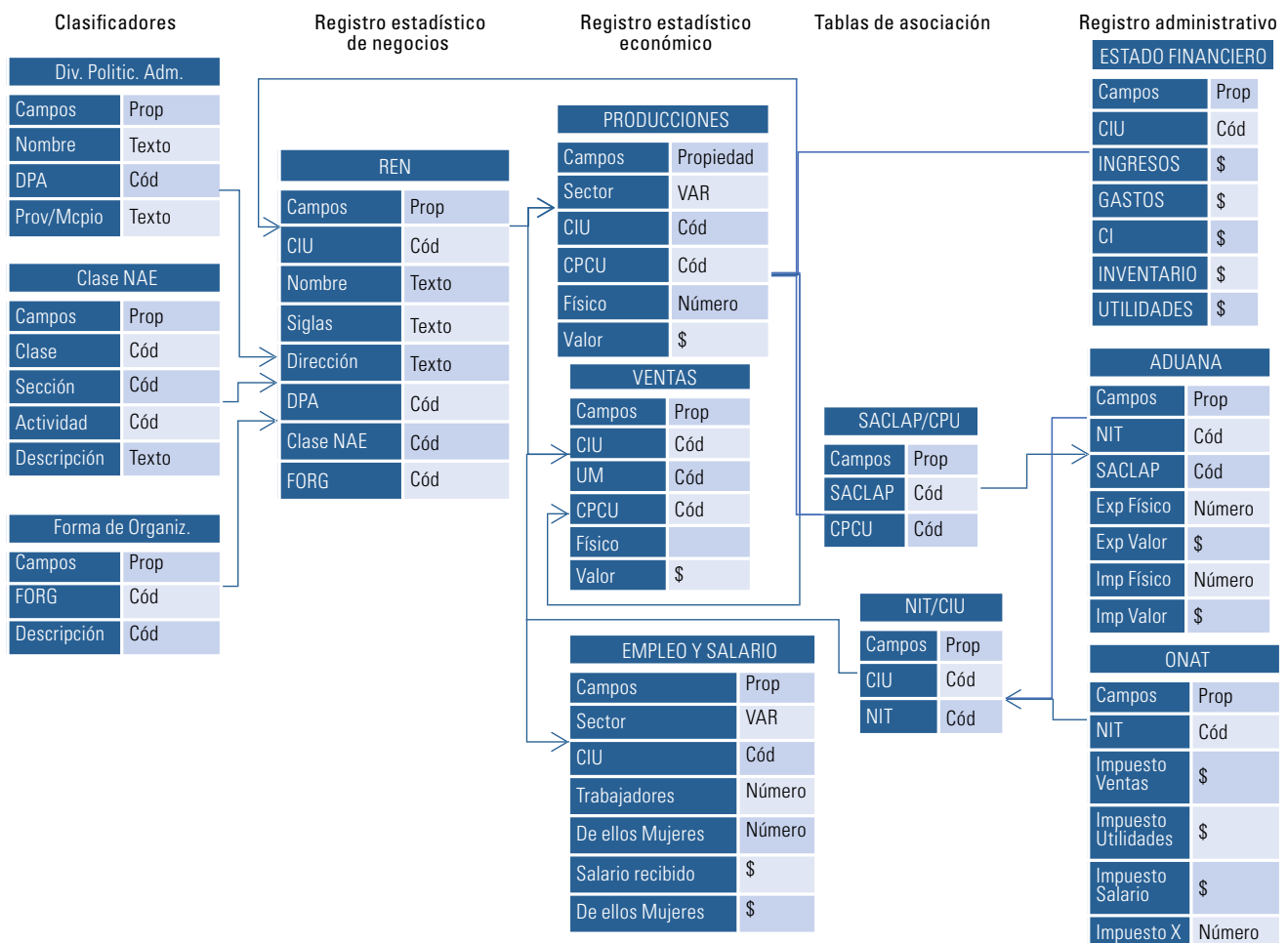
Lo óptimo sería que existiera un sistema de aviso ante una modificación de cualquiera de los datos obtenidos para la unidad básica. Este sistema funcionaría mediante un aviso emitido por propietarios del registro administrativo o de otras fuentes empleadas, con destino a las ONE, con el contenido y la descripción de las modificaciones que se han realizado. Para que se cumpla este aviso con la efectividad necesaria y requerida para los análisis que se emitan desde el registro estadístico de negocios, el sistema de aviso debe estar normado y regulado por leyes de carácter obligatorio o establecido en contratos que definan la obligatoriedad de los contrayentes.

Otra forma de mantener los datos actualizados sería asumir el trabajo de actualizar el registro estadístico de negocios con frecuencia periódica, lo que redundaría en una recarga de trabajo adicional para los registradores. Se acuerda una entrega de los datos de las fuentes actualizados con cierta periodicidad y, haciendo un trabajo minucioso para determinar cuáles son las modificaciones introducidas en las fuentes desde la última actualización, se definen y actualizan en el registro estadístico de negocios.

En cuanto al mantenimiento de los vínculos del registro estadístico de negocios y sus BDI con las bases de datos del registro estadístico económico, el trabajo futuro con vistas a su mantenimiento estaría regulado por una alerta que tenga en cuenta el vínculo ya establecido entre ambos registros, cada vez que haya cualquier actualización: nuevas inclusiones, modificaciones o eliminaciones. La alerta señala si se mantiene o adiciona el enlace con el registro estadístico de negocios, para que funcionen por el mismo sistema de enlace sin errores. Esta facilidad en el método de mantenimiento de las bases de datos viene dada por que las ONE son propietarias tanto del registro estadístico de negocios como de los registros estadísticos económicos con los que se puede enlazar (véase el diagrama IV.4).

Diagrama IV.4

Ejemplo de modelo de base de datos del sistema de registros administrativos económicos



Fuente: Oficina Nacional de Estadística e Información (ONEI) de Cuba.

E. Diseño de los sistemas de producción y flujos de trabajo

Este estudio debe ser realizado por las oficinas nacionales de estadística. Ellas serán las responsables de los registros estadísticos administrativos en su confección, mantenimiento y actualización.

Las figuras dentro de la ONE que son las responsables para ejecutar esta función pueden ser nombradas de acuerdo con su propio sistema organizativo o, como alternativa, por un equipo de trabajo *ad hoc*, ya sea temporal o permanente, que reúna especialistas necesarios para cumplir estas funciones.

1. Pasos para la selección y utilización de las fuentes

Los objetivos de este grupo de trabajo en una primera etapa de diseño del sistema también caracterizarían el contenido inicial de las funciones a desarrollar, las cuales serían:

- Recaudar toda la información legal que ampara el proceso de formación y mantenimiento de los registros. Evaluar si la legislación vigente es suficiente, o si es necesario reformar o aprobar nuevas leyes y resoluciones que garanticen cubrir las necesidades que se plantean.
- Definir los criterios de selección y aceptación en cada etapa del proceso de elaboración de los registros.
- Realizar un inventario de las fuentes dentro del país que pueden retribuir la información necesaria para la elaboración del registro estadístico de negocios. Posteriormente, evaluar cuáles cumplen con los criterios de selección prefijados.
- Dirigir y asesorar el procesamiento informático para la inserción de los procesos de creación y actualización de las bases de datos necesarias dentro de los sistemas existentes o la creación de un nuevo sistema que se enlace con el existente.

Concluida la primera etapa, o también puede ser en paralelo a esta, según vayan apareciendo definiciones del modo de actuar, se comienza un proceso de elaboración del marco legal y metodológico sobre el cual se basarán los intercambios de información, la conformación y el mantenimiento de las bases de datos, así como la seguridad en el manejo y publicación de la información recopilada. Este proceso se caracteriza por:

- Aprobación de las reformas y leyes que garanticen el marco legal de todo el proceso y, como parte de este proceso, el marco legal y metodológico con el que se realizarán los intercambios de información con las fuentes, los gestores y las personas usuarias de la información.
- Elaboración de los metadatos de registros y fichas de variables e indicadores que se van a utilizar.
- Gestión de la primera entrega de datos, homologación de nombres e identificadores de datos y variables, midiendo que se cumpla la cobertura y el rango especificados en la evaluación de las fuentes, desechando la información que no se va a utilizar y realizando los enlaces con las bases de datos según se vayan conformando (véanse más detalles en la sección IV.E.2, sobre pasos en el diseño de enlace entre los datos que aportan las fuentes y las bases de datos del registro estadístico).
- Medición y evaluación final del registro conformado. Hacer ejercicios de análisis posibles, partiendo de las bases de datos creadas.
- Reporte final del diseño con las fortalezas y debilidades con vistas a realizar mejoras.
- Elaboración de un manual operativo del sistema para guía y consultas de funcionalidad en modos de trabajo de mantenimiento y actualización de las bases de datos.
- Adiestramiento de registradores y operadores que quedaran a cargo de la operatividad del sistema con las bases de diseño finales.
- Adiestramiento de especialistas estadísticos y personas usuarias de todo tipo en las posibilidades de análisis de la información que brindan las bases de datos publicadas que puedan servir de consultas o como base a nuevas posibilidades de informes estadísticos con fines gubernamentales, asesorías financieras, estudios económicos, estudios educativos, entre otras cosas.

En todas las etapas de los procesos se debe garantizar que solo sean públicos los análisis globales o sectoriales, sin que haya forma de acceder a las bases de datos que puedan brindar información clasificada de las unidades básicas del registro.

2. Pasos en el diseño de enlace entre los datos que aportan las fuentes y las bases de datos del registro estadístico

El primer paso sería revisar la calidad y cobertura de los datos administrativos, chequeando varios puntos, entre los que se encuentran:

- El archivo tiene el tamaño esperado con relación al número de registros y variables.
- Las variables principales tienen un formato válido en lo que se refiere al rango, la nominación de “texto” o “numérica”, entre otras.
- La cobertura de las variables principales tiene valores para todas las unidades básicas.
- Las principales variables son confiables, en el sentido de que las numéricas corresponden a un rango de valores admisibles y los códigos usados son válidos, entre otras cosas.

Si este primer análisis de calidad revela que las variables de las fuentes tienen la calidad suficiente y pueden ayudar a aumentar la calidad del registro estadístico de negocios, es necesario elaborar un conjunto de procedimientos que dicten cómo realizar un pretratamiento de las variables del registro administrativo para cumplir los requisitos que debe tener antes de su transferencia al registro estadístico de negocios. Para ello se deben seguir las definiciones reguladas en el metadato del registro en cuanto a las definiciones estadísticas y la aplicación, entre otras cosas.

Este conjunto de procedimientos debe incluir los pasos necesarios, tales como la creación de algoritmos y bases de datos seguras para convertir variables y clasificaciones, como pueden ser la conversión de códigos en la clasificación actividades, los detalles del código de localización geográfica, entre otras cosas.

Capítulo V

Construcción

Introducción

Esta fase tiene por objeto la construcción y prueba de la infraestructura informática, los componentes, aplicaciones y servicios de *software*, para crear un ambiente operacional completo que permita ejecutar la producción de información, así como las pruebas que lo acrediten.

Dentro de la lógica de estandarización de los procesos, los insumos proporcionados por el proceso de diseño facilitan la construcción de diferentes componentes de un ambiente operativo que permita generar la operación estadística. Estos deben ser flexibles, puesto que requieren una mejora continua que atienda las cambiantes necesidades de información y de incorporación de fuentes de información definidas en el proceso de diseño.

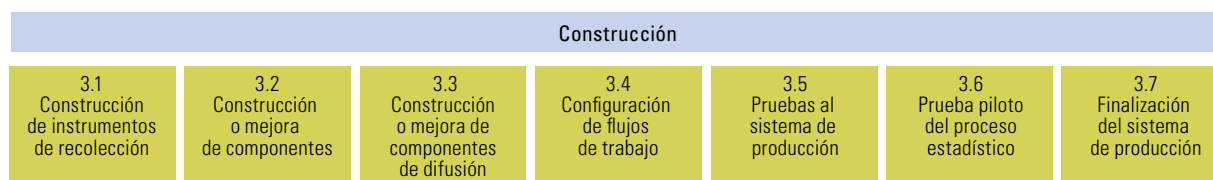
A. Definiciones a nivel de proceso

El MGPIE indica que dentro de este proceso se construye y prueba la solución de producción al punto en que está listo para ser utilizado en un ambiente “en vivo”. De esta manera, los insumos proporcionados por el proceso de diseño deben facilitar la construcción de un ambiente operativo que permita llevar adelante la operación estadística. Dentro de la lógica de estandarización de procesos, los componentes que se construyan en esta etapa deben ser en lo posible reutilizables, siendo mejorados según los requerimientos definidos en el proceso de diseño.

Este proceso se divide en siete subprocesos, según se indica en el diagrama V.1.

Diagrama V.1

Subprocesos incluidos en la fase de construcción



Fuente: Comisión Económica para Europa (CEPE), “Modelo genérico del proceso estadístico GSBPM (versión 1.0 en español, septiembre 2016)”, Ginebra, 2016.

Para entender una lógica de organización en el ámbito de los registros administrativos que se definen en este documento, se presenta el detalle de cada uno de los subprocesos del MGPIE.

1. Construcción de instrumentos de recolección

Dentro del MGPIE se define que este subproceso describe las actividades para construir los instrumentos que se usarán durante la fase de “recolección”.

Si bien se puede entender que el instrumento de recolección hace referencia a un cuestionario, el MGPIE también hace referencia a distintos modos de recolección, pues indica que “los instrumentos de recolección pueden ser también rutinas de extracción usados para recolectar datos de conjuntos de datos estadísticos o administrativos existentes”.

En ese sentido, la ONE, en acuerdo con las entidades proveedoras de datos, debe establecer un entorno para el intercambio de registros administrativos. Este debe considerar aspectos de seguridad y almacenamiento para dar respuesta a todos los requerimientos necesarios.

Además, las rutinas de extracción pueden ser automatizadas, semiautomatizadas o manuales, dependiendo del avance informático de las diferentes entidades proveedoras de datos, pero es necesario que la ONE se adapte a todos los escenarios posibles, con tal de capturar los datos necesarios para la construcción del registro estadístico de negocios, así como todos aquellos que permitan profundizar el uso de los registros administrativos económicos.

De esta manera, el intercambio de datos podría producirse en un entorno informático que sea proporcionado por entidades proveedoras de datos, como en uno proporcionado por la ONE.

Cuando en los registros administrativos a recolectar se considere un volumen menor, estos podrían ser traspasados a la ONE con otros medios digitales, como correo electrónico o discos magnéticos. De igual manera, se debe velar siempre por una mayor automatización y por un entorno seguro para el resguardo de la información sensible que se recibe.

De esta forma, en este subproceso se desarrollan las herramientas tecnológicas que soporten las actividades de transmisión y recepción de las variables de los registros administrativos, así como las variables derivadas, archivos o bases de datos estadísticos o no estadísticos que serán objeto de estudio. Esto significa desarrollar los medios tecnológicos de transmisión e interoperabilidad de datos provenientes de registros administrativos, archivos o bases de datos estadísticos o no estadísticos, desarrollar las técnicas y las herramientas para verificar que la información enviada o recibida cumpla con las condiciones de calidad establecidas para el intercambio de datos, y elaborar el plan de compras, así como las licitaciones, los contratos, convenios o acuerdos necesarios para el desarrollo de la operación estadística.

Se debe considerar la importancia para la definición de desarrollo de medios tecnológicos de transmisión e interoperabilidad. Ello se convierte en un desafío para el intercambio de datos, en el sentido de que cada vez se cuenta con un mayor volumen de información para el desarrollo de las operaciones estadísticas. De esta manera, las ONE deben estar preparadas para enfrentar estos desafíos en sus diferentes realidades.

Las ONE deben estar preparadas para la captura de diferentes tipos de datos. Estos pueden ser clasificados según su origen, su rango, su estructura, entre otros. Según la estructura se pueden definir tres tipos de datos: estructurados, semiestructurados y no estructurados. Los datos estructurados son típicos de la mayoría de las bases de datos relacionales. Estas bases de datos se caracterizan por tener un esquema determinado que define cómo son las tablas en las que se almacenan los datos, qué tipo de campos tienen y cómo se relacionan entre ellas. Son gestionados mediante el lenguaje SQL (Structured Query Language), diseñado precisamente para administrar y recuperar información de los sistemas de gestión de bases de datos relacionales.

Por otra parte, los datos semiestructurados, no relacionales o NoSQL no encajan en un formato de tablas, filas y columnas, sino que se organizan mediante etiquetas (*tags*) que permiten agruparlos y crear jerarquías. Así, las bases de datos NoSQL permiten almacenar información que no se adapta bien al formato de registro o tabla, por ejemplo, el texto de longitud variable, y también facilitan el intercambio de datos entre distintas bases de datos. Un ejemplo de datos semiestructurados son los archivos delimitados, que contienen elementos que permiten dividir los datos en jerarquías independientes.

En el caso de los datos no estructurados, su volumen y crecimiento son radicalmente superiores al de los datos estructurados. Su origen es diverso, por ejemplo, la información extraída de la web mediante técnicas de raspado web (*web scraping*). Debido a su estructura, no es posible utilizar la arquitectura relacional, por lo que es necesario trabajar con herramientas de macrodatos (*big data*), además de hacerse críticos en estas arquitecturas los aspectos relacionados con la escalabilidad y el paralelismo.

Es importante señalar que, históricamente, las ONE solo se han centrado en la extracción y análisis de información de datos estructurados. Sin embargo, con el actual crecimiento de la información semiestructurada y no estructurada, las ONE deben buscar soluciones que puedan ayudarles a analizar los tres tipos de datos. Esto representa una ventaja analítica y, de hecho, la capacidad de extraer valor de ellos es uno de los principales motores del rápido crecimiento de los macrodatos.

2. Construcción o mejora de componentes

Este subproceso describe las actividades para construir nuevos componentes y mejorar los existentes, así como servicios necesarios para las fases de procesamiento y análisis, de acuerdo con lo que se haya determinado en la fase de diseño.

El MGPIE hace referencia a elementos centrales para el trabajo que involucran los registros administrativos en términos de las características del sistema informático, donde se pueden “incluir funciones y características del tablero, servicios de información, funciones de transformación, marcos de referencia del flujo de trabajo, así como servicios de gestión de las entidades proveedoras y los metadatos” (CEPE, 2016).

Dado que parte del foco de la fase de construcción radica en el desarrollo, actualización, mejoras y prueba de los sistemas informáticos que deben sostener la operación estadística, en esta sección se presenta una serie de módulos que deben considerarse como parte relevante de la gestión de los registros administrativos económicos y del registro estadístico de negocios.

Es importante comprender que, en términos de sistemas, no solo deben considerarse aspectos relacionados con la captura, almacenamiento y procesamiento de los registros administrativos, sino que también hay aspectos relacionados con la gestión de entidades proveedoras, la evaluación de los datos, la integración de sistemas de datos, las herramientas para la documentación de los registros, entre otros.

Por lo tanto, se debe comprender que las soluciones informáticas deben proveer soluciones a la producción estadística basada en registros administrativos con una mirada transversal del proceso de quienes lo ejecutan.

A continuación, se presentan diferentes aspectos que se deben considerar para el desarrollo de los sistemas informáticos que sostienen el sistema de registros administrativos económicos.

a) Tecnología

La tecnología debe ser lo suficientemente flexible para evolucionar con nuevos requisitos. Por ello se recomienda que el sistema a desarrollar se ajuste a la arquitectura de tecnología de la información de la oficina nacional de estadística y que sea lo más compatible posible con otros sistemas, como los de adquisición de datos administrativos y de recolección de encuestas comerciales.

También, si la ONE cuenta con estándares definidos, estos deben seguirse. Si no, se sugiere pensar en una metodología ágil para mostrar resultados de manera concreta y regular.

Además, el documento describe que los requerimientos para el sistema se pueden diferenciar en dos: los funcionales y los no funcionales. Los primeros hacen referencia a lo que el sistema debe hacer o ejecutar, mientras que los segundos guardan relación con el cumplimiento de los estándares técnicos, el rendimiento, la escalabilidad, la seguridad, la disponibilidad, la accesibilidad, la calidad y la usabilidad de la organización. Para abordar las soluciones tecnológicas que darán soporte al sistema de registros administrativos económicos, se debe plantear una serie de interrogantes que deben ser resueltas con el objetivo de considerar las soluciones *ad hoc* según el estado de desarrollo de las diferentes ONE.

La evaluación de estos aspectos y sus soluciones influirán directamente en cómo deba diseñarse el sistema. Los autores Wallgren y Wallgren (2012) mencionan factores como la tecnología de almacenamiento usada por el registro, la frecuencia de actualización, la cantidad de datos, la seguridad y las tecnologías de comunicación.

Respecto de la tecnología de almacenamiento se indica si es posible transferir los registros administrativos directamente a la ONE. Respecto de la frecuencia de actualización se indica que, mientras más frecuente sea la transferencia, se debe pensar en un mayor nivel de automatización.

Relacionado con el punto anterior, si la cantidad (volumen) de datos es grande, se requiere construir una solución acorde para estos propósitos. Además, con respecto a la seguridad se deben establecer aspectos de resguardo de confidencialidad de los datos, así como de vulnerabilidad del sistema.

Respecto de las tecnologías de comunicación, se deben establecer posibles soluciones para los factores mencionados previamente, con el fin de elegir las mejores alternativas disponibles para la ONE.

b) Bases de datos

Parte del desarrollo de un sistema que almacene registros administrativos, y todos los datos que soportan al sistema de registros administrativos económicos, se debe considerar, dentro de las definiciones centrales, la especificación de cómo se relacionan las bases de datos entre sí.

En ese sentido, existen diferentes soluciones como Microsoft SQL Server, Oracle, entre otras. Si bien en años recientes ha aumentado el aprovechamiento de los datos no estructurados, en una etapa inicial se debe establecer un modelo de datos relacional, que permita establecer vínculos entre distintos registros mediante identificadores definidos para estos propósitos. Junto con esto, para el desarrollo del sistema de registros administrativos económicos y del registro estadístico de negocios, se debe contar con identificadores únicos que permitan establecer vínculos entre las unidades estadísticas definidas y los datos disponibles dentro del sistema. Este es un elemento central para vincular el registro estadístico de negocios y los registros administrativos económicos.

Para la construcción del sistema, la definición de los requisitos en términos de las bases de datos puede tener implicancias en los costos, ya sean por las licencias de las aplicaciones de *software* o por el tipo de almacenamiento requerido, los requisitos de cantidad y capacidad de los servidores de base de datos, entre otros. Respecto a la descripción del ítem de bases de datos, es importante considerar cómo se debe construir el modelo de bases de datos, qué tablas y variables almacenar y cómo van a interactuar las tablas. En línea con esto, y como se mencionó previamente, el uso de identificadores únicos para establecer vínculos entre los registros administrativos permite establecer relaciones entre ellos y potenciar el aprovechamiento estadístico. También podría facilitar la integración entre sistemas, toda vez que facilite las actividades de vinculación de los datos a analizar.

Un modelo de base de datos relacional tiene aspectos positivos en tanto se pueden establecer requerimientos de integridad del modelo que permitan asegurar la coherencia entre las tablas que se relacionan. También se pueden establecer restricciones y criterios para cargar los datos que formarán parte del sistema. Por lo tanto, un modelo de datos relacional permite el establecimiento y cumplimiento de criterios de calidad de los datos a utilizar.

En Wallgren y Wallgren (2012) también se menciona la relevancia del diseño de las bases de datos para el funcionamiento eficaz del sistema. En ese sentido indican que “el diseño de una base de datos para un sistema informático que trabajará con estadísticas basadas en registros no difiere en lo fundamental del diseño de bases de datos para otros sistemas informáticos destinados a la producción de estadísticas” (pág. 213). Por lo tanto, nuevamente se presenta la idea de que las ONE deben considerar sus definiciones de arquitectura y sistema para desarrollar los requerimientos informáticos en torno a los registros administrativos.

Respecto a las bases de datos, mencionan tres tipos principales: base de datos de entrada, de procesamiento y de salida.

La base de datos de entrada debe “tener capacidad para recibir todo tipo de actualizaciones, incluidos datos de entrada totalmente incorrectos” (Wallgren y Wallgren, 2012, pág. 213). A las bases de datos de entrada se les pueden aplicar reglas de consistencia básicas.

La base de datos de procesamiento es aquella que se normaliza y se le aplican reglas de consistencia más estrictas. Por su parte, a la base de datos de salida no se le deben hacer cambios y está optimizada para las salidas más utilizadas.

Además, Wallgren y Wallgren (2012) plantean la necesidad de definir dentro del modelo los aspectos relacionados con el desempeño de las bases de datos. Por ejemplo, si existen grandes volúmenes de información, introducir ciertos cambios puede empeorar el desempeño, por lo que pueden establecerse otros tipos de bases de datos que permitan optimizar el diseño.

Por otra parte, Wallgren y Wallgren (2012) plantean que “no es posible establecer un sistema de registros en una sola base de datos física, probablemente ni siquiera en un solo servidor” (pág. 216). Si bien el servidor puede asegurar un funcionamiento más veloz, respecto al trabajo que puede desarrollarse conectando diferentes servidores, se deben tomar decisiones respecto de cuál es la mejor forma de distribuir los datos, y cuán necesario es tener datos distribuidos en diferentes servidores.

Finalmente, en cuanto a las bases de datos se debe tener una indexación correcta para el buen desempeño del sistema. Esto significa que se deben establecer índices apropiados que permitan hacer más eficiente el cruce de datos. De esta manera se pueden indexar las claves primarias y externas, con el fin de ubicar de manera más rápida los elementos que serán utilizados para los cruces de datos. Este es otro de los aspectos a definir dentro de un modelo de base de datos.

c) Documentación

Es importante indicar que se necesita establecer la documentación relacionada con el sistema de registros administrativos económicos para garantizar su funcionamiento a largo plazo. Esto permite desarrollar el historial y guiar y entender los cambios que se han producido, tanto a nivel de sistema como de los datos que se administran. Ello implica que se debe documentar y hacer gestión sobre los documentos para establecer con claridad cómo funcionan los módulos y procesos desarrollados. También se deben documentar los cambios dentro de los registros administrativos que existen dentro del sistema.

Además, se recomienda realizar un seguimiento de las actualizaciones de fuentes administrativas. Por ejemplo, un registro administrativo económico que sea automatizado debe incluir la fecha y la fuente de actualización. Esta información puede resultar útil para solucionar problemas con un mantenimiento específico o procesos de actualización.

En línea con el enfoque de la CEPE, Wallgren y Wallgren (2012) también plantean la lógica de establecer una documentación completa del sistema, y los datos que contiene, como uno de los aspectos centrales a ser considerados. Estos autores plantean un conjunto de metadatos formalizados y otros documentos.

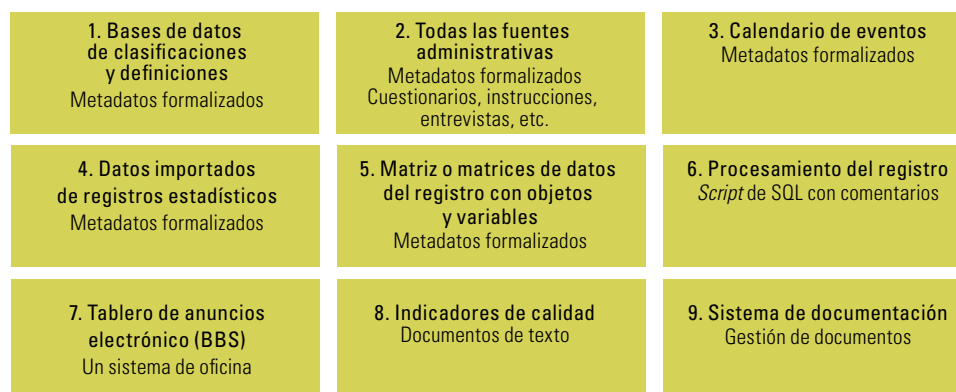
En el primer grupo se encuentran las bases de datos de clasificaciones y definiciones, la documentación de bases de datos procedentes de fuentes administrativas, el calendario de eventos (de cambios en el registro), los datos importados de los registros estadísticos y la documentación de las matrices de datos en el registro estadístico (la población de registro, el tipo de objetos y las variables).

Entre los otros documentos se encuentran la información de los sistemas administrativos, los cuestionarios administrativos con instrucciones y las minutas y notas de reuniones con entidades proveedoras de registros y los *scripts* de SQL con comentarios que describen el procesamiento del registro.

Si bien los autores hacen mención de SQL, las oficinas nacionales de estadística pueden trabajar con otros lenguajes informáticos para el almacenamiento, procesamiento y análisis de los datos dentro de un sistema de registros administrativos económicos (véase el diagrama V.2).

Diagrama V.2

Diferentes tipos de metadatos y herramientas en la documentación de un registro



Fuente: A. Wallgren y B. Wallgren, *Estadísticas basadas en registros: aprovechamiento estadístico de datos administrativos*, Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

Por lo tanto, el sistema debe permitir la gestión de estos documentos que facilitan el entendimiento de las funcionalidades y de los registros contenidos para el aprovechamiento estadístico.

3. Construcción o mejora de componentes de difusión

Este subproceso describe las actividades para construir nuevos y mejores componentes y servicios existentes que sean necesarios para la difusión de los productos estadísticos como se diseñaron en el subproceso 2.1 (productos del diseño).

De esta manera, se deben considerar todos aquellos elementos de difusión que involucren las operaciones estadísticas relacionadas con los registros administrativos económicos.

En este subproceso se considera el desarrollo de todas las herramientas que soportarán los productos estadísticos definidos a partir del sistema de registros administrativos económicos y también del registro estadístico de negocios.

Entre los componentes de difusión a desarrollar se encuentran bases de datos, informes de análisis, documentos y metadatos, así como herramientas geoespaciales. Las herramientas de difusión deben ir en línea con aquellas que disponga cada institución para sus diferentes operaciones estadísticas, ya sean basadas en registros administrativos, censos o encuestas.

4. Configuración de flujos de trabajo

Este subproceso “configura el flujo de trabajo, sistemas y transformaciones utilizadas dentro del proceso estadístico, desde la recolección de los datos hasta la difusión de los mismos” (CEPE, 2016).

Para una adecuada gestión de los registros administrativos económicos se hace necesario que las ONE tengan definidos, dentro de su estructura organizacional, los roles y responsabilidades de todas las unidades que están involucradas con la gestión de los registros administrativos económicos y las entidades proveedoras de estos.

Es importante destacar que, dentro del flujo de trabajo, deben considerarse perfiles enfocados para un aprovechamiento estadístico relevante de todos los datos existentes en el sistema de registros administrativos económicos. A esos efectos, se debe considerar a profesionales con conocimientos especializados en estadísticas, métodos científicos y análisis de datos. Además, teniendo en cuenta la relevancia del trabajo con grandes volúmenes de datos, es importante que dentro del flujo de trabajo se considere un soporte constante de los equipos informáticos de la oficina estadística.

5. Pruebas del sistema de producción

Este subproceso “se ocupa de probar los servicios ensamblados y configurados y los flujos de trabajo relacionados a ellos” (CEPE, 2016).

Además, este subproceso “define el alcance de la prueba que se realizará, costea las pruebas y establece los productos esperados, define los instrumentos que se deberían probar (instructivos, manuales, formatos de control o aplicativos informáticos), especifica las condiciones bajo las cuales se desarrollarán las pruebas, así como los formatos utilizados, su tiempo y responsables de estas y se elabora el informe que documente los resultados, las conclusiones y las recomendaciones de las pruebas realizadas” (DANE, 2020, pág. 42). Teniendo en cuenta que el trabajo con registros administrativos tiene un importante componente informático, se deben considerar dentro de este subproceso todas las pruebas de componentes y servicios involucrados, como podría ser la interacción de sistemas de transferencia de archivos con entidades proveedoras de información.

6. Prueba piloto del proceso estadístico

Este subproceso describe “las actividades relacionadas con gestionar una prueba de campo o piloto del Modelo Genérico de Proceso Estadístico”. Típicamente, este incluye “una recolección de datos a menor escala para probar los instrumentos de recolección seguido del procesamiento y análisis de los datos recolectados para asegurarse de que el Modelo Genérico de Proceso Estadístico funcione de la manera esperada” (CEPE, 2016).

Si bien dentro de esta descripción existe una lógica de una operación estadística que utiliza cuestionarios para la captura, en el marco del trabajo a desarrollar con los registros estadísticos económicos se deben realizar todas las pruebas necesarias que permitan establecer un proceso continuo a partir de la recolección de datos, toda vez que se requiera desarrollar un trabajo en torno a los registros administrativos.

7. Finalización del sistema de producción

Este subproceso incluye “las actividades para colocar los procesos y servicios ensamblados y configurados, incluyendo aquellos que hayan sido modificados o sean nuevos, en producción lista para ser usada en áreas de negocio” (CEPE, 2016).

B. Módulos del sistema

Se debe considerar que el sistema o sistemas informáticos que soporten al sistema de registros administrativos económicos y al registro estadístico de negocios deben contener diferentes elementos que permitan llevar adelante una adecuada gestión, teniendo en cuenta los desafíos que se presentan para el manejo de grandes volúmenes de información.

1. Gestión con entidades proveedoras de datos

Esta se define como el conjunto de actividades en que se promueve y fomenta el relacionamiento efectivo de entidades proveedoras públicas y privadas de registros administrativos con las oficinas nacionales de estadística, la creación de acuerdos y la ejecución de intercambios de registros administrativos y su transformación en registros estadísticos. Esto se aplica también para los registros administrativos económicos.

Considerando que las ONE se enfrentan a los desafíos de gestionar volúmenes de datos desde diferentes entidades proveedoras, ello implica que se deben establecer mecanismos que permitan llevar de mejor manera

este trabajo. Así, este módulo debe permitir establecer información básica de entidades proveedoras de datos, por ejemplo: nombre de la entidad proveedora, detalles del acuerdo de provisión de datos, contrapartes establecidas dentro de la institución, forma de transferencia de datos, entre otros.

Idealmente, cada entidad proveedora de datos debe identificarse con un identificador único, al igual que cada registro administrativo que proporcione. Estos identificadores facilitarán la vinculación con la documentación que soporta este módulo.

Además, de este módulo se puede extraer el inventario de entidades proveedoras de datos, que consiste en una caracterización de las fuentes de datos disponibles en el sistema de registros administrativos económicos, y que facilita a los equipos productores de estadísticas conocer todos los antecedentes que tienen disponibles para el aprovechamiento estadístico de los registros administrativos económicos.

2. Acopio de registros administrativos

Como se mencionó en la sección referente al subproceso de construcción de instrumentos de recolección, se deben desarrollar los medios tecnológicos de transmisión e interoperabilidad de datos provenientes de registros administrativos, archivos o bases de datos estadísticos o no estadísticos, y desarrollar las técnicas y las herramientas para verificar que la información enviada o recibida cumpla con las condiciones de calidad establecidas para el intercambio de datos.

Es muy probable que los medios tecnológicos puedan proporcionarse con diferentes herramientas de soporte, por ejemplo, medio magnético, protocolo de transferencia de archivos (FTP), servicios web, archivo sin formato, hojas de cálculo de Excel, base de datos, entre otros. No obstante, es importante para la gestión del sistema de registros administrativos económicos poder gestionar dentro un sistema todos los antecedentes relacionados con la transferencia de archivos, entre los que se encuentran: registro administrativo capturado, tamaño de los archivos, medio de captura, identificar si se requiere algún permiso o si existen restricciones para acceder al registro administrativo, entre otros.

3. Transformación de los registros administrativos

Una vez capturado el registro administrativo económico, las oficinas nacionales de estadística deben desarrollar un conjunto de actividades para su transformación a registro estadístico. De esta manera, a nivel de sistema informático, se debe contar con procedimientos establecidos para el desarrollo de actividades de consistencia de datos, estandarización y normalización, y desarrollo de las llaves únicas.

Como se ha mencionado a lo largo del capítulo, es posible que algunas de las actividades a desarrollar no se encuentren en un solo sistema informático, sino que también se utilicen otros servicios disponibles en las ONE. Por ejemplo, para la estandarización de domicilios se recurre a un servicio establecido en la oficina estadística que cumpla esos propósitos. También, para la detección de anomalías, podrían desarrollarse rutinas en un *software* estadístico, como SAS o R.

Por lo tanto, es importante entender que para el desarrollo del sistema informático se pueden establecer interacciones entre módulos específicos del sistema con aplicaciones de *software* estadístico.

Para el funcionamiento del sistema de registros administrativos económicos, existe un elemento fundamental que consiste en el desarrollo de las claves únicas. Es dentro de este módulo donde a cada unidad estadística existente dentro del sistema de registros administrativos económicos se le asigna una variable de identificación única, que permite la vinculación unívoca de la unidad con las diferentes fuentes de datos donde se encuentren.

Puede darse el caso, tal como proponen Wallgren y Wallgren (2012), de que, para garantizar la privacidad de cada unidad estadística dentro del sistema, se generen llaves únicas ficticias que permitan resguardar información sensible de las unidades. De esta manera, la vinculación dentro del sistema de registros administrativos económicos, tanto dentro del registro estadístico de negocios como de los registros administrativos económicos, se produce mediante estas llaves únicas ficticias.

4. Documentación de los datos

Como se mencionó en el punto 2 de la sección A de este capítulo (construcción o mejora de componentes) es fundamental establecer la documentación del sistema de registros administrativos económicos para un aprovechamiento estadístico.

Por lo tanto, dentro del sistema debe establecerse un repositorio que permita almacenar todos los documentos relacionados con el registro estadístico de negocios y los registros administrativos económicos, y con la ejecución de los procesos asociados a su transformación en registros estadísticos. Este módulo debe dar soporte a la carga de documentos técnicos, convenios, diagnósticos de calidad, entre otros.

5. Administración

El sistema informático debe permitir a los responsables de la gestión del sistema de registros administrativos económicos tener elementos de mantenimiento y control para poder gestionar las personas usuarias, los parámetros (como clasificadores y codificadores), entre otros.

Capítulo VI

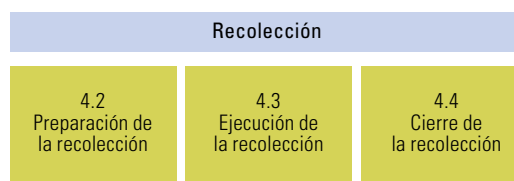
Recolección

Introducción

Conforme a la estructura del modelo del proceso estadístico MGPIE, la fase de recolección abarca los siguientes subprocesos, que enmarcan las actividades que deben llevarse a cabo en ese proceso para generar estadísticas oficiales (véase el diagrama VI.1).

Diagrama VI.1

Subprocesos incluidos en la fase de recolección



Fuente: Comisión Económica para Europa (CEPE), “Modelo genérico del proceso estadístico GSBPM (versión 1.0 en español, septiembre 2016)”, Ginebra, 2016.

Nota: Se omite el subproceso 4.1, “Conformación del marco muestral y selección de la muestra”, que no se aplica a registros administrativos, debido a que se orienta a aspectos que abarcan otros métodos tradicionales para elaborar estadísticas, como el caso de las encuestas por muestreo.

Esta fase tiene por objeto captar los datos que habrán de aprovecharse estadísticamente. Se ejecutan las actividades que fueron definidas, diseñadas, construidas y probadas durante las fases previas.

En la recolección se pueden obtener diversas fuentes de datos, como las tradicionales provenientes de las operaciones estadísticas. También se acopian datos generados por las unidades productoras de información a partir de los registros administrativos, por ejemplo, información geográfica, imágenes, entre otros, derivados a bases de datos o microdatos.

Ante la realidad de cada país y de cada unidad poseedora de registros administrativos, se presentan varios escenarios para el requerimiento de datos y su aprovechamiento estadístico. Sin embargo, es importante diferenciar dos situaciones o niveles que se dan en este proceso, a saber, el levantamiento de los datos por parte de las unidades poseedoras de los registros administrativos y el acceso de las ONE a los datos o a la información para su uso estadístico.

Considerando las instancias tratadas en las fases anteriores, el proceso de captación de la información debe estar garantizado por una metodología establecida para este fin, que contemple los procesos, la disponibilidad de la tecnología adecuada, los materiales y los demás elementos pertinentes durante la preparación y ejecución de la recolección.

Es indispensable establecer pautas para el proceso de captación de los datos necesarios en la producción de información estadística económica. Se deben convenir los contactos con los organismos productores de registros administrativos y las ONE para el envío de los datos en tiempo y forma, así como la realización de una validación básica de la estructura y la integridad de los datos recibidos, de forma que se cumpla con las especificaciones técnicas requeridas y los protocolos de confidencialidad.

Muchos organismos del Estado disponen de registros administrativos generados con base en metodologías y necesidades de información ligadas al ámbito de operación de las ONE. A partir de esta premisa se describen los mecanismos recomendables para el intercambio o acopio de la información.

A. Aspectos normativos o legales

Los avances tecnológicos y su incorporación a los diferentes procesos administrativos propician la generación masiva de información proveniente de los registros administrativos. Aunque estos no hayan sido creados con fines estadísticos, se les puede aprovechar con este fin. De hecho, algunos países, ajustando y aprovechando sus respectivas legislaciones, les dan un uso regular y los tienen como fuente estratégica en la producción de sus estadísticas oficiales. Basados en estas experiencias y dependiendo de las necesidades detectadas en cuanto a la definición de equipos técnicos, los países de la región deben plantearse el fortalecimiento o, en su defecto, la introducción de ajustes paulatinos o puntuales en sus marcos normativos. Estos deben ser flexibles y garantizar el acceso de las oficinas nacionales de estadística a los registros administrativos de las entidades generadoras de registros de carácter público, privado o mixto que forman parte del SEN. Estas a su vez deben tener acceso a retroalimentaciones que sean útiles y que contribuyan al mejoramiento continuo en términos técnicos e incluso conceptuales, para así fortalecer la generación de información estadística útil para los tomadores de decisiones en materia de política pública del país.

En cuanto a la adecuación de los respectivos marcos, en el caso de que algún país lo considere necesario, la región cuenta con una herramienta actualizada y homogenizada con base en las necesidades y requerimientos de los países. Se trata de la Ley Genérica sobre Estadísticas Oficiales para América Latina, la cual proporciona un modelo referencial para los países de América Latina que necesiten elaborar o reformular la base jurídica que define el funcionamiento de los SEN y la producción de estadísticas oficiales.

En particular, el artículo 17 de la mencionada Ley Genérica trata sobre el acceso a los registros administrativos y hace énfasis en la obligatoriedad que tienen las entidades poseedoras de registros administrativos. En este contexto el mencionado artículo dicta cuanto sigue: “Los proveedores de registros administrativos están obligados a proporcionar a los productores de estadísticas oficiales, de forma gratuita, los datos/registros administrativos que obran en su poder, con el nivel de detalle necesario para la producción de estadísticas oficiales y los metadatos, cuando sea posible, que permitan evaluar la calidad de los datos. No se podrán invocar las disposiciones de confidencialidad o de secreto de otra legislación, a menos que dicha legislación excluya explícitamente la utilización de datos para fines estadísticos” (CEPAL, 2020, pág. 22).

En países en los que este proceso no esté contemplado en la legislación o lo esté de manera tácita o no reglamentada, entre los mecanismos legales que son eficientes en su aplicación para acceder a la información, se encuentran los convenios interinstitucionales o los memorandos de entendimiento, los cuales se deben establecer considerando el interés de las partes y buscando que todas sean beneficiadas.

1. Convenios interinstitucionales o memorandos de entendimiento

Para el diseño de los operativos de captación, es necesario considerar los acuerdos iniciales entre la unidad poseedora de registros administrativos y la oficina nacional de estadística, sobre todo los puntos relativos a las responsabilidades que se asumen en cuanto a la participación con recursos humanos, materiales y financieros. También se deben considerar los procedimientos para el registro y transferencia de datos a la oficina nacional de estadística, la participación en el cumplimiento de funciones relacionadas con el control de la calidad de los datos recabados y el intercambio de asistencia técnica para la generación de las estadísticas, entre otros. En lo posible, estos acuerdos deben establecerse a largo plazo, al margen de los cambios administrativos o de funcionarios.

Los convenios y la concertación de apoyos pueden darse entre áreas de una misma dependencia u organismo, o entre áreas de instituciones diferentes e independientes.

En cualquiera de estas situaciones, la concertación y firma de convenios se realiza a nivel directivo, donde se concilian intereses y se establecen las condiciones propicias para la posterior toma de decisiones durante el desarrollo del proyecto estadístico y, en particular, para el operativo de captación, considerando lo siguiente:

- Acordar una estructura funcional y un programa de trabajo.
- Determinar las responsabilidades en cuanto a la participación de recursos humanos, materiales y financieros.
- Determinar la tecnología y los procedimientos para la transferencia de los datos, de las unidades poseedoras de registros administrativos a la ONE.
- Determinar la participación en el cumplimiento de funciones relacionadas con el control de la calidad de los datos recabados para la generación de las estadísticas, entre otros.
- Establecer compromisos y esquemas de control respecto de la captación y entrega de lotes de información en tiempo y forma por parte de las unidades poseedoras de registros administrativos.

Los reglamentos, políticas y normas que fundamentan jurídicamente los registros que las dependencias y organismos de la administración pública realizan como parte de su función son los que dan continuidad al proceso de registro y, por lo tanto, posibilitan la implementación de este tipo de proyectos estadísticos.

Este marco legal lo conforman básicamente:

- la Constitución nacional de los países;
- la ley orgánica de la creación de la institución;
- el plan nacional de desarrollo, y
- la ley de transparencia.

Asimismo, se complementa con otro tipo de legislación y reglamentos internos de las dependencias, por los que se rige la operación de las unidades involucradas.

En general las estrategias para la recopilación de datos en forma más eficiente dependen de:

- El reconocimiento de un organismo rector que norme, coordine y vigile el cumplimiento de la ley o los reglamentos nacionales en materia de estadística, para garantizar que la información de interés público se ponga a disposición de los responsables de la estadística.
- Iniciativas promovidas por la unidad de estadística cuando así proceda, tendientes a la modificación de las leyes o los reglamentos jurídicos o administrativos nacionales pertinentes, a fin de establecer bases sólidas para el mejoramiento de la calidad y la producción oportuna de las estadísticas.
- El respeto de estas leyes implica, por un lado, la obligatoriedad que compromete a las unidades de registro a entregar su información en el tiempo y la forma acordados; por parte de la unidad de estadística, la obligación consiste en salvaguardar el principio de confidencialidad de la información.
- El apoyo de ciertos organismos o secretarías de Estado vinculados con las unidades de registro, quienes condicionan la autorización de trámites para la operación de sus actividades, a cambio de la entrega de su información.

Los cambios en la legislación afectan al proceso estadístico, específicamente a la captación, por lo que es necesario determinar con claridad los ámbitos de competencia, estatal y municipal, e identificar las instancias legales que puedan aprovecharse para sensibilizar a los interesados sobre la importancia de la estadística en cuestión y la entrega de la información con los requisitos de calidad y confidencialidad acordados.

B. Aspectos metodológicos

En el formato de intercambio, la oficina nacional de estadística debe recibir sistemáticamente los registros administrativos y, en este contexto, se debe cumplir lo siguiente:

- Con base en un calendario acordado, se procede al intercambio de la información. La unidad poseedora del registro administrativo transfiere los datos necesarios a la ONE y esta, a su vez, realiza una retroalimentación.
- Asimismo, siempre y cuando la ONE esté en condiciones de hacerlo, debe brindar apoyo, con desarrollo de *software* o asesoría técnica, para facilitar el trabajo a las unidades poseedoras de registros administrativos y mejorar la calidad de la información. Por ello es tarea primordial promover y fomentar el impulso y el desarrollo de la infraestructura informática de estas últimas.
- En caso de que en el diagnóstico previo se haya identificado que los registros no cumplen con los requerimientos necesarios, se definirán de común acuerdo las delimitaciones técnicas, conceptuales y de infraestructura tecnológica. Esto implica realizar los ajustes necesarios en el instrumento de recolección utilizado por la unidad poseedora del registro administrativo, con sus respectivos respaldos conceptuales y operativos.

1. Estructura e información necesaria

Es preciso establecer en el acuerdo interinstitucional cuál será la estructura y la información mínima o básica de la que se necesita disponer para fines estadísticos. Esto debe estar fundamentado en los hallazgos del diagnóstico de la información requerida con respecto a la información disponible en las diversas fuentes.

La estructura se definirá con base en la evaluación de aprovechamiento del registro administrativo. En ella se consideran los aspectos relacionados con el metadato y esencialmente con el soporte tecnológico, puesto que se debe conseguir la correspondencia de los formatos y las variables. Asimismo, en esta estructura se priorizan las necesidades fundamentales de información del registro estadístico de negocios. En caso de que no se cumpla a cabalidad, se establecen fuentes alternativas o complementarias que podrán completar o suplir lo requerido.

Una vez definida la estructura adecuada a todas las partes, la unidad poseedora del registro administrativo que suministrará la información a la oficina nacional de estadística, lo hará con el correspondiente metadato y en el formato acordado.

Por último, la estructura definida y la información que se compartirá deben constar explícitamente en el documento de acuerdo interinstitucional, en el que además se expliciten el formato y la modalidad del intercambio, teniendo en cuenta las legislaciones a las que están sujetas ambas partes.

2. Periodicidad

Es necesario coordinar un calendario de trabajo en el que conste la periodicidad de:

- La recolección por parte de la entidad poseedora del registro administrativo de los datos que se necesitan para los fines estadísticos, junto con el período de referencia de los datos.
- La transferencia de los datos a la oficina nacional de estadística (recolección) y, en consecuencia, de la retroalimentación por parte de esta en los términos acordados.

3. Oportunidad

Para garantizar la oportunidad es fundamental el cumplimiento de los acuerdos de periodicidad y las fechas establecidas para la disposición de los datos. Ello debe estar claramente establecido en el convenio interinstitucional o memorando de entendimiento. Además, es importante establecer claramente los criterios técnicos que se precisan, de ambas partes, en cuanto al cumplimiento de los controles previos a la disponibilidad que debe incluir la crítica y codificación de las variables. Estos criterios deberán acordarse previamente en el convenio marco entre las partes.

4. Medio o canal de intercambio

En cuanto a los medios o canales de intercambio de los registros administrativos, es necesario considerar estos aspectos:

- Formato, equipamientos y periféricos: se refiere a una homologación en términos de formato y de modalidad de la entrega, pues se debe asegurar que las instituciones involucradas tengan disponibilidad de acceso al formato, a la fuente o forma de envío y, sobre todo, a la capacidad tecnológica de albergar el registro administrativo salvaguardando la seguridad y la confidencialidad. En este contexto será necesario definir, como mínimo, los requerimientos siguientes:
 - i) Sistema operativo con su versión especificada.
 - ii) Permisos para la instalación o desinstalación de *software*.
 - iii) Cantidad de espacio en disco duro requerido (para la instalación).
 - iv) Cantidad de espacio necesario en disco duro para el envío y recepción de información.
 - v) Tipo de conexión a Internet y ancho de banda necesario.
- Legales: cada país tiene su legislación, e internamente las instituciones también tienen normas específicas de confidencialidad. En este contexto se deben considerar las estrategias para salvar las discrepancias (si las hubiere) e incluir explícitamente en el acuerdo las aclaraciones y conciliaciones que respaldan el intercambio de la información.

5. Confidencialidad

En este punto se deben prever las condiciones y los niveles de acceso del funcionariado de la unidad poseedora del registro administrativo y de la oficina nacional de estadística. Se deberá firmar una carta de compromiso de confidencialidad en la que consten puntualmente sus responsabilidades y se comprometan a cumplir el mandato de la ley en términos del respeto de la confidencialidad de la información a la que se accede. Esta carta deberá ser firmada en tres copias que se destinarán: una al funcionario, una a la entidad o unidad poseedora del registro administrativo y una a la oficina nacional de estadística. Además, la carta será renovada cada un año calendario a partir de la firma.

Una vez designado el personal autorizado para la transferencia de datos por parte de la unidad poseedora del registro administrativo, este deberá ser registrado como tal ante la ONE y dicho procedimiento formará parte del convenio o memorando de entendimiento. Cumplido este procedimiento, le será emitido un certificado o cédula firmado por las máximas autoridades de las instituciones involucradas en el convenio.

Este personal autorizado recibirá una notificación por correo electrónico, mediante la que se le informará su designación como responsable para realizar el envío y recepción de la información, y que contiene:

- un vínculo para la descarga del instalador del sistema informático;
- la clave de activación (requerida durante la instalación del sistema), y
- la guía del usuario (este documento) que le orientará en la descarga, instalación, registro y operación del sistema.

C. Mecanismo tecnológico

A efecto de que las unidades poseedoras de registros administrativos provean a la oficina nacional de estadística la información de su ámbito de competencia, esta dispondrá de los mecanismos tecnológicos necesarios y suficientes para que se pueda realizar este proceso.

En este sentido, el desarrollo de la herramienta informática para el intercambio queda a cargo de la ONE y con ese fin se deben considerar los metadatos de la unidad poseedora del registro administrativo, previendo de esta manera la compatibilidad y la practicidad en el momento de la implementación. En lo posible, esta herramienta deberá formar parte de una plataforma que abarque las actividades de recolección primaria de la información por parte de la unidad poseedora del registro administrativo y la transmisión de los datos compilados por dicha unidad a la ONE, así como los protocolos para su almacenamiento, acceso y seguridad.

El sistema debe garantizar la seguridad de la información que se intercambia y, en consecuencia, esta debe ser sometida a procesos de cifrado y compactado mediante contraseña segura y encriptado del canal. De este modo se previenen pérdidas o accesos indebidos, con lo que se garantiza el carácter confidencial de la información.

La implantación deberá ir acompañada del plan de entrenamiento al personal vinculado con esta tarea, teniendo en cuenta los roles y los perfiles del talento humano, las temáticas correspondientes, la intensidad horaria y el método (virtual, presencial, semipresencial o de aprendizaje mixto).

D. Capacitación

Las oficinas nacionales de estadística deberán llevar a cabo capacitaciones continuas para el adiestramiento de los responsables en las entidades poseedoras de registros administrativos, con el propósito de mejorar el procesamiento de la información y los procedimientos técnicos.

Se deberá establecer una estrategia de capacitación y programar un calendario para su ejecución de manera coordinada en los diferentes niveles de la estructura organizacional, integrada por técnicos asignados según las funciones a desempeñar

Capítulo VII

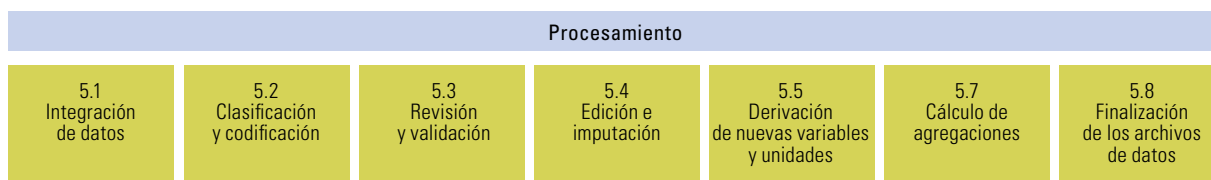
Procesamiento

Introducción

La de procesamiento es la fase posterior a la de recopilación de los datos de los registros administrativos económicos, con la finalidad de prepararlos para el análisis y de que puedan convertirse en información utilizable para la producción estadística. Está compuesta de varios subprocesos mediante los cuales se verifican, limpian y transforman los datos de entrada, se estandarizan las clasificaciones y códigos, se revisan, validan, editan e imputan los datos, y finalmente se integran. También se podrán crear nuevas variables, unidades, así como ponderadores y agregados, y luego organizar estos datos procesados en una base de datos para su análisis y difusión como resultados estadísticos (véase el diagrama VII.1).

Diagrama VII.1

Subprocesos incluidos en la fase de procesamiento



Fuente: Comisión Económica para Europa (CEPE), "Modelo genérico del proceso estadístico GSBPM (versión 1.0 en español, septiembre 2016)", Ginebra, 2016.

Nota: Se omite el subproceso 5.6, "Cálculo de ponderadores", que no se aplica a registros administrativos, debido a que se orienta a aspectos que abarcan otros métodos tradicionales para elaborar estadísticas, como el caso de las encuestas por muestreo.

Es recomendable que antes de la recopilación y procesamiento de los datos del registro administrativo económico, identificado con el inventario, se le realice una evaluación de calidad con la finalidad de conocer su potencial para el uso estadístico en el marco de las necesidades de la información objetivo.

Como ya se ha expuesto en los capítulos anteriores, los registros administrativos económicos provienen de diversas fuentes de la administración pública y de entidades privadas. Son generados con objetivos, variables, códigos y procesos para el control de sus actividades, que son distintas a los fines estadísticos que se persiguen, debiendo conocerse el estado de situación y el nivel de calidad de estos.

La calidad estadística de los registros administrativos debe evaluarse considerando las dimensiones o atributos que los organismos internacionales han reconocido que se deben cumplir en las actividades estadísticas. La Organización de Cooperación y Desarrollo Económicos (OCDE) prioriza ocho dimensiones o atributos, en su documento *Quality Framework and Guidelines for OECD Statistical Activities* (que contiene el marco y las directrices de calidad para actividades estadísticas), los cuales se muestran en el recuadro VII.1.

Recuadro VII.1

Dimensiones clave en la calidad de las estadísticas

De acuerdo con el documento *Quality Framework and Guidelines for OECD Statistical Activities* de la OCDE (2012), las dimensiones más relevantes para asegurar la calidad de las estadísticas son:

1. Pertinencia: las estadísticas cumplen con las necesidades de los usuarios.
2. Precisión: las estadísticas representan con precisión y fiabilidad la realidad.
3. Credibilidad: confianza depositada por los usuarios en los productos estadísticos.
4. Oportunidad y puntualidad: las estadísticas son publicadas de manera puntual y oportuna.
5. Accesibilidad: los datos se encuentran almacenados en locaciones seguras y conocidas por los usuarios que los demandan.
6. Interpretabilidad: los datos cuentan con toda la información complementaria para ser entendidos por los usuarios.
7. Coherencia y comparabilidad: las estadísticas son coherentes internamente, en el tiempo y en el espacio, y es posible combinar y hacer uso conjunto de datos relacionados provenientes de diferentes fuentes.
8. Costo-eficiencia: si bien no es una dimensión de calidad propiamente dicha, es un factor que se debe considerar en los análisis de calidad.

Fuente: A. Rojo, "Del dato administrativo de firmas al dato estadístico: desafíos a vencer en América Latina", *CAF Working Paper*, N° 2018/17, Caracas, Banco de Desarrollo de América Latina (CAF), 2018; Organización de Cooperación y Desarrollo Económicos (OCDE), *Quality Framework and Guidelines for OECD Statistical Activities* (STD/QFS(2011)1), París, 2012.

Se propone hacer en dos partes la evaluación de la calidad estadística de los registros administrativos³:

- La primera, con el propósito del análisis conceptual o del proceso de gestión del registro, para observar el cumplimiento de las dimensiones o atributos de la OCDE, referidos a la pertinencia, accesibilidad, credibilidad, oportunidad, coherencia y comparabilidad, y en la que se podrá considerar la aplicación de un cuestionario *ad hoc*. A este cuestionario se deberá anexar la documentación que se tenga del registro. También se puede sostener una entrevista con el responsable del registro.

Entre los principales aspectos de este análisis se revisarán las definiciones, las principales variables, la cobertura del registro, los métodos de recolección, los tipos de clasificación y codificación, el uso de tecnología, los criterios de validación y las personas usuarias.

- La segunda parte está orientada al análisis de la base de datos del registro administrativo. Se recomienda utilizar herramientas estadísticas por la magnitud de los datos. Tiene como fin revisar el cumplimiento de las dimensiones de precisión y coherencia.

El análisis se inicia con la revisión de la consistencia interna para observar errores, duplicados, cambios en la codificación, falta de información, entre otras cosas. Es imprescindible cuando los registros administrativos carecen de documentación o tienen muy poca. Luego se revisará la consistencia externa, es decir, se harán comparaciones con datos externos sobre estructura, cantidades y composiciones con otras fuentes estadísticas disponibles.

Los resultados de la evaluación de la calidad permiten el desarrollo de un plan de acción de mejora, que debe ser socializado con responsables del registro administrativo con el fin de corregir los errores o inconsistencias a corto plazo. Asimismo, se coordinarán las mejoras que fortalezcan el registro administrativo y que se realizarán en un mayor plazo.

Actualmente, existen cuestionarios o instrumentos diseñados para la evaluación de la calidad estadística de los registros administrativos, entre ellos:

- i) La Herramienta de Evaluación de la Calidad de los Registros Administrativos (HECRA) es un instrumento que fue elaborado con el apoyo del Banco Mundial, utilizado por el INEGI de México, entidad que le ha efectuado mejoras.

³ Para ampliar la lectura de la evaluación de la calidad estadística de los registros administrativos, sírvase consultar Rojo (2018).

La HECRA ofrece a las personas usuarias la posibilidad de tener elementos de autodiagnóstico para establecer sus áreas de oportunidad con el fin de mejorar la calidad de aquellos registros administrativos que sean susceptibles de aprovechar para generar información estadística y geográfica⁴.

- ii) El Cuestionario para Evaluar la Calidad de los Registros Administrativos (CECRA) es el instrumento elaborado en el marco del proyecto “Estadísticas de población e inmuebles a partir del uso de registros administrativos oficiales en la Comunidad Andina” con el apoyo del Banco Interamericano de Desarrollo (BID), en el marco de su programa de bienes públicos regionales (BPR).

El CECRA tiene como objetivo ser el instrumento mediante el cual se capte información de las fuentes administrativas con el fin de elaborar un diagnóstico de estos registros administrativos a nivel nacional y subregional. Las etapas de desarrollo son: diseño, producción, análisis, entrega y difusión.

A continuación, se describirán los subprocesos que conforman la fase del procesamiento, tomando como base el modelo genérico del proceso estadístico MGPIE.

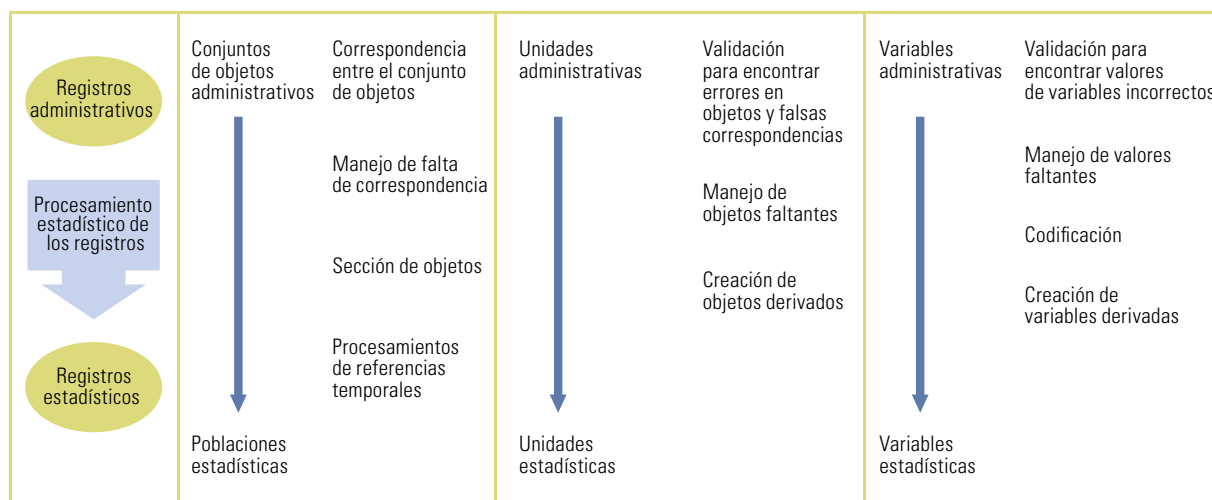
A. Integración de datos

Antes de realizar la integración de registros administrativos, se deberá hacer la transformación o la conversión para pasar de registros administrativos a registros estadísticos.

Esta labor implica preparar el registro administrativo para integrarlo con otros registros. En el diagrama VII.2 se describe el proceso de transformación de los registros administrativos, que pueden ser uno o varios en un registro estadístico.

Diagrama VII.2

Registros estadísticos a partir de registros administrativos



Fuente: A. Wallgren y B. Wallgren, *Estadísticas basadas en registros: aprovechamiento estadístico de datos administrativos*, Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

En la transformación de los registros administrativos, se estandarizan y normalizan variables, clasificadores y códigos, que luego se revisan y validan. También se desarrollan llaves únicas, se revisan clasificadores y códigos, se depuran y transforman los datos de entrada, a fin de ser utilizados con fines estadísticos. Es decir, se debe contar con reglas de consistencia preestablecidas (véase la sección B.3 del capítulo V, relativa a la transformación de registros administrativos).

⁴ Véase Instituto Nacional de Estadística y Geografía (INEGI), “Herramienta para la evaluación de la calidad de los registros administrativos (HECRA)” [en línea] <https://www.inegi.org.mx/mejorespracticas/hecra/>.

Una vez que la base de datos esté limpia, podrá integrarse con otros registros administrativos o estadísticos.

La integración es el subproceso que se inicia una vez culminada la etapa de depuración de los datos. Requiere que los registros administrativos económicos cuenten con una clave única de identificación, de forma que se facilite su vinculación.

Casi todos los países que tienen sistemas administrativos tributarios identifican a las empresas con un código tributario único, mediante el cual la entidad tributaria realiza el control de la gestión de esas entidades. Este código único asignado a las empresas es sumamente importante, pues permitirá la integración de los registros administrativos para disponer de una mayor y mejor información económica que cubra las demandas de los usuarios, una vez que haya culminado el procesamiento y el análisis de la información.

En el Perú, este código es el Registro Único del Contribuyente (RUC), generado por la Superintendencia Nacional de Aduanas y de Administración Tributaria (SUNAT), que ya se encuentra estandarizado. Actualmente esta fuente de información es utilizada en el Directorio Central de Empresas y Establecimientos del INEI.

Para la integración de uno o varios registros administrativos, se recomienda que se haya realizado una buena evaluación de su calidad estadística. Se debe tener en cuenta que, al unir los registros, se tendrá una combinación de los resultados de los subprocesos de la fase de recolección. Por esa razón los datos de entrada, al originarse de diversos registros, podrían tener métodos de recolección diferentes.

La integración de los registros administrativos podrá generar un sistema de registros administrativos económicos o incorporarlos al registro estadístico. Según el caso, hará también que se creen nuevas variables y unidades.

Es recomendable anonimizar los registros para garantizar la confidencialidad de la información antes de la actualización en la base de datos de la producción estadística. Esto puede hacerse después de la integración.

Segui (2016) manifiesta que, para integrar registros (unirlos o emparejarlos), se aplican métodos determinísticos y probabilísticos.

La unión de los registros consta de tres etapas: preunión, unión y postunión. La preunión corresponde a la normalización de variables de texto y estandarización de otras variables (realizada en la fase 5, de construcción). La unión de registros se realiza aplicando algoritmos estadísticos estandarizados y ampliamente utilizados en diversas disciplinas. Finalmente, se procede a la postunión o proceso de evaluación de los resultados (siguiente subproceso).

Como pasos después de la integración de registros, el autor señala los siguientes:

- i) evaluar los resultados de la integración;
- ii) derivar nuevas variables y unidades;
- iii) aplicar la protección de datos;
- iv) crear o actualizar el registro estadístico en la base de producción de datos, y
- v) crear o actualizar el registro base.

B. Clasificación y codificación

Este subproceso permite la revisión del sistema de clasificación y codificación que utilizan los registros administrativos que integran el registro estadístico económico a fin de conocer si estos corresponden a los estándares para la producción estadística o son códigos o tablas *ad hoc* elaboradas para su propio uso. En esta revisión se evaluarán si se cumplen las dimensiones o atributos de la OCDE referidos a la precisión, credibilidad, oportunidad y puntualidad.

La revisión de la clasificación y codificación utilizada en los registros administrativos que alimentan el registro estadístico es de suma importancia porque permitirá identificar si el registro se relaciona con las clasificaciones estándar de las oficinas nacionales de estadística, de modo que puedan compararse con la información que estas instituciones producen. En caso de que esto no ocurra, se deben elaborar tablas que se puedan relacionar con los registros administrativos para su conversión. Estas tablas se conocen en la producción estadística como tablas de equivalencias, que facilitan su homologación para la conversión con fines estadísticos.

Un sistema de clasificación y codificación único permite guardar o almacenar la información en forma ordenada, de manera que se pueda obtener información de una parte o de la totalidad del registro. Por ejemplo, en el caso del Directorio Central de Empresas y Establecimientos, se puede obtener información correspondiente a una región del país, o tal vez a una actividad económica específica de la Clasificación Industrial Internacional Uniforme de Todas las Actividades Económicas (CIIU) rev. 4. También se puede obtener información sobre género, es decir, la proporción de mujeres que son propietarias de negocios o que tienen cargos gerenciales.

Una clasificación estadística puede tener una estructura plana o lineal, por ejemplo, cuando se quiere conocer el tipo de organización de las empresas, si es persona natural o persona jurídica.

La clasificación estadística puede tener una estructura jerárquica, es decir, corresponder a varios niveles. Por ejemplo, cuando se utiliza la Clasificación Industrial Internacional Uniforme de Todas las Actividades Económicas (CIIU) (revisión 4), se tiene información a nivel de sección, división, grupo y clase. En este grupo también se consideran los clasificadores de la distribución territorial a nivel de país, que permiten conocer la información económica por áreas geográficas. En el caso del Perú, su división político-territorial está estructurada en departamentos, provincias y distritos.

Cuando las clasificaciones literales se transforman a numéricas, se hace mediante un sistema de codificación, a fin de facilitar la tabulación que permitirá la presentación de los datos estadísticos en forma de tablas o cuadros. En los registros administrativos es común encontrar variables clasificadas en categorías, que usualmente se codifican con valores numéricos. En algunos registros el sexo se codifica como 1 si es varón y 2 si es mujer, pero en otros registros administrativos se emplean las letras del alfabeto (M y F o H y M). De ahí la necesidad de que en este proceso se verifique la codificación utilizada para asegurarse de que sea la correcta. Además, si se trata de varios registros administrativos, se debe revisar que estén armonizados con los que utilizan las ONE. Las fallas en la codificación trascienden en los resultados.

Las modalidades de codificación más utilizadas son:

- Codificación manual, generalmente se utiliza en los registros administrativos que no son automatizados y son asignados por los responsables del registro.
- Codificación automática, son asignados mediante un programa de codificación preestablecido en los catálogos electrónicos.
- Codificación semiautomática, se asignan códigos con catálogos manuales y catálogos electrónicos.

C. Revisión y validación

Este subproceso consiste en la revisión de los datos de entrada para detectar potenciales problemas en el registro estadístico económico, que se ha conformado con registros administrativos.

Se identifican los errores, incongruencias y discrepancias que pudieran existir. Estas últimas pueden ser valores atípicos, respuestas sin datos, datos faltantes y errores en la codificación. Esta revisión se efectúa sobre los datos que ingresan de los registros administrativos.

Este subproceso también se conoce como “validación de datos entrantes” y se inicia cuando los datos ingresan al sistema. Generalmente, esta validación se realiza de manera iterativa, es decir, de forma continua y las veces que sea necesaria.

Es recomendable que, antes de la revisión y validación, se elabore un manual con consistencias básicas en que se describan las reglas que los datos deben cumplir en relación con los flujos de información, los rangos de los datos, las reglas lógicas de validación, las reglas para la cobertura, los controles para la verificación de la codificación, la generación de tablas para la actividad económica y para la distribución geográfica, el control de verificación de valores extremos y la identificación de valores faltantes, entre otras cosas. Para ello será de mucha utilidad el análisis previo realizado en el diagnóstico de la calidad estadística.

La revisión y validación de datos puede hacerse de forma manual o automática. Debe aplicarse antes y después de la integración de los registros administrativos y a cualquier fuente de datos que alimente el registro estadístico.

Cuando es automatizado, se producirán informes con los registros que no fueron validados, los que deben ser revisados por los responsables en las oficinas nacionales de estadística mediante un análisis exploratorio para identificar el tratamiento específico que requiere la información.

En el Perú, la revisión y validación de datos de los registros administrativos de la SUNAT se realiza en forma automatizada. Se cuenta con manuales de consistencia básica cuyas reglas han sido programadas con el apoyo de personal informático. Los informes de inconsistencias son luego revisados por un equipo de personas capacitado para el desarrollo de esta labor.

D. Edición e imputación

Este subproceso tiene como objetivo mejorar la precisión y credibilidad de los datos del registro económico. Consiste en la revisión de los datos de entrada y se realiza con la finalidad de detectar datos erróneos, faltantes, incompletos, incongruentes o poco confiables que sea necesario corregir, pudiéndose también insertar nuevos valores.

Cabe resaltar que el registro estadístico ha sido creado con datos recopilados de varios registros administrativos, que disponen de una clave de identificación común y cuyos datos corresponden a las necesidades del tema estadístico que se está tratando. Esos datos ya han pasado por procesos de edición realizados por la entidad generadora, que priorizó esa información para su uso administrativo.

Por esta razón, se recomienda que, cuando se realice la edición en el registro estadístico, las oficinas nacionales de estadística pongan atención a la revisión de los datos que son de mayor utilidad para el uso con fines estadísticos.

Con la edición, se busca que el registro estadístico esté limpio y coherente, a fin de que no haya disparidades entre los registros individuales. En el caso de que se presenten incongruencias es necesario revisarlas, porque podrían deberse a errores del sistema en los registros administrativos.

Como habitualmente las ONE realizan la edición en sus operaciones estadísticas, también para el registro estadístico económico, primero se controlará cada registro mediante comprobaciones lógicas internas a fin de detectar errores o duplicados. Estas podrán referirse a reglas de flujo de información, reglas lógicas de validación de datos y reglas para el control de cobertura o la verificación de valores extremos, entre otras.

Los resultados de aplicar estas reglas de consistencia orientarán lo que se debe hacer para corregir los errores o duplicados que se puedan encontrar. Si se detectaran errores sistemáticos producidos en los registros administrativos económicos de donde se obtuvieron los datos, es primordial que se corrijan y habrá que acudir a los responsables de dichos registros en las entidades generadoras, por lo que es importante mantener un buen contacto con ellos.

En los registros estadísticos económicos también es usual observar datos faltantes. Esto se puede deber a diversas situaciones, tales como: i) cuando existen datos incongruentes que provienen de los registros administrativos, en este caso es necesario revisar la documentación y el análisis del proceso de gestión realizado al registro para asegurarse de que realmente el dato no tiene un valor y es necesario realizar una imputación; ii) cuando el dato no tiene un valor, por lo que también será necesario realizar un análisis exploratorio para decidir su corrección, y iii) cuando, al unir varios registros administrativos, se observa que no hay vinculación entre ellos, lo que puede deberse a que algunas unidades económicas no se encuentren. En este caso, se recomienda que el análisis exploratorio se efectúe a ese nivel.

De acuerdo con cada situación, si se comprueba que verdaderamente existen datos faltantes, que no se hayan podido corregir a través de los responsables de los registros administrativos en las entidades generadoras o mediante procesos de edición que corrijan incongruencias, se podrá realizar imputaciones en los datos faltantes a fin de completar el registro estadístico.

A continuación se mencionan algunos de los posibles métodos de imputación (Useche y Mesa, 2006):

- Técnicas determinísticas, cuando al repetir la imputación en varias unidades bajo las mismas condiciones se producirán las mismas respuestas (métodos de imputación de la media o moda, imputación por regresión y de modelos de tiempo).
- Técnicas fundamentales en información externa, cuando se basan en variables relacionadas con una encuesta perteneciente a otras bases de datos o reglas previas (métodos deductivos, tablas de consulta).

E. Derivación de nuevas variables y unidades

Este subproceso tiene como objetivo crear nuevas variables que no se encuentran en ninguno de los registros administrativos o estadísticos que se integraron y no estuvieron programadas en la recolección de los datos, pero que son necesarias para atender las demandas de los usuarios.

Estas variables y unidades se crean a partir de otras variables de los registros administrativos o registros estadísticos que se han vinculado. Tienen como propósito mejorar la precisión y la credibilidad, así como la coherencia y la comparabilidad, dimensiones o atributos priorizados para la calidad por la OCDE.

La derivación de nuevas variables y unidades puede ser repetitiva durante el proceso, debido a que se basarán en otras variables derivadas, por lo que se recomienda poner mucha atención a fin de que responda a un orden correcto.

Dado que la obtención de estadísticas basadas en registros administrativos se realiza a partir del diseño para la recolección de los datos de las oficinas que los generan, no es posible para las oficinas nacionales de estadística definir las variables necesarias con fines estadísticos, por lo que se tiene que recurrir a la creación de variables agregadas, basadas en las variables disponibles de los registros administrativos.

Seguí (2016) menciona que las variables derivadas o agregadas se crean de seis modos diferentes:

- i) Por medio de cálculos aritméticos o procedimientos lógicos.
- ii) Mediante agrupamiento de valores.
- iii) Mediante codificación.
- iv) Por asociación de otras variables.
- v) Mediante agregación de otras variables.
- vi) Mediante modelos estadísticos.

También señala que las variables derivadas se pueden generar combinando estos métodos, incluso aplicando cálculos más complejos con procedimientos basados en ciertas reglas, y recomienda que deben ser documentadas según el estándar definido para tales fines. Es importante describir la fórmula o regla de cálculo utilizada cuando se crea la variable.

F. Finalización de los archivos de datos

Este subproceso incluye las actividades de cierre realizadas en cada uno de los subprocesos de esta fase. Se debe considerar lo siguiente:

- Los archivos de microdatos y los formatos de recolección de datos del registro administrativo deben ser remitidos para su resguardo legal.
- Los archivos de microdatos y bases de datos para la fase de presentación de resultados serán remitidos para el análisis de la información.

- Centralizar la documentación gestionada en esta fase, así como el informe del procesamiento, que será de utilidad para el análisis de datos.
- En los archivos documentarios, deben estar los programas que han servido como fuentes, así como los catálogos para los clasificadores, para las llaves de identificación, de integración, los métodos de imputación, la creación de las variables derivadas que se hayan generado, y todo lo referente a la carga de los datos y metadatos recopilados en un entorno electrónico.
- Respaldos de los sistemas informáticos y la información en medios computacionales para su archivo.
- Con toda esta información se debe crear una carpeta de datos (usualmente macrodatos) que permitirá utilizarla como base para la fase de análisis o, en caso de que sea un proyecto, que se le pueda dar continuidad.

Capítulo VIII

Análisis

Introducción

Esta fase tiene por objeto asegurar que la información producida sea apta para su propósito, es decir, que esté lista para su uso y difusión.

Se deberá examinar el microdato en sus dimensiones básicas tanto en la captura en el registro administrativo como en el procesamiento de los datos para obtener resultados confiables acordes a la visión del registro estadístico de negocios y del sistema de registros administrativos económicos.

A. Análisis del registro de empresa

Este capítulo tiene como objetivo principal examinar el microdato en sus dimensiones básicas, tanto en la captura como en el procesamiento de los datos, con el fin de obtener resultados estadísticos confiables y acordes con la visión del registro estadístico de negocios y de un sistema de registros administrativos económicos (véase el capítulo II). También se pretende establecer los parámetros mínimos de la estructura de los documentos a divulgar.

En esta fase se analizan los procesos relacionados con el instrumento de recolección de información utilizado por la institución responsable del registro empresarial para lograr una visión holística de cara a evaluar la calidad del registro, uno de cuyos problemas puede ser, por ejemplo, el subregistro.

Se busca validar el uso de los clasificadores y las normas de calidad, tales como la CIIU y la certificación ISO 27000, para comprobar la seguridad de la información a fin de generar confianza en los responsables de la fuente y en los titulares de las empresas que deben suministrar la información según el registro administrativo. La implementación y ulterior certificación de estas normas de calidad para los procesos de producción debe contribuir a preservar la integridad y confidencialidad de los datos de las empresas y de todos los procesos institucionales.

Es imperioso contar con la documentación de la operación estadística del registro administrativo según el inventario que realiza la institución rectora del SEN, en el que se visualiza el marco legal que respalda la fuente, los estándares o el clasificador utilizado, la periodicidad o corte de la base datos, entre otras cosas. También se debe especificar el tipo de empresa y los registros administrativos asociados con la empresa a registrar, a saber:

- Sociedades de responsabilidad limitada (SRL), sociedades anónimas y sociedades por acciones simplificadas (SAS)

Conlleva el análisis de los siguientes registros administrativos:

- Formulario de Declaración Jurada para el Registro y Actualización de Datos de Sociedades (RC-02)
- Certificado de registro mercantil.
- Recibo de pago por constitución de compañía.
- Certificado de nombre comercial emitido por la Oficina Nacional de la Propiedad Industrial (ONAPI).
- Empresas individuales de responsabilidad limitada (EIRL)
- Formulario de Declaración Jurada para el Registro y Actualización de Datos de Sociedades (RC-02).
- Acto constitutivo registrado en la Cámara de Comercio.

- Certificación de depósito bancario
- Certificado del registro mercantil.
- Certificado de nombre comercial emitido por la ONAPI.
- Sociedades accidentales o en participación
 - Formulario de Declaración Jurada para el Registro y Actualización de Datos de Sociedades (RC-02).
 - Contrato firmado y sellado por el Registro Civil o el Ayuntamiento del Distrito Nacional.
 - Recibo de pago del capital social autorizado.
- Sociedades en nombre colectivo, en comandita simple y en comandita por acciones
 - Formulario de Declaración Jurada para el Registro y Actualización de Datos de Sociedades (RC-02).
 - Certificado de registro mercantil.
 - Certificado de nombre comercial emitido por la ONAPI.
 - Recibo de pago del capital social autorizado.
- Negocios de único dueño
 - Carta de solicitud firmada, requiriendo la inscripción al negocio de único dueño.
- Sociedades extranjeras
 - Formulario de Declaración Jurada para el Registro y Actualización de Datos de Sociedades (RC-02).
 - Certificado de registro mercantil.
 - Acto constitutivo registrado en la Cámara de Comercio.
- Consorcios
 - Formulario de Declaración Jurada para el Registro y Actualización de Datos de Sociedades (RC-02).
 - Contrato de consorcio registrado por el Registro Civil y Conservaduría de Hipotecas.
 - Certificado del nombre comercial emitido por la ONAPI.
 - Recibo de pago del capital social autorizado.
- Entidades acogidas a leyes de incentivo
 - Formulario de Declaración Jurada para el Registro y Actualización de Datos de Sociedades (RC-02).
 - Resolución o certificación del organismo rector.
 - Certificado de registro mercantil.

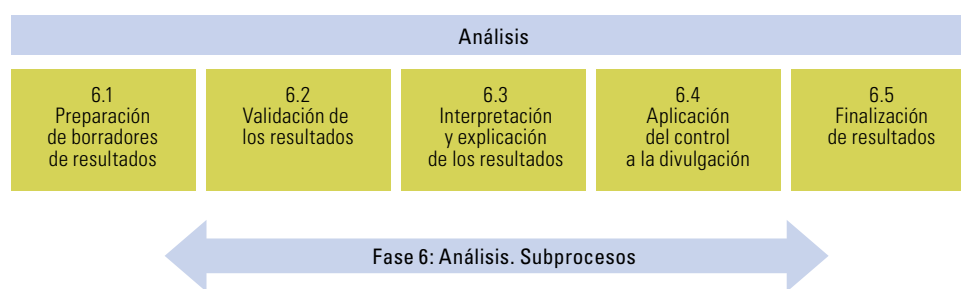
La coherencia entre cada registro administrativo contribuirá a la construcción del registro estadístico de negocios y al análisis que consecuentemente deberá realizarse en cada caso. Para estos fines, la fase 6 del MGPIE se ha dividido en cinco subprocesos, que se describen en el siguiente punto.

B. Subprocesos de la fase de análisis

La fase de análisis se divide en cinco subprocesos relacionados entre sí, organizados secuencialmente del uno al cinco. Sin embargo, pueden realizarse de forma concomitante y hacerlos interactuar entre sí, considerando que cada uno tiene un resultado independiente respecto al otro, tal como se visualiza en el diagrama VIII.1.

Diagrama VIII.1

Subprocesos incluidos en la fase de análisis



Fuente: Comisión Económica para Europa (CEPE), "Modelo genérico del proceso estadístico GSBPM (versión 1.0 en español, septiembre 2016)", Ginebra, 2016.

1. Preparación de borradores de resultados

Partiendo de la demanda identificada en los instrumentos de planificación nacional, el plan estadístico nacional, el plan estadístico sectorial o la priorización estadística del país, y considerando las normas y procedimiento para la elaboración de los productos estadísticos económicos y ambientales, en este subproceso se determinan los indicadores de interés para la toma de decisiones tanto del sector público como del privado. También se identifican las variables para la realización del tabulado, el diseño de tablas, los gráficos y la integración geoespacial de los datos y, si fuese necesario, el nivel de desagregación de los datos de las empresas: territorial, tamaño de la empresa, número de empleados por sexo, actividad económica, tipo de sociedad, tipo de contribuyente, entre otros.

Con aspectos como estos se pueden generar los datos capturados y procesados provenientes de la fase 5 (procesamiento) y obtener resultados estadísticos a partir de ellos. Sin embargo, estos resultados deben ser analizados y comprobarse su calidad.

2. Validación de los resultados

En este subproceso se analiza el comportamiento de los datos procesados, en busca de patrones y tendencias implícitas. También se evalúa si la calidad del dato es óptima para la publicación.

En ese sentido, es importante determinar el subregistro y la tasa de respuesta según formularios o registro administrativo. Se procede a la comparación con los resúmenes de otras instituciones que han servido como fuentes para el período de estudio, la verificación del resultado de las consultas y los cruces de variables de las diferentes tablas según los registros administrativos que componen el registro de la empresa (véase el tipo de empresa en el apartado 8.1)⁵. Por último, se realiza la comprobación de la declaración de la empresa respecto al número de empleados, tipo de empresa, entre otros.

Un aspecto fundamental es la verificación de la ubicación de la empresa respecto a la dirección declarada. Esto se refiere principalmente a las empresas medianas y las microempresas, que tienden a cambiar de dirección si sus propietarios se mudan de una comunidad a otra. Ello repercute negativamente en el registro estadístico, por ejemplo, cuando se cruzan estos datos con las variables ambientales y el impacto en el territorio, el aporte al producto interno bruto (PIB) según comunidad, el empleo según territorio, las políticas de movilidad, la oferta educativa a nivel de grado académico o formación técnica, entre otros.

⁵ Objeto en una base de datos que se organiza en filas y columnas, parecido a una hoja de cálculo de Excel.

La ubicación geográfica y la actividad económica deben clasificarse de acuerdo con los catálogos oficiales. En el proceso de captura de datos, los registros administrativos de las empresas deben pasar por la validación instantánea de las características de ubicación geográfica, mediante un sistema de catálogos que asegure la jerarquía de las entidades geográficas. La oficina nacional de estadística para estos fines debe contar con el código geográfico de las empresas y establecimientos, a través de la versión oficial de la división territorial del país.

Por otra parte, los clasificadores deben verificarse con las descripciones de la actividad económica, determinando si estas tienen los elementos necesarios para la clasificación en alguna de las clases de actividad económica de la CIIU.

Una vez detectadas las inconsistencias o errores, se debe determinar el tratamiento para su corrección, basado en metodologías debidamente establecidas, así como preparar la documentación de todo el proceso hasta contar con resultados óptimos.

3. Interpretación y explicación de los resultados

Es evidente que una correcta interpretación de los resultados depende de la capacidad analítica de los técnicos responsables del tema. Sin embargo, dejar una tarea de suma importancia a criterios personales es un riesgo, por lo que se precisa de una metodología documentada que incluya los métodos económicos más apropiados para explicar el dato, indicando el soporte y su referencia o la fuente utilizada para la interpretación.

Cada resultado conlleva inferencias basadas en los reportes, tablas y gráficos con análisis del comportamiento de las variables según el tipo de empresa, asociados a los indicadores macroeconómicos y al desarrollo empresarial. Se debe visualizar el crecimiento, las fortalezas y las debilidades del sector empresarial según la rama económica.

Al integrar los datos con la información geoespacial se podrá obtener una visión holística del comportamiento de las variables que componen el registro estadístico.

4. Aplicación del control a la divulgación

Dos aspectos fundamentales para garantizar la continuidad del registro estadístico de negocios son: a) la confidencialidad, y b) la anonimización.

Al definir si el documento se dará a conocer por medios impresos o digitales, así como la disponibilidad de la base de datos y el nivel de desagregación (ya sea en microdatos o valores agregados), el documento a divulgar debe estar libre de información identificativa de la empresa. En ese sentido, la oficina nacional de estadística debe procesar los datos correspondientes a cada empresa y excluir de los resultados este tipo de información. Es un derecho de los informantes, salvo que por escrito se especifique lo contrario.

Dado que la información a publicar debe estar libre de identificación de la empresa, los datos deben pasar por un proceso de anonimización, incluida la desagregación territorial y las informaciones según el tipo de empresa.

De acuerdo con lo anterior, el marco legal de la oficina nacional de estadística y las disposiciones regulatorias para el acceso a la información son instrumentos que pueden facultar a dicha oficina para garantizar estos aspectos fundamentales.

5. Finalización de resultados

Se procede a la elaboración de un informe anual acorde a la estructura definida previamente, el marco conceptual y metodológico, los antecedentes, la definición y los objetivos, la cobertura por actividad económica, las definiciones y variables incluidas, la división territorial usada, la guía de validación, la clasificación y codificación de las actividades económicas y actividades auxiliares, entre otros factores.

Sobre las empresas, se debe definir su actividad económica, el sector de industrias manufactureras, el tamaño de las empresas, su distribución geográfica, entre otros factores.

Capítulo IX

Difusión

Introducción

En esta fase se maneja la publicación de los productos estadísticos a las personas usuarias. Incluye todas las actividades asociadas con ensamblar y publicar un rango de productos dinámicos y estáticos por medio de distintos canales. Estas actividades ayudan a las personas usuarias a acceder a los datos publicados por la organización estadística y a utilizarlos.

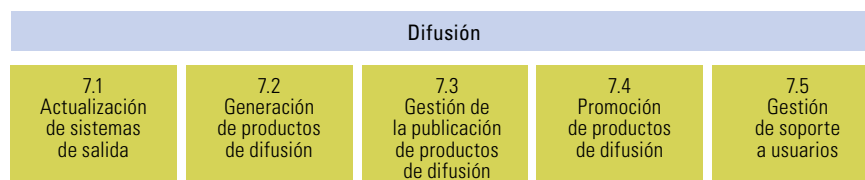
Esta fase representa la interlocución entre la oficina nacional de estadística y las personas usuarias finales de la producción estadística, que pueden ser:

- organismos públicos en los diferentes niveles de gobierno;
- empresas y entidades representativas de sectores económicos;
- investigadores, y
- la sociedad en general.

Esta fase consta de cinco subprocesos que generalmente son secuenciales, pero que también pueden ocurrir en paralelo y ser iterativos. Estos subprocesos se detallan en el diagrama IX.1.

Diagrama IX.1

Subprocesos incluidos en la fase de difusión



Fuente: Comisión Económica para Europa (CEPE), "Modelo genérico del proceso estadístico GSBPM (versión 1.0 en español, septiembre 2016)", Ginebra, 2016.

A. Actualización de sistemas de salida

Este subproceso contempla la producción final de las tablas de resultados con fines de disseminación y actualización de los metadatos. Incluye:

- dar formato a los datos y metadatos listos para colocarlos en bases de datos de salida;
- cargar datos y metadatos en las bases de datos de salida, y
- asegurarse de que los datos estén vinculados con metadatos relevantes.

Los procedimientos de formatear, cargar y vincular los metadatos deben realizarse preferentemente en fases anteriores, pero este subproceso incluye una revisión final para asegurarse de que todos los metadatos necesarios estén en su lugar y listos para ser difundidos.

Los resultados estadísticos deben ser acompañados por metadatos adecuados que permitan a las personas usuarias recuperarlos, interpretarlos correctamente y formar una opinión sobre su calidad.

De acuerdo con el *Data and Metadata Reporting and Presentation Handbook* (OCDE, 2007, pág. 126), es importante incluir la información siguiente en los metadatos: i) la fuente de los datos administrativos; ii) una

descripción de las finalidades para las que los datos fueron originalmente compilados y recolectados por el órgano administrativo; iii) un análisis de los puntos fuertes y débiles de los datos en términos de su aplicación estadística (cobertura, diferencias en los conceptos en relación con las recomendaciones internacionales, utilización de clasificaciones no estandarizadas, entre otras cosas); iv) una descripción del procesamiento por la oficina nacional de estadística después de recibir los datos administrativos, y v) una descripción de la confiabilidad de los datos, incluida la adhesión a las normas y estándares internacionales y las posibles limitaciones sobre su uso estadístico.

B. Generación de productos de difusión

En este subproceso se obtienen los productos pertinentes, tal y como se habían diseñado en el subproceso 2.1, para cubrir las necesidades del usuario (por ejemplo, publicaciones impresas, comunicados de prensa y sitios web). Los productos pueden adquirir muchas formas, incluidos gráficos interactivos, tablas, conjuntos de microdatos de uso público y documentos descargables. Entre los pasos típicos de este proceso se incluyen:

- preparar los componentes del producto (textos de explicación, tablas, gráficos, declaraciones de calidad, entre otros);
- ensamblar los componentes en productos, y
- editar los productos y revisar que cumplan con los estándares de publicación.

C. Gestión de la publicación de productos de difusión

Este subproceso se asegura de que los elementos a ser divulgados estén en su lugar. Incluye gestionar el momento en que serán publicados, informar a grupos específicos como la prensa o los ministros y coordinar los arreglos para cualquier embargo antes del lanzamiento; la provisión de productos a los suscriptores y la gestión del acceso a datos confidenciales a grupos de personas usuarias autorizadas (como los investigadores). En ocasiones, una oficina nacional de estadística debe retirar un producto (por ejemplo, si se descubre un error). Esa posibilidad también se incluye en este subproceso.

D. Promoción de productos de difusión

Mientras que la mercadotecnia puede considerarse como un proceso transversal, este subproceso se enfoca en la promoción activa de los productos estadísticos producidos en un proceso estadístico global para que pueda llegar a la mayor cantidad de personas usuarias de información. Implica el uso de herramientas de gestión de relaciones para establecer de forma efectiva las personas usuarias potenciales del producto, así como otras herramientas tales como sitios web, wikis y blogs para facilitar el proceso de comunicar la información estadística al usuario.

Este subproceso incorpora las actividades de promoción activa de los productos estadísticos de las encuestas estructurales por empresa. Incluye la elaboración de *banners* virtuales, la utilización de infográficos en las publicaciones de redes sociales, la creación de podcasts y pequeños videos para medios como la radio y la televisión, además de la elección de los principales destaques que van a componer el lanzamiento y otros elementos de comunicación social de la oficina nacional de estadística. Este proceso, naturalmente, presupone la interlocución directa con periodistas del Instituto, así como profesionales del área de diseño y de *marketing*.

E. Gestión de soporte a las personas usuarias

Este subproceso permite asegurarse de que se registren las dudas y pedidos del cliente relacionados con servicios, como el acceso a microdatos, y que se proporcionen respuestas puntuales dentro de un rango de tiempo establecido. Las consultas y solicitudes deben revisarse periódicamente para proporcionar un insumo al proceso transversal de gestión de la calidad, ya que pueden indicar necesidades nuevas o cambios en las necesidades de las personas usuarias.

Capítulo X

Evaluación

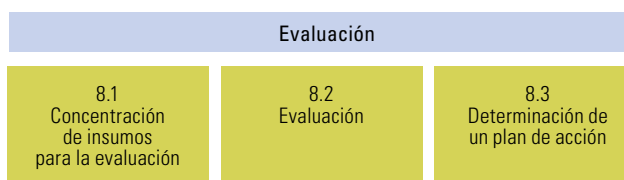
Introducción

Esta fase tiene por objeto decidir si el siguiente ciclo de producción de información debe llevarse a cabo utilizando las mismas especificaciones de necesidades, diseño y construcción, o si se requiere implementar alguna mejora.

En la evaluación de los procesos y subprocesos desarrollados en la construcción del producto estadístico se debe tener en cuenta desde la identificación de la necesidad estadística hasta un producto final satisfactorio y su difusión. Esto permitirá dejar en evidencia las debilidades y fortalezas del proceso y proporcionará información útil para establecer un plan de acción que propenda a la mejora continua del proceso de producción estadística (véase el diagrama X.1).

Diagrama X.1

Subprocesos incluidos en la fase de evaluación



Fuente: Comisión Económica para Europa (CEPE), "Modelo genérico del proceso estadístico GSBPM (versión 1.0 en español, septiembre 2016)", Ginebra, 2016.

A. Características de la evaluación

La evaluación del proceso desarrollado para obtener un producto estadístico implica, en primera instancia, la consolidación de los insumos de entrada y salida que atienden a los aspectos conceptuales, tecnológicos y analíticos sobre los que se construyó el producto estadístico.

Asimismo, para realizar la evaluación se deben considerar:

- las condiciones organizacionales y políticas de la entidad que generó el producto estadístico, y
- los recursos y la logística que permitirían el desarrollo de la evaluación.

Es importante destacar estos aspectos y advertir que tal vez no estén dadas las condiciones propicias para la evaluación.

En este sentido, se propone una evaluación basada en las dimensiones que se describen a continuación.

B. Viabilidad de la evaluación

La formulación explicada, estructurada y normada de este proceso será una condición importante de la factibilidad de la evaluación.

La viabilidad de la evaluación se define como “la posibilidad que tiene una intervención pública para ser evaluada” (MIDEPLAN, 2012, pág. 22; INAP, 2019).

En este sentido, además de los aspectos conceptuales y analíticos de la propuesta a evaluar, deberán tenerse en cuenta:

- las condiciones organizacionales y políticas para realizar la evaluación, y
- los recursos, la logística y la organización necesaria para realizar la evaluación.

Es importante destacar estos aspectos y advertir a los encargados de realizar esta tarea, ya que puede ocurrir que no se cuente con las condiciones propicias para la evaluación.

1. Dimensión institucional

La línea 1 de evaluación tiene como objetivo analizar y explicar las condiciones institucionales e interinstitucionales que promueven el desarrollo de interoperabilidades que otorgan sentido a la utilización de un producto o insumo, como los datos, de los sectores público y privado. Se favorecen y se potencian los procesos de una cultura estadística más rica y amplia, y a la vez menos costosa, para los actores intervinientes.

2. Objetivos específicos

- Identificar las condiciones que favorezcan u obstaculicen el desarrollo de la interoperabilidad institucional e interinstitucional.
- Indagar sobre las percepciones que tienen los distintos actores institucionales sobre la metodología de trabajo.
- Analizar si el desarrollo del proceso ha favorecido el mejoramiento de las capacidades preexistentes mediante la creación de aprendizajes efectivos en los equipos de trabajo que laboran en conjunto para el desarrollo de la interoperabilidad, y si ha contribuido a una mejor relación entre la oficina nacional de estadística y las entidades poseedoras de registros administrativos.
- Describir las configuraciones interinstitucionales que sostienen y consolidan los procesos de sinergia favorables al desarrollo de las capacidades de las ONE y de las entidades poseedoras de registros administrativos en lo que respecta a la calidad de estos.
- Estudiar las implicancias sobre los costos o cargas de trabajo para los informantes y las ONE.

3. Dimensión técnico-metodológica

La línea 2 de investigación tiene como objetivo analizar aspectos de la estrategia técnico-metodológica de implementación de la fuente de registros administrativos, como insumos de las estadísticas oficiales relativas a un registro estadístico de negocios y de las prácticas administrativas responsables del insumo base para este.

4. Objetivos específicos

- Revisar los subprocesos para el análisis de la cobertura de las necesidades de información estadística o geográfica requeridas.
- Estudiar si los procedimientos aplicados permiten mantener la comparabilidad y coherencia conceptual y con los resultados obtenidos a partir de otras fuentes de información e incluso mediante otros métodos.
- Validar el cumplimiento de los estándares internacionales, los subprocesos implementados para la selección de fuentes y la evaluación de la calidad, la estandarización de datos, la vinculación de registros y la transformación de los registros administrativos a registros estadísticos.

- Verificar si en los procesos llevados adelante se cumplieron los principios de confidencialidad y seguridad de la información.
- Indagar sobre la satisfacción de las necesidades de información entre las personas usuarias internas y externas.
- Verificar la implementación de una solución informática alineada con las normas establecidas y la interoperabilidad del sistema.

5. Abordaje metodológico de la evaluación

a) Línea de base

Para lograr los objetivos propuestos se deberá establecer una línea de base que permita:

- Describir la dinámica actual de funcionamiento de los procesos de captura de datos de los registros administrativos y la construcción de datos estadísticos a partir de esos procesos.
- Construir los indicadores pertinentes.
- Identificar las fuentes de información de las que se obtendrán los datos.

Son ejemplo de ello el estudio de los costos y cargas para informantes y para las oficinas nacionales de estadística en cuanto a la generación de los insumos y datos, mediante operativos tradicionales como censos o encuestas. También lo son la oferta final a las personas usuarias y la información preexistente referida a unidades económicas.

b) Universo de estudio

El universo que se estudia será el conformado por los organismos estatales que elaboran registros administrativos, a partir de los cuales se construyen o pueden construir datos estadísticos, y la oficina nacional de estadística.

Para la construcción de la muestra es necesario preguntarse:

- i) ¿A quiénes se referirá la evaluación?
- ii) ¿Qué niveles institucionales serán considerados?
- iii) ¿Qué actores serán seleccionados?
- iv) ¿Qué tipo de personas usuarias serán incluidas?
- v) ¿Quiénes serán las personas usuarias de los resultados de la evaluación?

Se resalta que las personas usuarias identificadas en el inciso e) serán fuente primaria de información. En cuanto a las fuentes secundarias, se recurrirá a la experiencia de otras ONE en relación con la temática, así como con la información existente en las ONE o registros internos, para conformar la línea de base.

c) Instrumentos de recolección de datos

Para recolectar la información necesaria se acudirá a las técnicas e instrumentos de investigación cualitativa y cuantitativa que resulten pertinentes a los fines de la evaluación.

Se proponen entrevistas y notas de campo. Las entrevistas serán semiestructuradas, con una guía temática y cuestionarios que contienen preguntas que pueden ser abiertas o precodificadas (véase el anexo A8).

d) Entrevista

Se recomienda realizar entrevistas a:

- Directivos y técnicos de la oficina nacional de estadística que formen parte del diseño de los acuerdos interinstitucionales y del trabajo técnico en cuanto al registro estadístico de negocios, respectivamente.
- Directivos de las entidades poseedoras de registros administrativos y administrativos que trabajen con los registros administrativos y formen parte del acuerdo.
- Personas usuarias externas de la información producida por la oficina nacional de estadística.
- Personas usuarias internas de la ONE que utilicen como insumo el producto del registro estadístico de negocios.

e) Notas de campo

Como complemento de las entrevistas, pueden utilizarse las notas de campo en las que se registran los hechos observados directamente por el entrevistador o entrevistadora. Se trata de la interpretación que describe la acción y los entornos, quién habla con quién, entre otras cosas. Se deben incluir comentarios potencialmente útiles, por ejemplo, sobre las características del entorno en el que se reunieron los datos y el tratamiento que se da a su consolidación.

f) Equipo evaluador

La especificidad del proceso a evaluar requiere conocimientos mínimos de sus fases y subfases. Por lo tanto, sería recomendable que quienes lo conformen sean parte de la oficina nacional de estadística o el SEN, aunque no los mismos que realizan la tarea de generar estadísticas de un registro estadístico de negocios.

6. Características a considerar en la evaluación de las dimensiones

En los cuadros X.1 y X.2 se presentan las dimensiones de evaluación que expresan los aspectos estructurales del proceso. Por tratarse de un documento genérico y no de un caso concreto, la propuesta de operacionalización en la columna "Indicadores relativos a" se refiere a componentes sobre los que se deberán construir los indicadores y no a indicadores definidos.

7. Convenios o acuerdos

El acceso a la fuente de registros administrativos de los organismos públicos resulta irregular debido a los cambios de gobierno, así como a los cambios de gestión de un organismo en el marco de un mismo gobierno. Evaluar su implementación resulta importante ya que permitirá determinar las ventajas y limitaciones que posibiliten realizar los ajustes necesarios para favorecer la planificación y la institucionalidad. Es decir, su redacción debe contemplar las necesidades a largo plazo de las estadísticas oficiales, y se han de evitar cláusulas que permitan la interrupción de los acuerdos a criterio de los funcionarios de turno.

8. Necesidades

La puesta a disposición de las diferentes personas usuarias de productos estadísticos nuevos o renovados puede satisfacer las necesidades planteadas, pero también hacer que aparezcan otras nuevas. Así, la evaluación del comportamiento de las necesidades de las personas usuarias permitirá retroalimentar y mejorar el proceso, al mismo tiempo que determinar las secuencias de necesidades-oferta y también de oferta-necesidades. Esto permitirá a las oficinas nacionales de estadística no solamente ceñirse a los planteamientos de las personas usuarias, sino expandir su oferta estadística a partir de fuentes de datos subexplotadas.

Cuadro X.1

Matriz de dimensiones institucionales

Línea I institucional				
Dimensión	Indicadores relativos a	Fuente de información	Técnica o método	Observaciones
Convenios o acuerdos	<ul style="list-style-type: none"> – Cantidad de convenios o acuerdos – Vigencia de los convenios o acuerdos – Organismos con convenios o acuerdos – Tipo de cláusulas – Grado de cumplimiento de los convenios o acuerdos 	<ul style="list-style-type: none"> – Área de relaciones institucionales de las oficinas nacionales de estadística (ONE) – Directivos de las ONE – Directivos de organismos poseedores de registros administrativos – Técnicos de áreas sustantivas de las ONE – Administrativos de organismos poseedores de registros administrativos 	<ul style="list-style-type: none"> – Registros internos de las ONE – Cuestionarios 	<ul style="list-style-type: none"> – Se espera que los acuerdos favorezcan la institucionalidad. – Que contemple las necesidades a largo plazo de las estadísticas oficiales.
Necesidades	<ul style="list-style-type: none"> – Requerimientos anteriores – Cumplimiento de requerimientos anteriores – Nuevos requerimientos – Inventario de fuentes – Relación entre diseño y necesidades anteriores 	<ul style="list-style-type: none"> – Personas usuarias internas de las ONE – Personas usuarias del SEN – Personas usuarias externas 	<ul style="list-style-type: none"> – Registros internos – Cuestionarios – Reuniones 	
Recepción de registros administrativos	<ul style="list-style-type: none"> – Costos y mantenimientos 	<ul style="list-style-type: none"> – Área sustantiva de las ONE – Área sustantiva de organismos poseedores de registros administrativos – Técnicos de áreas sustantivas de las ONE – Administrativos de organismos poseedores de registros administrativos 	<ul style="list-style-type: none"> – Registros internos y de intercambio de las ONE y los organismos poseedores de registros administrativos 	
Confidencialidad	<ul style="list-style-type: none"> – Secreto estadístico – Secreto administrativo – Otros secretos – Cláusulas en convenios o acuerdos – Identificación de subfases con datos no anonimizados 	<ul style="list-style-type: none"> – Área jurídica de las ONE – Área jurídica de organismos poseedores de registros administrativos – Técnicos de áreas sustantivas de las ONE – Administrativos de organismos poseedores de registros administrativos 	<ul style="list-style-type: none"> – Marcos legales de las ONE y los organismos poseedores de registros administrativos – Registros internos de las ONE – Registros internos de organismos poseedores de registros administrativos – Cuestionarios 	<ul style="list-style-type: none"> – Se deberán considerar las rigideces y flexibilidades de los marcos legales que dificulten o favorezcan el acceso a las fuentes administrativas.
Costos y beneficios	<ul style="list-style-type: none"> – Costo de operaciones estadísticas generales – Costo de operaciones estadísticas de registros administrativos – Frecuencia de consulta a informantes por áreas de las ONE – Dispersión o centralización de consultas – Personas usuarias de las ONE, del SEN, externas – Tiempos de acceso a los datos – Operaciones estadísticas de registros administrativos con relación al total 	<ul style="list-style-type: none"> – Área contable de las ONE – Administrativos del área – Técnicos de áreas sustantivas de las ONE – Administrativos de organismos poseedores de registros administrativos 	<ul style="list-style-type: none"> – Registros internos de las ONE – Cuestionario 	<ul style="list-style-type: none"> – El costo deberá considerarse desde el punto de vista del costo-beneficio global para las ONE y no solo como una ecuación monetaria unitaria.

Fuente: Instituto Nacional de Estadística y Censos (INDEC) de la Argentina.

Cuadro X.2
Matriz de dimensiones técnico-metodológicas

Línea II técnico-metodológica				
Dimensión	Indicadores relativos a	Fuente de información	Técnica o método	Observaciones
Diseño conceptual y analítico	<ul style="list-style-type: none"> - Inventario de fuentes - Metadatos definidos en el diseño - Salidas diseñadas - Porcentaje disponible con relación al diseño y las necesidades de las oficinas nacionales de estadística (ONE) - Armonía conceptual - Cobertura geográfica - Identificador utilizado - Clasificadores 	<ul style="list-style-type: none"> - Área sustantiva: ONE - Área sustantiva de organismos poseedores de registros administrativos - Técnicos de áreas sustantivas de las oficinas nacionales de estadística - Técnicos de áreas sustantivas de organismos poseedores de registros administrativos 	<ul style="list-style-type: none"> - Registros internos y de intercambio entre las ONE y los organismos poseedores de registros administrativos - Cuestionarios 	
Recepción de registros administrativos	<ul style="list-style-type: none"> - <i>Software</i> o plataforma de tecnología de la información y las comunicaciones (TIC) - Instrumentos de captación - Identificador(es) - Almacenamiento - Volumen de datos - Actualización - Seguridad de la información - Confidencialidad - Entidades proveedoras de datos - Control de calidad de entrada - Documentación - Manuales - Clasificadores - Tiempos utilizados 	<ul style="list-style-type: none"> - Áreas de informática de las ONE - Áreas de informática de organismos poseedores de registros administrativos - Áreas sustantivas de las ONE - Técnicos de áreas sustantivas y de informática de las ONE y los organismos poseedores de registros administrativos 	<ul style="list-style-type: none"> - Registros internos de las ONE - Cuestionario 	- Se deberán considerar aspectos que favorezcan la interoperabilidad.
Integración de registros administrativos	<ul style="list-style-type: none"> - Control de calidad - Análisis conceptual - Consistencias de procesamiento - Validación interna - Validación externa - Cambios propuestos en los registros administrativos - Duplicados - Conversión clasificadores - Errores - Imputaciones - Tiempos utilizados 	<ul style="list-style-type: none"> - Áreas de informática de las ONE - Áreas de informática de organismos poseedores de registros administrativos - Áreas sustantivas de las ONE - Técnicos de áreas sustantivas y de informática de las ONE y los organismos poseedores de registros administrativos 	<ul style="list-style-type: none"> - Registros internos de las ONE - Cuestionario 	
Análisis y difusión	<ul style="list-style-type: none"> - Validación externa - Consistencias - Cobertura de población - Series - Metadatos - Clasificaciones - Anonimización - Datos y metadatos de salida - Componentes de productos - Calendario de difusión - Interlocución con las personas usuarias y prensa - Soporte a las personas usuarias 	<ul style="list-style-type: none"> - Técnicos de áreas sustantivas de las ONE - Técnicos de difusión de las ONE - Personas usuarias externas - Periodistas acreditados 	<ul style="list-style-type: none"> - Registros internos de las ONE - Cuestionario 	
Asistencia técnica y registración	<ul style="list-style-type: none"> - Metadatos de las ONE - Metadatos de organismos poseedores de registros administrativos - Oferta de capacitaciones - Capacitaciones realizadas - Reuniones intercambio - Solicitudes de cambios en los registros administrativos - Plazos de realización 	<ul style="list-style-type: none"> - Área de relaciones institucionales de las ONE - Directivos de las ONE - Directivos de organismos poseedores de registros administrativos - Técnicos de áreas sustantivas de las ONE - Administrativos de organismos poseedores de registros administrativos 	<ul style="list-style-type: none"> - Registros internos - Cuestionarios 	

Fuente: Instituto Nacional de Estadística y Censos (INDEC) de la Argentina.

9. Diseño conceptual y analítico

Fase clave de todo el proceso estadístico donde se diseñan los productos estadísticos, los conceptos, las metodologías, los instrumentos de recolección y los procesos operativos, entre otros aspectos. Se pone a prueba al trabajar con insumos diseñados por otros actores y otros fines. La evaluación de estos aspectos permitirá determinar la calidad de una fuente externa y su adaptación a las necesidades de la oficina nacional de estadística.

10. Recepción de los registros administrativos

El flujo de datos de fuentes externas a la oficina nacional de estadística repercute en la calidad de su trabajo y en la oferta veraz y oportuna de sus productos, así como en la dinámica interna del proceso de trabajo.

El seguimiento y evaluación de esta fase resulta fundamental para el buen desempeño, mejora y planificación de la producción de estadísticas oficiales.

11. Integración de los registros administrativos

A partir de la propuesta de diseño de la oficina nacional de estadística, esta fase se enfrenta con los datos reales con diseños diferentes a las necesidades estadísticas. El control de consistencia, de calidad y validación será más intenso comparado con el efectuado en la fase de captura. La evaluación de esta etapa supondrá la posibilidad de puesta a punto para lograr productos estadísticos acordes al diseño original.

12. Análisis y difusión

El cambio de fuente o la incorporación y combinación con las ya preexistentes puede requerir esfuerzos interpretativos y explicativos, tanto en el equipo de trabajo interno como a nivel de las personas usuarias. Al evaluar el impacto en el proceso de trabajo y en los resultados finales se podrán identificar aspectos para cambiar en favor de un mejor resultado.

13. Confidencialidad

Una de las limitantes a la hora de disponer de los datos de fuentes de registros administrativos tiene que ver con los marcos legales que establecen diferentes tipos de secreto que existen según el fenómeno en cuestión: fiscal, administrativo, médico, estadístico, entre otros.

Este aspecto se convierte en central para poder evaluar su impacto en un proceso que pretende superar la coyuntura y transformarse en una política pública.

14. Costos y beneficios

El enfoque economicista que concede primacía a los factores económicos sobre los de cualquier otra índole resulta insuficiente a la hora de evaluar un proceso como el que se describe en este documento.

Se deberá considerar el enfoque de la relación costo-beneficio global para el SEN y, por lo tanto, para el conjunto de las personas usuarias. Se refiere al costo-beneficio entendido como una mayor y mejor oferta estadística oficial a partir de la posibilidad de disponer de un caudal mayor de datos, tanto cuantitativamente como en términos de una mayor variedad de fenómenos y también de alcances geográficos.

15. Asistencia técnica y registración

La generación de datos propios por parte de las oficinas nacionales de estadística se logra mediante las fuentes tradicionales: los censos y las encuestas. Estas operaciones estadísticas entrañan, entre otros aspectos metodológicos, el diseño de los instrumentos de captación y los metadatos.

En el caso de los registros administrativos y, como se menciona a lo largo del documento, su diseño corresponde a los organismos poseedores de registros administrativos.

Si se utilizan como fuente estos registros administrativos, resulta necesario que en la dinámica de intercambio con los organismos poseedores de registros administrativos se vayan diseñando los mecanismos de conversión de los instrumentos administrativos a los diseños metodológicos de las ONE. Esto se debe hacer sin incidir ni interferir en las funciones específicas de los entes administrativos y siempre considerando el interés mutuo.

La evaluación de estos aspectos resulta crucial para el mejor y constante desempeño del proceso.

Capítulo XI

Conclusiones y recomendaciones

Introducción

Las oficinas nacionales de estadística de los países de América Latina y el Caribe enfrentan dificultades similares al tratar de utilizar los registros estadísticos administrativos, lo que trae consigo problemas para explotar su aprovechamiento estadístico.

El presente documento está encaminado a compartir recomendaciones con los países de América Latina y el Caribe para tratar de buscar alternativas que permitan aprovechar de la mejor manera posible el potencial estadístico de los registros administrativos, ya que son una alternativa muy buena para la generación de estadísticas.

El documento presenta como visión la construcción de los cuatro registros base que identifican Wallgren y Wallgren (2007 y 2012). Sin embargo, se circunscribe al tema económico, y presenta la propuesta de construir un sistema de registros administrativos económicos.

Es muy importante tener en cuenta la visión de un sistema de registros administrativos económicos, tal y como lo proponen Wallgren y Wallgren (2007 y 2012). Ello permitirá a las ONE de los países de América Latina y el Caribe dar pasos firmes en dicha dirección, sin importar el nivel de desarrollo estadístico que tengan ni la disponibilidad y calidad de los registros administrativos, ya que la identificación de los problemas es el primer paso para tratar de corregirlos.

Cada país, en función de sus recursos, tendrá la posibilidad de atender algunas de las recomendaciones propuestas y los ejemplos de buenas prácticas les permitirán conocer la manera en que otros países han resuelto algún problema similar.

Sin lugar a duda, la participación de los países de la región resultó de gran importancia para la elaboración del documento, ya que refleja la realidad y los retos que enfrentan durante la generación de estadísticas económicas, que en gran medida pueden ser resueltos con el desarrollo del sistema de registros administrativos económicos.

A. Hacia un sistema de registros administrativos económicos

El documento presenta la visión del sistema de registros administrativos económicos, en el cual el registro estadístico de negocios es el principal componente.

El hecho de contar con un registro estadístico de negocios robusto permitirá, entre otras cosas, crear marcos de muestreo para la realización de encuestas económicas, generar indicadores de demografía de los negocios, extraer directorios con información de las unidades económicas correspondientes a encuestas probabilísticas y no probabilísticas y obtener estadísticas económicas que combinan información de varias fuentes.

La conformación de un único sistema de registros administrativos económicos basado en un registro estadístico de negocios y en registros administrativos que sea mantenido por la oficina nacional de estadística de cada país permitirá a los países contar con una infraestructura actualizada de información económica, con lo que podrán obtener marcos de muestreo, directorios y la posibilidad de generar nuevas o más oportunas estadísticas económicas.

La idea fundamental es que los países de América Latina y el Caribe tengan elementos que les permitan avanzar en la consolidación de un registro estadístico de negocios. Este puede actualizarse periódicamente mediante registros administrativos económicos y, en los países donde se realicen, por los censos económicos.

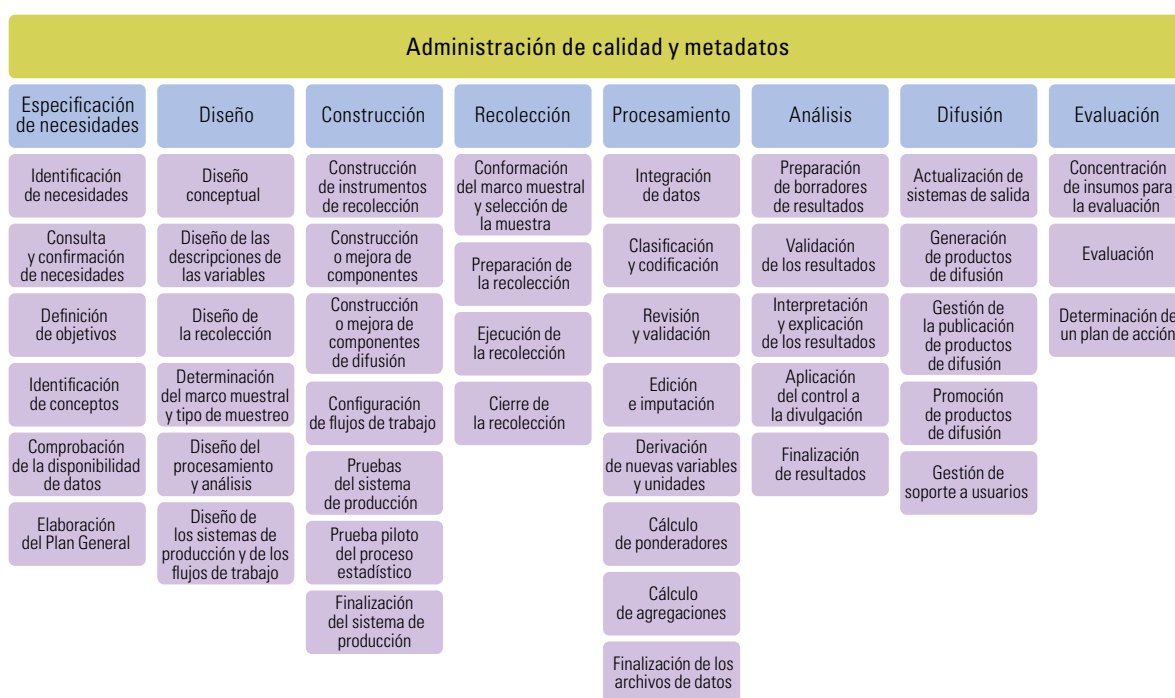
B. Modelo del proceso estadístico (MGPIE)

Para la elaboración de las recomendaciones se siguió el modelo establecido para estandarizar los procesos estadísticos desarrollado por la CEPE y que se denomina MGPIE, particularmente la versión 1.0 (español).

En el diagrama XI.1 se muestran las ocho fases del MGPIE y los subprocesos correspondientes.

Diagrama XI.1

Fases y subprocesos del Modelo Genérico de Procesos Institucionales Estadísticos (MGPIE)



Fuente: Comisión Económica para Europa (CEPE), "Modelo genérico del proceso estadístico GSBPM (versión 1.0 en español, septiembre 2016)", Ginebra, 2016.

C. Detección y análisis de las necesidades

Para la construcción del sistema de registros administrativos económicos es necesario analizar las demandas de información que se tienen a nivel de país, a nivel institucional o mediante alguna solicitud explícita de algún organismo internacional o de un ente gubernamental. Es importante incluir las necesidades de información de todas las personas usuarias y evitar la duplicación de esfuerzos. Dichas necesidades deben ser debidamente documentadas y priorizadas, con el propósito de identificar fuentes de información disponibles para atenderlas de manera eficaz y eficiente.

Se recomienda no pretender desde un inicio tratar de atender todas las demandas de información. Es necesario considerar metas alcanzables en corto tiempo, sobre la base de la información de los registros administrativos y los datos estadísticos accesibles, que reúnan los requisitos de calidad esperados, y posteriormente definir un plan de mejoras sobre la base de objetivos y metas para alcanzar a largo plazo.

Es muy importante incluir como parte de las necesidades de información que todas las variables posibles se desagreguen por sexo, lo que permitirá hacer visible el papel de la mujer en diferentes ámbitos de la economía del país.

Se sugiere seguir todos los subprocesos identificados en el modelo MGPIE: a) identificación de las necesidades; b) confirmación de las necesidades; c) definición de los objetivos; d) identificación de los conceptos; e) comprobación de la disponibilidad de datos, y f) elaboración del plan general.

D. Diseño

Una vez que se han documentado y priorizado las necesidades de las personas usuarias, tanto internas como externas, y que se ha identificado la disponibilidad de los datos para satisfacerlas, se realiza el diseño del sistema de registros administrativos económicos, cuyos principales objetivos son los siguientes:

- Atención a las necesidades de las personas usuarias con base en las prioridades y disponibilidad.
- Identificación de las unidades estadísticas y uso de un identificador único para cada una de ellas.
- Conformación de un registro estadístico de negocios que se actualice periódicamente mediante registros administrativos y, en su caso, mediante censos económicos.
- El sistema de registros administrativos económicos solo se crea para ser usado con propósitos estadísticos. Esto se debe a la necesidad de mantener la confidencialidad de la información individual recabada por los sistemas informativos.

En esta etapa se perfilan las estrategias de desarrollo para las fases de construcción, recolección, procesamiento análisis y difusión, por lo que deben priorizarse las metas alcanzables en períodos cortos de tiempo y después definirse un plan de mejoras a un plazo más largo.

E. Construcción

El MGPIE indica que dentro de esta fase se construye y prueba la solución de producción, de manera de asegurar que pueda utilizarse en un ambiente real.

El tipo de solución depende de las necesidades de las personas usuarias y de las posibilidades de cada país. Por ese motivo, la oficina nacional de estadística, de acuerdo con entidades proveedoras de datos, debe establecer un entorno para el intercambio de registros administrativos. En este entorno se deben considerar aspectos de seguridad y almacenamiento para dar respuesta a todos los requerimientos.

Como se ha mencionado anteriormente, el componente principal del sistema de registros administrativos económicos es el registro estadístico de negocios. Para su construcción, se recomienda seguir las directrices de la CEPE (2015).

Las rutinas de extracción pueden ser automatizadas, semiautomatizadas o manuales, dependiendo del nivel de avance informático de las diferentes entidades proveedoras de datos. No obstante, es necesario que la ONE se adapte a todos los escenarios posibles, con tal de capturar los datos necesarios para la construcción del registro estadístico de negocios.

Asimismo, la ONE debe desarrollar los medios tecnológicos de transmisión e interoperabilidad de datos provenientes de registros administrativos, archivos o bases de datos estadísticos o no estadísticos y las técnicas y herramientas para verificar que la información enviada o recibida cumpla con las condiciones de calidad establecidas para el intercambio de datos.

F. Recolección

Ante la realidad de cada país y de cada unidad poseedora de registros administrativos para la oficina nacional de estadística, la recolección de la información puede llevarse a cabo a partir de diversas fuentes de datos, como las tradicionales provenientes de las operaciones estadísticas (censos o encuestas). También puede recurrirse al acopio de datos generados por las unidades productoras de información a partir de los registros administrativos, datos geográficos, imágenes, entre otros, derivados a bases de datos o microdatos.

En los casos en que se recolecte información a partir de registros administrativos, es de gran importancia respetar rigurosamente los acuerdos existentes entre las instituciones que poseen registros administrativos y los calendarios de entrega de la información. La información debe recolectarse con la mayor periodicidad posible.

La ONE debe ser impulsora en su respectivo ámbito del aprovechamiento estadístico de registros administrativos. Por ese motivo es muy recomendable que se conformen grupos de trabajo permanentes en el marco del SEN para promover el aprovechamiento estadístico de los datos que captan las diferentes instituciones de gobierno y privadas. Para ello, se recomienda la revisión de los marcos legales, tanto de la ONE como de las dependencias involucradas, para que permitan la elaboración de convenios de colaboración que faciliten el intercambio de información para fines estadísticos.

G. Procesamiento

Los datos del registro administrativo deben procesarse para el análisis con la finalidad de que puedan convertirse en información utilizable para la producción estadística.

Deben ejecutarse procesos para verificar, limpiar y transformar los datos de entrada, así como estandarizar clasificaciones y códigos, con miras a su integración. Además, los datos deben revisarse, validarse y, en su caso, se han de ejecutar procesos de edición e imputación. Los procesos pueden ser manuales, semiautomáticos o automáticos, pero siempre deben llevarse a cabo para garantizar la calidad de la información.

Cabe destacar la importancia de un identificador común en los registros administrativos para su integración. Después de ser identificados, los procesos deben someterse a una evaluación de calidad estadística.

Se recomienda utilizar alguna herramienta que apoye la evaluación de la calidad de los registros administrativos. Entre estas se encuentran la HECRA y la CECRA.

H. Análisis

La información producida debe ser pertinente y oportuna, es decir, que satisfaga en tiempo y forma las necesidades de las personas usuarias, tanto externas como internas. Para ello es importante revisar la documentación de las necesidades de las personas usuarias.

Además, debe asegurarse la calidad de la información generada, así como la protección de la confidencialidad de la oficina nacional de estadística. Para ello es necesario realizar las siguientes actividades: a) preparación de borradores de resultados; b) validación de resultados; c) interpretación y explicación de los resultados, y d) aplicación de controles de la difusión.

I. Difusión

De este subproceso se obtienen los productos tal y como se diseñaron previamente, para cubrir las necesidades de las personas usuarias (pueden incluir publicaciones impresas o digitales, comunicados de prensa y sitios web). Los productos pueden adquirir muchas formas, incluidos gráficos interactivos, tablas, conjuntos de microdatos de uso público y documentos descargables. Entre otros pasos típicos en este proceso cabe mencionar los siguientes:

- Preparar los componentes del producto (textos de explicación, tablas, gráficos, declaraciones de calidad, entre otros).
- Ensamblar los componentes en productos.
- Editar los productos y revisar que cumplan con los estándares de publicación.

J. Evaluación

Esta fase tiene por objeto determinar si el siguiente ciclo de producción de información debe llevarse a cabo utilizando las mismas especificaciones de necesidades, diseño y construcción, o si se requiere implementar alguna mejora.

La evaluación debe realizarse para todas las fases del ciclo del MGPIE. Se propone la evaluación del proceso en las siguientes dimensiones:

- Convenios o acuerdos de colaboración
- Necesidades
- Diseño conceptual y analítico
- Recolección
- Integración de registros administrativos
- Análisis y difusión
- Confidencialidad
- Costo-beneficio
- Asistencia técnica

Bibliografía

- Asamblea Legislativa de Costa Rica (2019), *Ley del Sistema de Estadística Nacional*, San José.
- Brasil, Presidencia de la República del (2011), *Lei N° 12.527, de 18 de novembro de 2011*, Brasilia.
- CEPAL (Comisión Económica para América Latina y el Caribe) (2020), *Ley Genérica sobre Estadísticas para América Latina* (LC/CEA.10/8), Santiago.
- ___ (2019), "Taller: Big Data, Economía Digital y Desarrollo Sostenible", Santo Domingo [en línea] <https://www.cepal.org/es/eventos/taller-big-data-economia-digital-desarrollo-sostenible>.
- CEPE (Comisión Económica para Europa) (2019), "Modelo genérico del proceso estadístico GSBPM (Versión 5.1, enero 2019)", Ginebra.
- ___ (2016), "Modelo genérico del proceso estadístico GSBPM (versión 1.0 en español, septiembre 2016)", Ginebra.
- ___ (2015), *Guidelines on Statistical Business Registers*, Ginebra.
- ___ (2011), *Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices*, Ginebra.
- ___ (2000), "Terminology on statistical metadata", *Conference of European Statisticians Statistical Standards and Studies*, N° 53, Ginebra.
- DANE (Departamento Administrativo Nacional de Estadística) (2020), *Lineamientos para el proceso estadístico en el Sistema Estadístico Nacional*, Bogotá.
- Gabinete de Ministros (2015), "Sistema Integral de Seguimiento y Evaluación de la Gestión (SISEG)", Buenos Aires.
- INAP (Instituto Nacional de la Administración Pública) (2019), "Seguimiento y evaluación de políticas públicas (IN31860/19)", Buenos Aires [en línea] <https://capacitacion.inap.gov.ar/actividad/seguimiento-y-evaluacion-de-politicas-publicas/>.
- INEGI (Instituto Nacional de Estadística y Geografía) (2020a), *Norma para el Aseguramiento de la Calidad de la Información del Instituto Nacional de Estadística y Geografía*, Aguascalientes.
- ___ (2020b), *Norma Técnica del Proceso de Producción de Información Estadística y Geográfica para el Instituto Nacional de Estadística y Geografía*, Aguascalientes.
- ___ (2019), "Norma Técnica para la incorporación y actualización de información en el Registro Estadístico de Negocios de México", *Diario Oficial de la Federación*, Aguascalientes, 27 de agosto.
- ___ (2018a), *Norma para la Difusión y Promoción del Acceso, Conocimiento y Uso de la Información Estadística que Genera el Instituto Nacional de Estadística y Geografía*, Aguascalientes.
- ___ (2018b), *Norma Técnica para la Clasificación Nacional de Delitos para Fines Estadísticos*, Aguascalientes.
- ___ (2018c), *Ley del Sistema Nacional de Información Estadística y Geográfica*, Aguascalientes.
- ___ (2017), *Presentación de datos Estadísticos en Cuadros y Gráficas: actualización 2017*, Aguascalientes.
- ___ (2015a), *Glosario del SNIEG 2015*, Aguascalientes.
- ___ (2015b), *Lineamientos para Incorporar la Perspectiva de Género en el Sistema Nacional de Información Estadística y Geográfica*, Aguascalientes.
- ___ (2015c), *Norma Técnica para la Elaboración de Metadatos para Proyectos de Generación de Información de Estadística Básica y de los Componentes Estadísticos Derivados de Proyectos Geográficos*, Aguascalientes.
- ___ (2012a), *Planeación y control administrativo en proyectos estadísticos*, Aguascalientes.
- ___ (2012b), *Procesamiento de la información*, Aguascalientes.
- ___ (2012c), *Proceso estándar para el aprovechamiento de registros administrativos*, Aguascalientes.
- ___ (2011a), *Diseño de la muestra en proyectos de encuesta*, Aguascalientes.
- ___ (2011b), *Presentación de datos estadísticos en cuadros y gráficas*, Aguascalientes.
- ___ (2010a), *Norma Técnica para la Generación de Estadística Básica*, Aguascalientes.
- ___ (2010b), *Diseño conceptual para la generación de estadística básica*, Aguascalientes.
- ___ (2010c), *Proceso estándar para encuestas por muestreo*, Aguascalientes.
- MIDEPLAN (Ministerio de Planificación Nacional y Política Económica) (2012), *Manual gerencial para el diseño y ejecución de evaluaciones estratégicas de gobierno*, San José.
- Ministerio de Cultura y Juventud (2019), *Glosario de términos del Ministerio de Cultura y Juventud en el marco del proyecto Sistema de Registros Administrativos de Cultura y Juventud*, San José.

- Ministerio de Educación (2011), *Nuevas voces, nuevos escenarios: estudios evaluativos sobre el Programa Conectar Igualdad*, Buenos Aires.
- Nirenberg, O., J. Brawerman y V. Ruiz (2003), *Programación y evaluación de proyectos sociales: aportes para la racionalidad y la transparencia*, Buenos Aires, Paidós.
- _____(2000), *Evaluar para la transformación: innovaciones en la evaluación de programas y proyectos sociales*, Buenos Aires, Paidós.
- OCDE (Organización de Cooperación y Desarrollo Económicos) (2012), *Quality Framework and Guidelines for OECD Statistical Activities* (STD/QFS(2011)1), París.
- _____(2007), *Data and Metadata Reporting and Presentation Handbook*, París.
- Rojo, A. (2018), "Del dato administrativo de firmas al dato estadístico: desafíos a vencer en América Latina", *CAF Working Paper*, N° 2018/17, Caracas, Banco de Desarrollo de América Latina (CAF).
- Segui, F. (2016), "Marco conceptual y metodológico que sustenta el diseño, desarrollo e implementación de un sistema integrado de registros estadísticos de población e inmuebles", Washington, D.C., Banco Interamericano de Desarrollo (BID).
- _____(2012), "Guía de la herramienta para la evaluación de la calidad de registros administrativos a ser usados con fines estadísticos", Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI).
- SIEMPRO y otros (Sistema de Información, Monitoreo y Evaluación de Programas Sociales y otros) (1999), *Gestión Integral de Programas Sociales Orientada a Resultados: Manual Metodológico para la Planificación y Evaluación de Programas Sociales*, Buenos Aires, Fondo de Cultura Económica (FCE).
- Useche, L. y D. Mesa (2006), "Una introducción a la imputación de valores perdidos", *Terra Nueva Etapa*, vol. 22, N° 31, Caracas, Universidad Central de Venezuela (UCV).
- Vale, S. (2013), "Quality and administrative data", Ginebra, Comisión Económica para Europa (CEPE) [en línea] <https://statswiki.unece.org/display/adso/Quality+and+Administrative+Data>.
- Wallgren, A. y B. Wallgren (2012), *Estadísticas basadas en registros: aprovechamiento estadístico de datos administrativos*, Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI).
- _____(2007), *Register-based Statistics: Administrative Data for Statistical Purposes*, Hoboken, Wiley.

Anexo A1

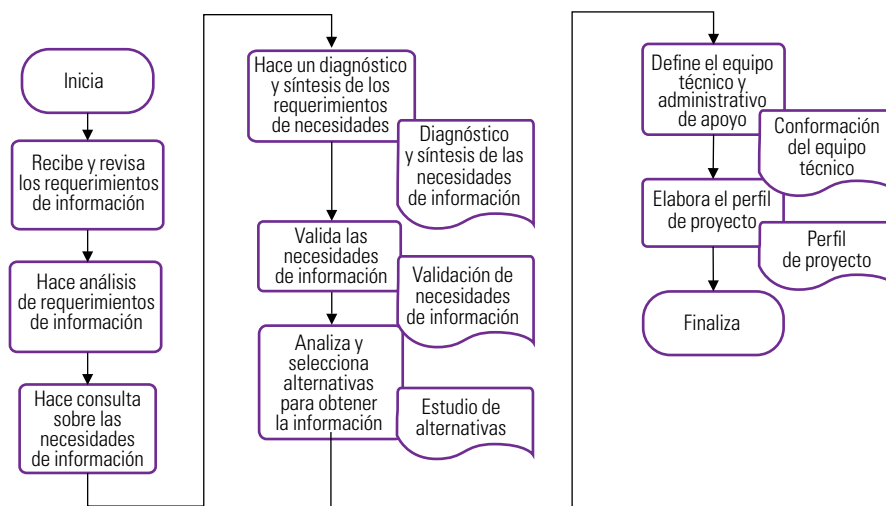
Costa Rica: ejemplos de casos prácticos para cada subprocesso de la especificación de necesidades

En la identificación de necesidades no solamente se incluye la consulta a las potenciales personas usuarias de la información, sino también a posibles entidades proveedoras de información y, además, una revisión de las experiencias con proyectos similares tanto a nivel nacional como internacional, de manera que puedan servir de referencia en la formulación del sistema de registros administrativos económicos.

En el diagrama A1.1 se muestra un ejemplo de un flujograma con las actividades que pueden realizarse en la especificación de necesidades y que permiten mostrar de manera práctica cómo se podría abordar cada subprocesso para la conformación de un sistema de registros administrativos económicos (más adelante se explica en qué consiste cada uno de ellos).

Diagrama A1.1

Flujograma de especificación de necesidades



Fuente: Instituto Nacional de Estadística y Censos (INEC) de Costa Rica, "Identificación de necesidades del Instituto Nacional de Estadística y Censos de Costa Rica (INEC)", San José, inédito.

1. Recibe y revisa los requerimientos de información

La oficina nacional de estadística recibe de las personas usuarias (internas o externas) los requerimientos de información para ser revisados.

Estos requerimientos pueden ser producto de una necesidad a nivel de país o institucional, o bien una solicitud explícita de algún organismo internacional o de un ente gubernamental.

Esta necesidad puede estar directamente relacionada con el uso de los registros administrativos empresariales o alguna otra que requiera una investigación, dado que no implica el uso de registros, o supone una combinación con algún otro método de recolección.

En este punto, lo ideal es contar con un documento que muestre todos los requerimientos iniciales de la solicitud de información, un estudio que evidencie la necesidad o algún otro que sirva de base para sustentar la necesidad y el uso que se le daría a la información, entre otros aspectos.

Para tener una mayor claridad de las necesidades y el sustento de estas, se puede compartir con las personas usuarias una matriz como la que se muestra en el cuadro A1.1⁶.

Cuadro A1.1

Matriz de consulta para la identificación de necesidades de información

Necesidad	Política, estudio o plan institucional

Fuente: Instituto Nacional de Estadística y Censos (INEC) de Costa Rica, "Identificación de necesidades del Instituto Nacional de Estadística y Censos de Costa Rica (INEC)", San José, inédito.

2. Análisis de requerimientos de información

La ONE analiza la solicitud de información con cada uno de los requerimientos, para tener mayor claridad sobre las necesidades y su sustento. De cualquier modo, es un insumo que será reforzado y validado posteriormente con la consulta.

Este análisis puede incluir una investigación más detallada acerca de las necesidades de información requeridas por las diferentes instituciones. Por ejemplo, puede realizarse una búsqueda más exhaustiva en las páginas web de cada una, con el fin de determinar si producen o no alguna información relacionada con las necesidades o si hay alguna otra institución que genere información al respecto, entre otras cosas. Lo anterior contribuye a evitar la duplicación de esfuerzos y avalar que las necesidades no han sido contempladas por ninguna otra institución o iniciativa.

En el cuadro A1.2 se presenta un ejemplo de cómo documentar este análisis de requerimientos.

Cuadro A1.2

Matriz de necesidades de información y existencia de la información

Institución	Necesidades de información	¿Produce información relacionada con su necesidad?	¿Qué información produce?	Otra información que produce relacionada sobre empresas y establecimientos	¿Otra institución produce información para suplir esta necesidad?	Nombre de la institución que produce la información	¿Qué produce?	Observación

Fuente: Instituto Nacional de Estadística y Censos (INEC) de Costa Rica, "Identificación de necesidades del Instituto Nacional de Estadística y Censos de Costa Rica (INEC)", San José, inédito.

3. Consulta sobre las necesidades de información

Al tener la ONE la solicitud de información y la matriz de necesidades de información, debe organizar una consulta. Esta puede consistir en reuniones o sesiones de trabajo con las personas usuarias de la información.

⁶ Esta matriz puede ser compartida internamente en la institución u organización, de manera que, en el momento de ser enviada a la oficina nacional de estadística, incluya las necesidades de manera integral.

En cuanto a estas personas, se deben tener en cuenta tanto los actores nacionales e internacionales como las personas usuarias internas (de la misma ONE). Las necesidades internas deben ser consideradas ya que el objetivo es evitar la duplicación de esfuerzos y no tener que realizar las mismas solicitudes de registros administrativos empresariales a una determinada institución.

El objetivo de las reuniones debe ser explicar el sistema de registros administrativos económicos y permitir a las personas usuarias que expliquen y detallen cada una de sus necesidades de información incluidas en la matriz de necesidades. Estas se deberán tener en cuenta siempre y cuando estén marcadas en los objetivos del proyecto y tengan un respaldo que permita argumentar de forma más robusta la inclusión de sus requerimientos en el diseño de la operación estadística.

4. Diagnóstico y síntesis de los requerimientos de necesidades de información

Una vez realizada la consulta a las personas usuarias de la información, en la ONE se debe realizar un diagnóstico que permita sintetizar los principales requerimientos.

Para ello, se puede elaborar un documento que posea los siguientes apartados⁷:

- i) Introducción.
- ii) Justificación (de dónde proviene el requerimiento).
- iii) Descripción de la solicitud inicial (es una transcripción del documento mencionado en el apartado *Recibir y revisar los requerimientos de información*).
- iv) Antecedentes (donde se detalle el resultado de la revisión bibliográfica temática y conceptual, de la legislación nacional y de las experiencias y recomendaciones nacionales e internacionales).

En este paso se debe realizar una primera identificación de los conceptos. Tal como se mencionó anteriormente, estos conceptos preliminares serán un insumo que se utilizará como parte de la temática del sistema de registros administrativos económicos y que requieren ser medidos. También se tendrá en cuenta el análisis de la información enviada por las personas usuarias para ampliar la temática de las necesidades de información, entre otras cosas.

Para identificar los conceptos, se puede utilizar una matriz como la que se presenta en el cuadro A1.3.

Cuadro A1.3

Matriz de identificación preliminar de conceptos

Temas principales preliminares	Conceptos asociados al tema principal (preliminares)	Definición de conceptos asociados al tema principal (preliminares) ^a	Enlace o fuente

Fuente: Instituto Nacional de Estadística y Censos (INEC) de Costa Rica, "Identificación de necesidades del Instituto Nacional de Estadística y Censos de Costa Rica (INEC)", San José, inédito.

^a Esto puede ser insumo para la fase del diseño conceptual.

A continuación, se detalla en qué consiste cada uno de los apartados de la matriz:

- Temas principales preliminares: son los temas que se lograron definir en la identificación de necesidades y que se podrían tratar en la formulación del proyecto.
- Conceptos asociados al tema principal (preliminares): conceptos que están relacionados con los temas preliminares y que pueden servir para explicarlos, o bien, que se pueden convertir en variables según cada temática.

⁷ Tomado de Instituto Nacional de Estadística y Censos (INEC), "Diagnóstico y síntesis de las necesidades de información", San José, inédito.

- Definición de conceptos asociados al tema principal (preliminares): corresponde a la definición de cada concepto identificado.
- Enlace o fuente: es un respaldo de donde se obtuvo cada definición.

Tal como se ha mencionado, es importante tener en cuenta que pueden existir varios conceptos de diversas fuentes. Más adelante en el diseño del marco conceptual, estas podrían ser un insumo para definir cuál de todas las definiciones implementar en el sistema de registros administrativos económicos.

- i) Requerimientos o necesidades de información (definir a las personas usuarias potenciales a nivel interno (institucional), nacional e internacional, las exigencias de actores (necesidades a satisfacer con respecto a la solicitud).
- ii) Aspectos metodológicos de la solicitud (abordar el alcance metodológico (en términos generales) que deberá tener el registro estadístico de negocios, basado en la identificación de necesidades).
- iii) Conclusiones y recomendaciones.

5. Validación de las necesidades de información

Una parte fundamental del proceso de identificación de las necesidades de información es tener claridad de los requerimientos y lo más importante, que estén dentro del objetivo del sistema de registros administrativos económicos.

Para ello, la ONE requiere realizar una validación de las necesidades de información con cada uno de los actores involucrados (personas usuarias que anteriormente externalizaron sus necesidades de información, con el llenado de la matriz de identificación de necesidades y posteriores reuniones). El objetivo es confirmar el alcance de los requerimientos o introducir los ajustes correspondientes según la consulta realizada a las personas usuarias.

Una estrategia propuesta es realizar un proceso de reconsulta con cada una de las personas usuarias, mediante una segunda o incluso una tercera reunión o sesión de trabajo en que se pueda analizar cada una de las necesidades enviadas en la matriz y ampliar con más detalle, así como aclarar dudas, entre otras cosas.

Para ello, se puede elaborar un documento con la siguiente estructura⁸:

- i) Siglas y acrónimos
- ii) Resumen ejecutivo
- iii) Introducción
- iv) Metodología de trabajo. Consiste en explicar cuáles fueron los pasos que se llevaron a cabo para la validación de necesidades.
- v) Descripción de las instituciones y sus requerimientos. Consiste en documentar con detalle la información de cada uno de los actores consultados en la identificación de necesidades, por ejemplo:
 - Nombre de la institución
 - Antecedentes (lo ideal es que sean relacionados con la solicitud de información)
 - Sustento de las solicitudes y necesidades de información (detallar el respaldo o sustento que proporcionó la institución en el momento de solicitar el requerimiento (esto es parte de la matriz propuesta))
- vi) Análisis de los requerimientos de información. Consiste básicamente en realizar una revisión conjunta de las necesidades de información de actores consultados.
 - Requerimientos de información que podrían obtenerse por medio de registros administrativos (necesidades de información que están relacionadas con los objetivos del sistema de registros administrativos económicos, que se identifican y podrían obtenerse de los registros administrativos).
 - Requerimientos de información descartados (necesidades de información que escapan de los objetivos del sistema de registros administrativos económicos, o bien no corresponden a su cobertura temática)

⁸ Véase Instituto Nacional de Estadística y Censos (INEC), "Caso práctico del proceso de identificación de necesidades de estadísticas de empresas", San José, inédito.

y metodológica, o incluso presentan algún inconveniente en su definición que impide considerarlas dentro del trabajo que debe realizarse; principalmente porque no pueden obtenerse de un registro administrativo, o porque sí existen en registros, pero no se pretende llegar hasta ese nivel de detalle, entre otros).

- Identificación de entidades proveedoras de registros administrativos (las instituciones u organizaciones que podrían ser entidades proveedoras de información y cuáles serían los registros administrativos que se podrían solicitar a cada una).

Ya teniendo una primera identificación de los conceptos, se debe comprobar la disponibilidad de los datos mediante una investigación acerca de cuáles son los registros administrativos existentes que podrían ser los insumos clave para la obtención de estos.

Para ello se sugiere realizar una revisión de las fichas técnicas del Inventario Nacional de Operaciones Estadísticas que, por lo general, disponen los SEN de cada país. Se sugiere, además, realizar búsquedas en la página web de cada persona usuaria o entidades proveedoras de registros que se hayan identificado y programar reuniones con los generadores de información para conocer la disponibilidad y el marco legal, entre otras cosas.

En el cuadro A1.4 se presenta un ejemplo de cómo documentar la comprobación de la disponibilidad de los datos.

Cuadro A1.4

Ejemplo para comprobar la disponibilidad de datos

Institución SEN	Operación estadística	Resumen de ficha técnica ^a	Marco legal

Fuente: Instituto Nacional de Estadística y Censos (INEC) de Costa Rica, "Identificación de necesidades del Instituto Nacional de Estadística y Censos de Costa Rica (INEC)", San José, inédito.

^a Incluye aspectos tales como: nombre de la operación o registro, objetivo, tema, cobertura geográfica, unidad de estudio y periodicidad.

En este paso también es importante analizar el marco legal, no solo de la institución productora del registro administrativo, sino también de la ONE.

- i) Conclusiones y recomendaciones.
- ii) Bibliografía.
- iii) Anexos (personas contactadas en cada institución o algún documento que fundamente alguna de las solicitudes).

6. Analiza y selecciona alternativas para obtener la información

Con el fin de identificar las alternativas, la ONE debe proponer una o varias soluciones para cada una de las necesidades detectadas que deben ser solventadas.

Consiste en valorar las distintas alternativas o estrategias planteadas. Debe tomarse en cuenta que se deben poseer criterios precisos que permitan seleccionar la mejor estrategia de intervención.

La selección de la mejor alternativa (óptima) se hace no solo en función de que sea la más factible en términos económicos, técnicos, legales y ambientales, sino también de que sea pertinente, eficiente y eficaz. Por ejemplo, se puede tomar en consideración aspectos como el presupuesto, el costo, la duración de la ejecución, la dependencia de recursos externos, la contribución a la solución del problema, o bien, en caso necesario, pueden incluirse otros criterios.

En el cuadro A1.5 se presenta un ejemplo de cómo se puede realizar un análisis de alternativas tomando en cuenta los aspectos antes mencionados.

Cuadro A1.5

Ejemplo del análisis de alternativas

c. Criterios de selección y categorías		d. Análisis de cada alternativa		
		Alternativa A	Alternativa B	Alternativa C
		Puntaje	Puntaje	Puntaje
El costo es...	Bajo = 3 Medio = 2 Alto = 1			
La duración es...	Baja = 3 Media = 2 Alta = 1			
La dependencia de recursos externos es...	Baja = 3 Media = 2 Alta = 1			
La viabilidad técnica (recurso humano competente) de la alternativa es...	Baja = 3 Media = 2 Alta = 1			
La complejidad operativa de la alternativa es...	Baja = 3 Media = 2 Alta = 1			
La viabilidad política de la alternativa es...	Baja = 3 Media = 2 Alta = 1			
La probabilidad de conseguir financiamiento para la alternativa es...	Baja = 3 Media = 2 Alta = 1			
La contribución al problema es...	Baja = 3 Media = 2 Alta = 1			
Puntaje total		0	0	0
b. Indicar una breve descripción de cada alternativa				
a. Anotar el nombre de las alternativas de solución	Alternativa A			
	Alternativa B			
	Alternativa C			

Fuente: Instituto Nacional de Estadística y Censos (INEC) de Costa Rica, "Identificación de necesidades del Instituto Nacional de Estadística y Censos de Costa Rica (INEC)", San José, inédito.

La alternativa que obtenga el puntaje más alto podría ser la más factible de implementar. No obstante, puede ocurrir que no sea la que finalmente se elija. Podría seleccionarse otra que se considere más factible en función de los criterios de selección valorados, por ejemplo, desde la perspectiva de su contribución a dar solución al problema de desarrollo⁹.

7. Define el equipo técnico y administrativo de apoyo

Consiste en indicar cuál será el equipo de trabajo básico en la ONE que se requerirá para abordar la fase de especificación de las necesidades y, de ser posible, cuál sería su participación.

No solamente se debe incluir el personal técnico, sino también el equipo de apoyo (financiero, de recursos humanos, de tecnologías de la información y las comunicaciones (TIC), entre otros).

8. Elabora el perfil del sistema de registros administrativos económicos

Este documento se puede estructurar de la siguiente manera:

- Identificación del problema o necesidades de desarrollo. En este caso se debe documentar el problema o la necesidad de desarrollo, de manera que se describa la situación negativa (necesidad) percibida por las y los involucrados, que ha de ser analizada y solucionada.

⁹ Tomado de Instituto Nacional de Estadística y Censos (INEC), "Instructivo para la identificación, formulación, ejecución y evaluación de proyectos", San José, inédito.

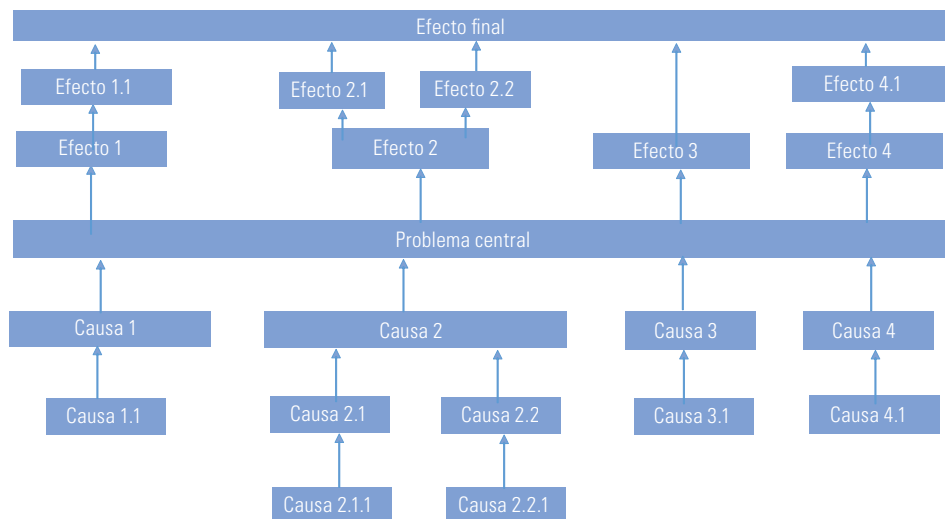
Hay que recordar que el problema o necesidad de desarrollo puede derivarse de diagnósticos o estudios previos, o a partir del criterio o juicio de las personas involucradas.

Una buena forma de documentar esas necesidades es mediante la determinación de las causas (conviene recordar que estas causas se convertirán en las posibles alternativas que se deben identificar para solventar la necesidad) y los correspondientes efectos (resultados o consecuencias de tener esa necesidad).

Una manera de visualizar los problemas detectados (necesidades) es estructurarlos como se presenta en el diagrama A1.2.

Diagrama A1.2

Ejemplo de esquema para el análisis de problemas o necesidades



Fuente: Instituto Nacional de Estadística y Censos (INEC) de Costa Rica, "Identificación de necesidades del Instituto Nacional de Estadística y Censos de Costa Rica (INEC)", San José, inédito.

- ii) Conformación del equipo básico. En este caso se debe documentar el equipo definido en el proceso de flujograma llamado *Define el equipo técnico y administrativo de apoyo*.
- iii) Identificación y análisis de grupos involucrados. Durante la identificación de necesidades se deben identificar y analizar los grupos involucrados (aquellos que tienen algún vínculo con el problema o necesidad de desarrollo con el sistema de registros administrativos económicos y que puedan ser parte de la solución).

Para la identificación de grupos de involucrados se puede utilizar el formato que se muestra en el cuadro A1.6.

Cuadro A1.6

Ejemplo para identificar el grupo de involucrados

Grupos involucrados en el problema o necesidad de desarrollo	Vínculo(s) que tiene el grupo de involucrados con el problema o necesidad ^a	Posición del grupo de involucrados en cada expectativa planteada		Observaciones
		Favorable	Negativa (riesgos)	

Fuente: Instituto Nacional de Estadística y Censos (INEC) de Costa Rica, "Identificación de necesidades del Instituto Nacional de Estadística y Censos de Costa Rica (INEC)", San José, inédito.

^a Ejemplo: autorización de presupuestos públicos, aprobación de normativas, decretos, leyes, autorización de plazas, préstamos de activos, apoyo de gremios, apoyo político, apoyo logístico, comunicación asertiva hacia la ciudadanía, cooperación internacional, capacitación, asesoría, entre otros.

Este análisis permite identificar la posibilidad de apoyo para la solución del problema o necesidad en desarrollo. Además, se puede realizar una evaluación de la viabilidad del sistema de registros administrativos económicos.

En caso de que se identifiquen percepciones negativas, estas deben considerarse posibles riesgos a la intervención futura. Por lo tanto, deberán atenderse con prontitud, de manera que se tomen acciones para revertirlos o contrarrestarlos e ir creando la viabilidad técnica, económica, política y operativa de la intervención.

- iv) Análisis de objetivos. La definición de los objetivos consiste en convertir las condiciones negativas de la identificación de necesidades en soluciones.

Pasos para construir el árbol de objetivos¹⁰:

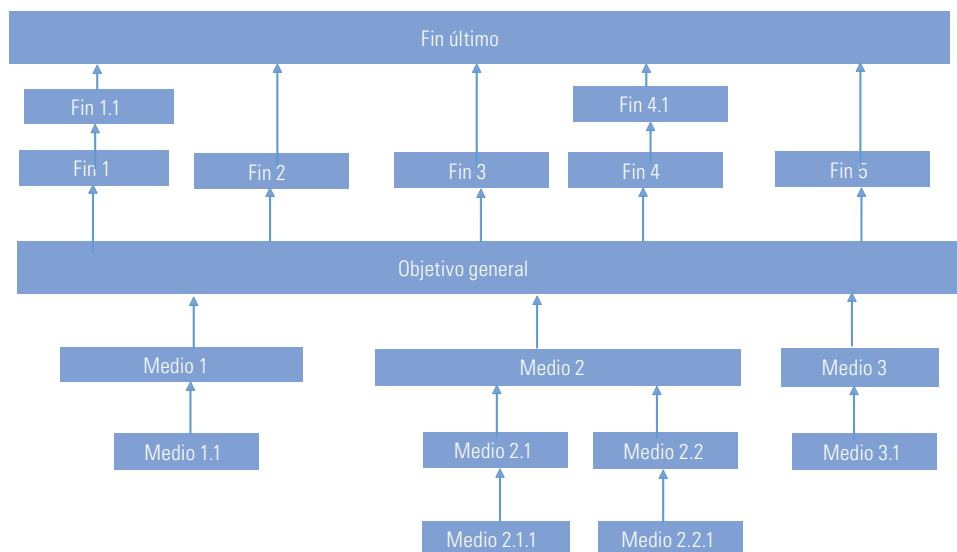
- Paso 1. Determinar el objetivo general. Esto se logra al convertir el problema de desarrollo o la necesidad identificada en una situación positiva.
- Paso 2. Identificar los fines. Es decir, se deben convertir en fines los efectos evidenciados en la identificación de necesidades, de manera tal que se aprecie el estado positivo esperado.

Los fines del objetivo general son las consecuencias positivas que se producen cuando se resuelve el problema identificado.

- Paso 3. Identificar los medios. Consiste en reemplazar las causas que le dan origen al problema o necesidad, por hechos opuestos que contribuyan a solucionarlo.
- Paso 4. Validar los objetivos. Una vez que se hayan construido los objetivos, se examinan las relaciones entre los medios y los fines establecidos para garantizar que el esquema de análisis sea válido e íntegro (véase el diagrama A1.3).

Diagrama A1.3

Ejemplo de esquema para determinar objetivos



Fuente: Instituto Nacional de Estadística y Censos (INEC) de Costa Rica, "Identificación de necesidades del Instituto Nacional de Estadística y Censos de Costa Rica (INEC)", San José, inédito.

- Identificación de alternativas de solución y selección de la más conveniente. En este caso se debe documentar el equipo definido en el proceso de flujograma llamado *Analiza y selecciona alternativas para obtener la información*.
- Acta constitutiva del sistema de registros administrativos económicos. Una vez reunida la información sobre el problema o necesidad y la alternativa elegida para su solución, se dispone de los elementos mínimos requeridos para someterlo a consideración por parte de la autoridad correspondiente (véase el cuadro A1.7).

¹⁰ Tomado de Instituto Nacional de Estadística y Censos (INEC), "Instructivo para la identificación, formulación, ejecución y evaluación de proyectos", San José, inédito.

Cuadro A1.7

Ejemplo de acta constitutiva

Acta constitutiva del sistema de registros administrativos económicos	
Nombre (alternativa seleccionada)	Sistema de registros administrativos económicos
Nombre de la persona que coordina el sistema de registros administrativos económicos	
Objetivo general	
Problema o necesidad de desarrollo	
Justificación o importancia	
Componentes	
Período en que se estima será ejecutado el registro estadístico empresarial	
Costo (de ser posible, se podría indicar una aproximación del costo, con base en proyectos similares)	
Grupos involucrados	
Riesgos	

Presentado por: _____
 Aprobado por: _____

Fecha: _____
 Fecha: _____

Fuente: Instituto Nacional de Estadística y Censos (INEC) de Costa Rica, "Identificación de necesidades del Instituto Nacional de Estadística y Censos de Costa Rica (INEC)", San José, inédito.

Anexo A2

Cuba: experiencia práctica del registro estadístico de negocios

1. Marco legal que ampara la creación del registro estadístico de negocios en Cuba

El Decreto Ley núm. 335 del Sistema de Registros Públicos de la República de Cuba estipula que la Oficina Nacional de Estadística e Información (ONEI) participa en la organización del Sistema de Registros Públicos, en tanto los integrantes de este constituyen fuente de información del Gobierno.

El Decreto Ley núm. 6 del Sistema de Información del Gobierno (SiGob), define el SiGob como el conjunto integrado de los sistemas que se organizan para satisfacer las necesidades informativas relacionadas con los objetivos y planes del Gobierno en todos los niveles organizativos de la sociedad y el ciudadano. También designa a la Oficina Nacional de Estadística e Información (ONEI) como la entidad nacional que tiene la misión de dirigir el SEN y responder por la dirección metodológica del Sistema de Información del Gobierno, incluida su organización, coordinación, integración y control.

El Decreto núm. 9, Reglamento del Decreto Ley núm. 6 del Sistema de Información del Gobierno, entre otras funciones de la Oficina Nacional de Estadística e Información, establece las de definir e implementar, a los fines estadísticos, las metodologías, clasificaciones y nomencladores, y asesorar en su uso; coordinar la integración coherente de los registros públicos y los registros administrativos que constituyen fuente del Sistema de Información del Gobierno (SiGob) e identificar, definir y controlar, según corresponda, a los fines estadísticos, los centros informantes que tributan al SEN.

Lo anteriormente expuesto asegura un marco legal para el aprovechamiento de los registros, la incidencia en su diseño, la implementación de estándares y codificadores definidos por la ONEI, el registro de actores del SiGob y la definición de un identificador único como parte de estos codificadores.

Es principio de ambos sistemas la interoperabilidad, de modo que tanto los registros como los sistemas se comuniquen entre sí, aspecto contemplado en el proceso de informatización de estos.

2. Flujo de creación del registro estadístico de negocios en Cuba

El registro estadístico de negocios en Cuba se denomina Directorio de Unidades Institucionales y Establecimientos (DUIE), que tiene como fuente principal los registros de personas jurídicas del país.

Debido a la baja informatización de los registros de personas jurídicas del país, la ONEI establece un registro directo, para lo que se requiere la presentación en la ONEI de los documentos constitutivos de la unidad legal.

A partir de los documentos de constitución, el registro recoge no solo los elementos de identificación y localización, sino que codifica un conjunto de atributos como la clase de actividad económica según el Clasificador Nacional de Actividades Económicas (CNAE), el sector institucional, la forma de financiamiento, la localización en zonas especiales de desarrollo, los grupos empresariales a que se asocia o subordina, los establecimientos, entre otros.

En el acto de registrar la unidad en el DUIE, se le asigna el número de identificación única otorgado por la ONEI y se emite un documento oficial de este registro a la unidad registrada. Este número de identificación única cumple el principio de que no se repite, aun cuando se haya extinguido la entidad. La inscripción en el registro se realiza de forma literal a partir de los documentos aportados por las entidades económicas. El número de identificación única otorgado a la unidad es de amplio uso en el país, tanto para la apertura de cuentas bancarias, como para el registro en la Oficina Tributaria y la concertación de contratos comerciales y de suministro de servicios.

La coordinación del uso de este código por parte de entidades con el banco y la oficina tributaria como principales aliados, que utilizan el mismo número identificador en sus bases de datos, garantiza que las entidades se registren de manera obligatoria en la ONEI y que se alcance una cobertura cercana al 100% de las personas jurídicas.

El código emitido por la ONEI identifica a actores del SiGob, por lo que, en principio, a través de este código único, puede relacionarse la información estadística con la existente en los registros administrativos del Gobierno.

El código de identificación no brinda información sobre las características de la entidad. Sin embargo, en la base de datos del registro administrativo están definidos los atributos de cada unidad jurídica, que permiten relacionar los diferentes tipos de unidades registradas con la información estadística que se reporta y con otras entidades asociadas, como los grupos empresariales y los establecimientos.

El DUIE inscribe a unidades de gobierno, empresas estatales, empresas de capital mixto y extranjero y sociedades mercantiles de capital cubano, así como cooperativas agropecuarias y no agropecuarias, e incluye algunas asociaciones sin fines de lucro. El DUIE contiene, además de las personas jurídicas, los establecimientos (aún en proceso de identificación), que se actualizan a partir de una encuesta realizada a las personas jurídicas en la localidad donde se establecen las estructuras subordinadas organizadas en estratos. A partir del estudio de cada unidad jurídica se va definiendo la apertura de esta hasta la subordinación en unidad local, por tipo de actividad o establecimiento, según sea el caso.

Un dato importante es la definición de la actividad económica principal indicada en los documentos constitutivos de la unidad jurídica, a partir de los objetos sociales o funciones establecidas en principio para cada unidad legal. Todo esto se contrasta con la información económica aportada en las diferentes operaciones estadísticas que la unidad reporta por el número identificador asignado.

Existe, además del número de identificación único de las unidades legales, el código de identificación tributario o Número de Identificación Tributaria (NIT), utilizado mayormente como identificador en el registro de aduana del país. No obstante, la ONEI ha estado trabajando en la correlación entre ambos códigos a los fines del aprovechamiento del registro de aduana para los análisis estadísticos que se fundamentan en dicho registro.

Anexo A3

Cuba: inventario de registros administrativos con fines estadísticos

Oficina Nacional de Estadística e Información de Cuba
Dirección de Metodología de la Información
Marzo de 2017

Objetivo

Establecer un instrumento que permita tener actualizado el inventario con las características (de la fuente y metadato) de los registros administrativos.

Guía para el inventario de registros administrativos

- 1) Datos generales
 - 1.1 Nombre de la institución: es el nombre de la institución donde se levanta la información del inventario.
 - 1.2 Misión de la institución: corresponde a la razón de ser de la institución, que debe ser redactada como conste en los documentos oficiales.
 - 1.3 Dirección geográfica: se registran las calles principal y secundaria, y el número del inmueble donde se encuentra asentada la institución.
 - 1.4 Persona de contacto del registro: datos de la persona de contacto del registro administrativo, es decir, nombres, apellidos, dirección y unidad administrativa donde labora, número de teléfono y correo electrónico.
- 2) Características del registro administrativo
 - 2.1 Nombre del registro administrativo: se debe registrar el nombre del (de los) registro(s) administrativo(s) que administra la institución.
 - 2.2 Tipo de registro: a fines de la encuesta se clasificarán los registros en población, actividades, empresas e inmuebles.
 - 2.3 Máxima normatividad legal que respalda la creación del registro: corresponde a la máxima normatividad legal que permite el funcionamiento del registro administrativo. Admite solo una respuesta.
 - 2.4 Objetivo del registro administrativo: se debe enunciar el objetivo general que cumple el registro administrativo en la institución. Población objetivo: corresponde a individuos u objetos de los cuales se registran los datos en el registro administrativo, que no corresponde a datos administrativos propios de la institución, sino de las personas usuarias de servicios.
 - 2.5 Cobertura: corresponde a la cobertura geográfica más alta que alcanza el registro administrativo.
 - 2.6 Año de creación: es el año en que se creó oficialmente el registro administrativo.
 - 2.7 Medio de captación: es el medio por el cual la institución capta la información en el registro administrativo, que puede ser mediante formulario físico, electrónico u otro medio. Indicar el medio principal por el cual se capta la información directamente desde el usuario.
 - 2.8 Periodicidad de captación: es la periodicidad con la que se capta (recoge) la información de las personas usuarias (individuos, objetos).
 - 2.9 Formato en que se disponen los datos del registro administrativo: corresponde al programa informático (*software* estadístico) en el que se disponen los datos captados en el registro administrativo.
 - 2.10 Bases de datos centralizados o descentralizados: corresponde a la disponibilidad de una base de datos centralizados o no en un solo repositorio, cuando el registro sea captado en distintos lugares geográficos.

- 2.11 Año de disponibilidad de la base de datos distribuida (BDD): corresponde al año desde el que se dispone de los datos captados del registro administrativo en una base de datos.
- 2.12 Período de corte para el almacenamiento: corresponde a la frecuencia con la que se realiza el corte para el almacenamiento de la información en la BDD.
- 2.13 Identificador único: corresponde a la disponibilidad o no de un identificador único para cada objeto (individuo) en la base de datos del registro administrativo.
- 3) Diccionario de variables
- 3.1 Diccionario de variables, ficha explicativa o algún documento que detalle la estructura de las variables del registro administrativo: corresponde a la disponibilidad o no de un diccionario de variables, ficha técnica explicativa o documento que contenga información sobre las características de los datos (variables) que capta el registro administrativo (véanse los cuadros A3.1 y A3.2).

Cuadro A3.1

Formulario para el levantamiento del inventario de registros administrativos

1. Datos generales	
1.1. Nombre de la institución	
1.2. Misión de la institución	
1.3. Dirección geográfica	
1.4. Persona de contacto del registro	Nombres: Apellidos: Dirección administrativa: Unidad donde labora: Teléfono: Ext.: Correo electrónico:
2. Características del registro administrativo	
2.1. Nombre del registro administrativo	
2.2. Tipo de registro	
2.3. ¿Cuál es la máxima normatividad legal que respalda la creación del registro administrativo?	Constitución política <input type="checkbox"/> 1 Tratado <input type="checkbox"/> 2 Ley internacional <input type="checkbox"/> 3 Decreto <input type="checkbox"/> 4 Resolución <input type="checkbox"/> 5 Otra <input type="checkbox"/> 6 ¿Cuál? Ninguna <input type="checkbox"/> 7
2.4. ¿Cuál es el objetivo general del registro administrativo?	
2.5. ¿Cuál es la población objetivo que se registra?	
2.6. ¿Cuál es la cobertura del registro administrativo?	Nacional <input type="checkbox"/> Provincial <input type="checkbox"/> Otro <input type="checkbox"/> ¿Cuál?
2.7. Año de creación del registro	
2.8. ¿Cuál es el medio de captación?	Formulario físico <input type="checkbox"/> 1 Adjuntar formulario Formulario electrónico <input type="checkbox"/> 2 Adjuntar formulario Otro <input type="checkbox"/> 3 ¿Cuál?
2.9. ¿Cuál es la periodicidad de captación?	Anual <input type="checkbox"/> 1 Semestral <input type="checkbox"/> 2 Cuatrimestral <input type="checkbox"/> 3 Trimestral <input type="checkbox"/> 4 Mensual <input type="checkbox"/> 5 Semanal <input type="checkbox"/> 6 Diario <input type="checkbox"/> 7 Otro <input type="checkbox"/> 8 ¿Cuál?

Cuadro A3.1 (conclusión)

2. Características del registro administrativo		
2.10. ¿En qué formato se disponen los datos del registro administrativo?	SPSS	<input type="checkbox"/> 1
	Excel	<input type="checkbox"/> 2
	Stata	<input type="checkbox"/> 3
	REDATAM	<input type="checkbox"/> 4
	Otra	<input type="checkbox"/> 5
	¿Cuál?	
	Ninguna	<input type="checkbox"/> 6
2.11. ¿Los datos del registro están centralizados o descentralizados?	Base de datos descentralizado	<input type="checkbox"/> 1
	Base de datos centralizado	<input type="checkbox"/> 2
2.12 Año de disponibilidad de la bdd.		
2.13. ¿Cuál es la periodicidad principal de corte para el almacenamiento de la bdd?	Anual	<input type="checkbox"/> 1
	Semestral	<input type="checkbox"/> 2
	Cuatrimestral	<input type="checkbox"/> 3
	Trimestral	<input type="checkbox"/> 4
	Mensual	<input type="checkbox"/> 5
	Semanal	<input type="checkbox"/> 6
	Diario	<input type="checkbox"/> 7
	Otra	<input type="checkbox"/> 8
	¿Cuál?	
2.14. ¿Dispone la bdd de identificador único para cada objeto (individuo)?	Sí	<input type="checkbox"/> 1
	¿Cuál?	
	No	<input type="checkbox"/> 2
2.15. ¿La institución genera estadísticas con la base de datos del registro administrativo?	Sí	<input type="checkbox"/> 1
	¿Cuáles?	
	No	<input type="checkbox"/> 2
3. Diccionario de variables		
3.1. ¿El registro administrativo cuenta con un diccionario de variables, ficha explicativa o algún otro documento que detalla la estructura de las variables del registro?	Sí	<input type="checkbox"/> 1
	Adjuntar archivos y llene la ficha técnica	
	No	<input type="checkbox"/> 2
	Llene la ficha técnica	

Fuente: Oficina Nacional de Estadística e Información (ONEI) de Cuba.

Cuadro A3.2

Ficha técnica del registro administrativo

Nombre de la variable (escribir de forma completa lo que figura en el formulario)	Definición del campo o variable (en caso de que corresponda, indicar la unidad de medida)	Tipo		Campo			Abierto		Cerrado			Categorías de la variable				
		Númerico	Númerico con decimales	Alfabético	Alfanumérico	NS/NR	Abierto	Cerrado	NC/NR	Codificado	No codificado		NC/NR	Excluyente	Multimarca	NC/NR

Fuente: Oficina Nacional de Estadística e Información (ONEI) de Cuba.

- a) Categorías de la variable según sus funciones
- Variable de identificación (clave)
 - Variable de comunicación (contacto)
 - Variable de referencia (claves exteriores)
 - Referencia en el tiempo
 - Variables técnicas (uso interno)
 - Variables estadísticas (edad, sexo, estado civil)

Anexo A4

Chile: la experiencia del INE

Como parte de la formulación de los proyectos estratégicos, derivada del proceso de generación del Plan Estratégico 2018-2022 del Instituto Nacional de Estadísticas (INE) de Chile (que se enmarca sobre tres ejes: valor público del dato estadístico, estadística nacional y excelencia organizacional), es que la institución ha identificado la necesidad de contar con un sistema que permita integrar los registros administrativos disponibles a nivel nacional y potenciar su uso en la producción estadística. Es así como se decide formular el proyecto estratégico “Sistema integrado para el uso estadístico de los registros administrativos”.

El proyecto tiene por objetivo idear, diseñar, implementar y documentar un sistema eficiente que integre el uso estadístico de los registros administrativos a la producción del INE, de acuerdo con estándares internacionales de calidad y seguridad. Esto no quiere decir que actualmente el INE no desarrolle operaciones estadísticas basadas en registros administrativos, pero se requiere dar un impulso mayor a esta fuente de datos, en línea con los requerimientos de información existentes en el mundo actual.

Para la construcción del sistema de registros estadísticos, el INE debe tener acceso a registros administrativos disponibles en diferentes oficinas administrativas. De esta manera, se deben establecer convenios y proyectos con diferentes ministerios y autoridades del nivel directivo, con el objetivo de promover una constante cooperación entre el INE y otras instituciones. Se deben estudiar formas para que las oficinas administrativas puedan acceder a los registros base y registros estadísticos, desde una perspectiva de estudios y análisis estadístico. Es relevante mencionar que en todo este proceso debe resguardarse la privacidad de los datos.

El mayor volumen de información requerido para la toma de decisiones implica que el trabajo de las oficinas nacionales de estadística se vuelva más costoso, lo que involucra a su vez una carga importante para las unidades informantes.

La carga estadística está directamente relacionada con los diferentes cuestionarios que reciben las unidades informantes. Es en este contexto donde los registros administrativos se transforman en una herramienta central para complementar la producción estadística de las naciones.

El proyecto “Sistema integrado para uso estadístico de los registros administrativos” busca establecer las bases metodológicas para la generación de un sistema de registros estadísticos sostenible en el tiempo. Esto requiere una transición de un sistema estadístico tradicional a un nuevo sistema basado en registros administrativos. Implica nuevos métodos sobre cómo utilizar los datos administrativos, nuevas formas de cooperación institucional y un trabajo a largo plazo para el desarrollo de nuevas metodologías de producción estadística.

Dentro del proyecto se han introducido algunas definiciones preliminares en torno al marco conceptual para el uso estadístico de los registros administrativos, y el modelo de gestión y producción estadística de los registros administrativos. En el modelo de gestión y producción estadística, se presenta un esquema que permite organizar los registros administrativos, desde el envío de estos por parte de una institución externa hasta la puesta a disposición de los datos para el aprovechamiento estadístico de los productores.

En línea con el proceso de construcción se presentan dos experiencias desarrolladas por el INE en el marco del proyecto. Están asociadas a los subprocesos de construcción de instrumentos de recolección y de construcción o mejora de componentes.

1. Construcción de instrumentos de recolección: transferencia de archivos

Para la construcción del sistema de registros estadísticos, el INE debe tener acceso a registros administrativos de diferentes fuentes administrativas. De esta manera, se deben establecer convenios y proyectos con diferentes ministerios y autoridades del nivel directivo, con el objetivo de promover una constante cooperación entre el INE y otras instituciones. Estas últimas deben estar involucradas con las operaciones estadísticas que realiza el INE, de manera que, en caso de existir cambios en los datos de la oficina administrativa, puedan entender el impacto que se generan en las estadísticas.

Toda vez que el INE haya definido un acuerdo de provisión con alguna entidad proveedora de datos que involucre la recepción de registros administrativos, se deben considerar los siguientes canales de intercambio para desarrollar la transferencia de datos:

- i) VPN (red privada virtual) proporcionada por entidades proveedoras de datos, donde el personal específico del INE debe realizar la conexión a los servidores externos para extraer la información y, posteriormente, almacenarla en un directorio compartido (servidor de archivos).
- ii) El INE proporciona una VPN, donde la entidad proveedora de datos es la encargada de transferir la información a una carpeta específica. Las credenciales del INE para los responsables de transferir los datos son entregadas mediante correo certificado.

Estas dos formas de capturar la información permiten establecer un entorno seguro para la transferencia de grandes volúmenes de datos, en un servidor definido para esos propósitos y con personal controlado que puede acceder para la extracción de esos archivos.

Con estas alternativas, el INE puede adaptarse a los requerimientos de la entidad proveedora de datos, o también establecer las condiciones del intercambio de datos. Estas formas de intercambio soportan todo tipo de formatos de archivos (texto sin formato, planillas Excel, entre otros).

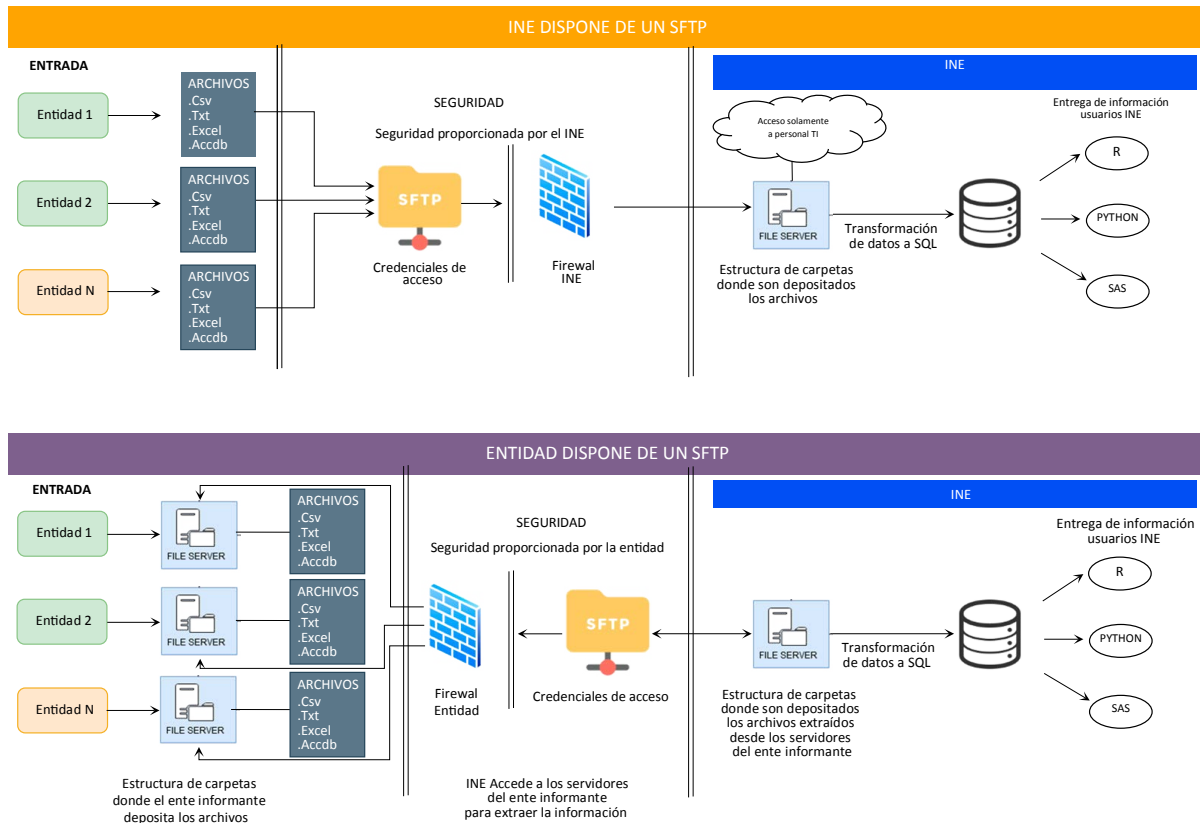
Si bien los archivos pueden traspasarse en un formato estructurado, el foco de las definiciones del INE se relaciona con el intercambio de archivos sin formato, dadas sus ventajas de reducir el volumen de los archivos.

Se debe destacar que, para el resguardo de la confidencialidad de los datos, un equipo definido de la unidad de tecnologías de información puede acceder directamente a la descarga de estos archivos dentro de la institución. Esta información se transformará posteriormente a registro estadístico para la puesta a disposición de los equipos productores. El alcance de esta transformación se explica en la siguiente sección (véase el diagrama A4.1).

Para la recopilación de los antecedentes asociados al canal de intercambio definido, se deben documentar los siguientes ítems: canal de intercambio, nombre del servidor donde se respaldan los datos, ruta del servidor donde se respaldan los datos, contacto de sistemas de la organización y contacto de sistemas por parte del INE. En el cuadro A4.1 se presenta en detalle cada ítem. Con estos antecedentes se facilita la gestión del intercambio de los datos definidos.

Diagrama A4.1

Esquema de transferencia de archivos



Fuente: Instituto Nacional de Estadísticas (INE) de Chile.

Cuadro A4.1

Ítems del canal de intercambio

Ítem	Descripción
Canal de intercambio	Se debe indicar el canal de intercambio por el que se transfiere la información. Estos pueden ser SFTP propio, SFTP externo, medio magnético, entre otros.
Nombre del servidor donde se respaldan los datos	Se indica el nombre del servidor donde se respaldan los datos que provienen de parte de la organización.
Ubicación del servidor donde se respaldan los datos	Ruta de donde se encuentra el respaldo del registro administrativo.
Contacto de sistemas de la organización	Indicar nombre, cargo, correo electrónico y teléfono del responsable de los sistemas de información de la organización asociado al traspaso de información definido.
Responsable de sistemas del INE	Indicar el nombre y cargo del contacto responsable desde los sistemas de información por parte del INE.

Fuente: Instituto Nacional de Estadísticas (INE) de Chile.

2. Construcción o mejora de componentes: almacén de datos para el trabajo con registros administrativos

De lo antes explicado, cuando se reciben archivos de texto sin formato o simple (en formato csv o txt¹¹) en el intercambio de información con entidades proveedoras de datos, se deben extraer los archivos del servidor. Después se desarrollan las actividades de delimitación y transformación de registro administrativo a estadístico¹².

¹¹ También se pueden recibir otros tipos de archivos de texto sin formato.

¹² También pueden ser otros formatos estructurados.

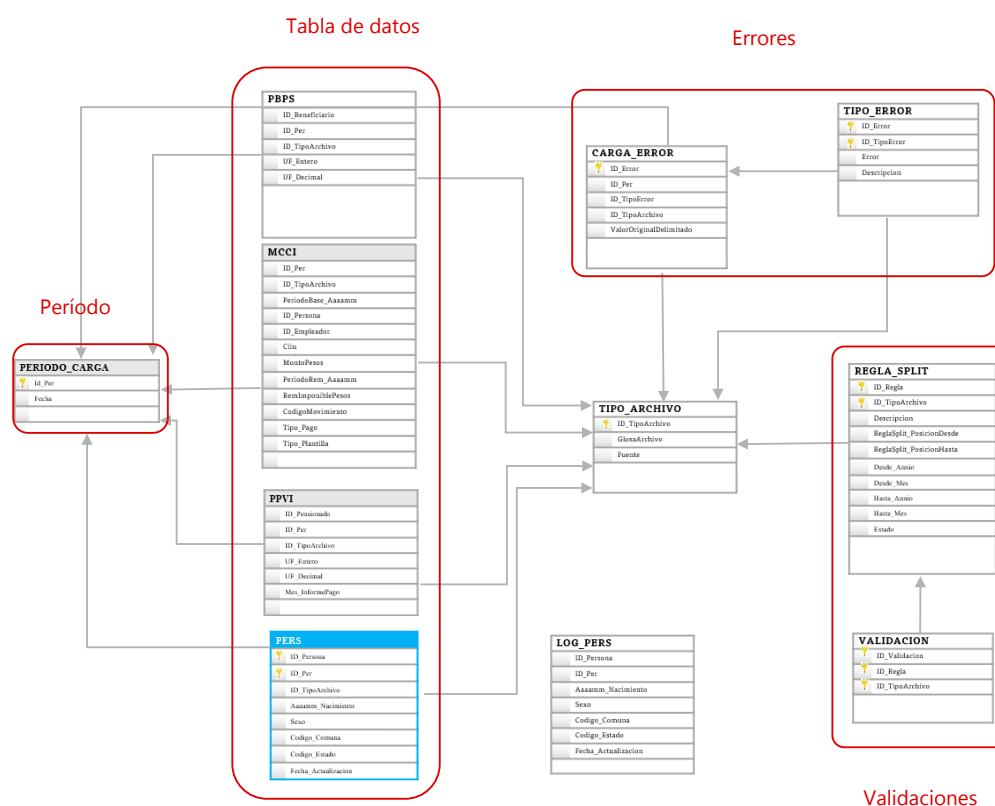
En primer lugar, en cada archivo se debe considerar la separación de las variables según la delimitación y atributos definidos. Una vez que los archivos se delimitan correctamente, se almacenan en un modelo único de bases de datos (almacén de datos) donde se distinguen cuatro grupos: tablas de datos, período, errores y validaciones. Las tablas de datos corresponden a los archivos proporcionados por las entidades proveedoras de datos, el período corresponde al período de carga definido, las validaciones hacen referencia a todas las reglas de negocio aplicada a los archivos, mientras que los errores corresponden al almacenamiento de todas las inconsistencias detectadas a partir de las reglas de negocio definidas.

Esta nueva forma de modelar el trabajo con los registros administrativos obedece a la necesidad de dar respuestas a los requerimientos para trabajar con mayores volúmenes de información y desarrollar un sistema de datos que aproveche las actuales competencias de los equipos informáticos de la institución, donde existe un amplio conocimiento para el trabajo en SQL.

En el diagrama A4.2 se muestra el modelo de almacén de datos aplicado a los datos de la Superintendencia de Pensiones¹³. Con diferentes registros administrativos, cada modelo de trabajo tendrá diferentes nombres de tablas contenidas dentro del recuadro que contiene la tabla de datos, pero este es el modelo de trabajo que se está aplicando a cada registro administrativo para la transformación a registro estadístico.

Diagrama A4.2

Esquema de modelado de almacén de datos con datos de la Superintendencia de Pensiones



Fuente: Instituto Nacional de Estadísticas (INE) de Chile.

Parte de las actividades de delimitación involucran validar el diseño definido. Esto significa que, si en el diseño se contempla que el archivo contenga cinco variables, toda vez que se reciban los archivos se debe mantener esa estructura.

¹³ Corresponde a uno de los registros administrativos que se ha trabajado bajo este modelo, dado su volumen mensual. Se cuenta con datos desde diciembre de 2009, con tablas de datos que rondan los 7 millones de filas mensualmente.

Si existe una estructura de delimitación diferente, se debe contactar a la entidad proveedora de datos e informar sobre las inconsistencias encontradas, y puede ser necesario modificar el diseño y construcción del *script* en etapas previas, para volver a ejecutar la delimitación. Con estos propósitos se establece un protocolo de documentación para cada archivo transferido que contenga la información de los cuadros A4.2 y A4.3.

Cuadro A4.2

Ítems de la estructura de datos para caracterizar archivos (tablas)

Ítem	Descripción
Título del archivo	Se debe indicar el título por el que se conoce el archivo.
Nombre del archivo	Se indica el nombre del archivo.
Formato del archivo	Se indica el formato del archivo (txt, csv, xlsx, entre otros).
Delimitación del archivo	Se debe indicar el tipo de delimitación (ancho fijo, delimitado por carácter, o si ya viene estructurado). En caso de tener un delimitador, se debe indicar cuál es el carácter utilizado.
Número de variables	Se indica el número de variables contenidas en el archivo.
Periodicidad de envío del archivo	Se debe detallar la periodicidad de recepción del archivo por parte del INE. En lo posible, se debe expresar cuál es el período en que se recibe, por ejemplo: último día hábil del mes, final de año, entre otros.
Tamaño del archivo (aproximado)	Se solicita entregar una estimación del archivo en megabytes (MB) con el objetivo de estimar el espacio a utilizar para su respaldo y posterior disposición en los servidores de trabajo.
Descripción del archivo	Se describen los aspectos principales de las características principales de las variables contenidas dentro del archivo.

Fuente: Instituto Nacional de Estadísticas (INE) de Chile.

Cuadro A4.3

Ítems de la estructura de datos para caracterizar el contenido de cada archivo

Ítem	Descripción
Nombre del archivo	Se indica el nombre del archivo.
Información del campo	Indicar descripción del campo o variable, junto con su contenido.
Tamaño del campo	Indicar el tamaño del campo.
Formato del campo	Indicar el formato del campo, si es alfanumérico, numérico, fecha, entre otros.
Notas adicionales del campo	Incluir notas adicionales respecto al contenido del campo, ya sean alcances de la variable, lista de categorías o valores que puede tomar la variable, entre otros.
Posición del campo	Indicar posición de la variable dentro del archivo. Debe ser un número.

Fuente: Instituto Nacional de Estadísticas (INE) de Chile.

Además, dentro de las actividades de transformación del registro administrativo a registro estadístico se realizan normalizaciones y se procede a la codificación y conversión de variables, con el objeto de facilitar la integración de las unidades y la comparación de valores mediante códigos únicos. Dentro de estas actividades se contempla la creación de nuevas variables. También se ejecutan validaciones dentro del modelo, ya sea respecto a intervalos lógicos o de consistencia dentro de las tablas, según se hayan definido dentro del diseño.

Se debe indicar, dentro de las transformaciones a realizar, si se desarrollan actividades para resguardar las identidades de las unidades estadísticas. Todos los registros administrativos que contengan variables identificadoras, donde se especifique que estas serán usadas para crear vínculos (lo que se debe definir para cada variable de cada registro administrativo en el diccionario de datos), deben ser transformadas a un identificador interno, que finalmente corresponde a la variable que permitirá la integración de registros.

Por otra parte, todas aquellas variables que contengan información explícita y donde se puedan identificar directamente las unidades estadísticas, son definidas como variables sensibles y serán encriptadas para mantener el resguardo y su indeterminación. Esto se realiza con el propósito de que dentro del sistema se mantenga bajo secreto estadístico toda la información sensible que identifique a la unidad estadística.

Finalmente, toda vez que haya finalizado la transformación del registro administrativo a registro estadístico, los datos pueden ser explotados por las y los productores del INE. Cada usuario que lo requiera debe solicitar acceso al servidor donde se encuentran los datos. Puede conectarse utilizando algún *software* estadístico, como SAS o R, para el desarrollo de sus respectivos análisis.

Anexo A5

Perú: experiencia del INEI en la fase de procesamiento

El Instituto Nacional de Estadística e Informática (INEI), por conducto de la Dirección Ejecutiva de Censos y Encuestas de Empresas y Establecimiento de la Dirección Nacional de Censos y Encuestas, actualiza cada año el Directorio Central de Empresas y Establecimientos (DCEE).

El DCEE es un sistema de información que tiene como fuente principal de información el Padrón de Contribuyentes de la Superintendencia Nacional de Aduanas y de Administración Tributaria (SUNAT), actualizado también por las investigaciones que realiza periódicamente el INEI y los registros de otras instituciones.

Es importante destacar que el DCEE constituye un instrumento base para la construcción de marcos muestrales con miras a la ejecución de encuestas a empresas y establecimientos en el SEN. Su mantenimiento o actualización es importante para mejorar la calidad de las estadísticas económicas.

1. Usos del directorio

- Se utiliza para la obtención del marco muestral de las encuestas referidas a empresas.
- Permite conocer la demografía empresarial (creación y muerte de empresas).
- Permite analizar la estructura empresarial del país.
- Permite conocer la ubicación de las empresas mediante la georreferenciación.

2. Procesamiento

El procesamiento del DCEE se realiza con periodicidad trimestral y anual. Las actividades que se llevan a cabo para el procesamiento de información se cubren en cinco fases, que se describen a continuación.

2.1. Integración de los datos

En esta fase, se realiza la integración de distintos registros administrativos, en el caso del Perú, el registro base es el Padrón de Contribuyentes que brinda la SUNAT, que es el ente fiscal. La información que brinda la SUNAT es de todas las empresas formadas o constituidas por personas naturales y personas jurídicas. También se incluyen fuentes como encuestas económicas y otros registros administrativos, siendo el Registro Único de Contribuyentes (RUC) la variable o código identificador de la empresa que permite integrar todos los registros o encuestas en el DCEE.

La SUNAT, que es la fuente principal, brinda dos tipos de información: la primera es de periodicidad anual (información del Padrón de Contribuyentes), y la segunda es de periodicidad diaria, que se refiere a las novedades y se denomina “altas y bajas de empresas”.

- a) Recepción y procesamiento de la información (con información diaria)

La información es emitida por el ente tributario de forma diaria. Se almacena en línea mediante SFTP seis tipos de archivos txt. Sin embargo, la actualización se realiza trimestralmente. En la actualización se realizan procesos que incluyen la recepción de la información, la extracción, las depuraciones, las imputaciones y las codificaciones, hasta obtener una base de datos normalizada.

- b) Recepción y procesamiento de la información (con información anual)

La información es emitida por el ente tributario de forma anual gracias a un convenio con esta institución. La información remitida permite actualizar variables como: organización jurídica, CIIU, direcciones, entre otras.

Es importante señalar que la cobertura del DCEE considera a todas las empresas que estén registradas en la SUNAT y que cuenten con un RUC y estén activas o en operación. Se excluyen las empresas formadas como personas naturales sin negocio. Tampoco se considera a las unidades económicas informales.

El sistema de codificación utilizada se basa en las variables que contiene el DCEE, por ejemplo, código de actividad económica CIIU revisión 4, código de ubicación geográfica (UBIGEO), entre otros.

Asimismo, existen criterios internos de validación en la carga de datos, tales como el requisito de que el código RUC cuente con 11 dígitos; de lo contrario se considera inválido el registro. En cuanto a los datos de variables, se evalúa que, por ejemplo, la CIIU cuente solo con cuatro dígitos y el UBIGEO con seis dígitos, que las fechas tengan un formato específico, entre otras cosas, a fin de garantizar la coherencia en la información. No obstante, existen errores o inconsistencia que no cumplen con las reglas y que se deben revisar manualmente antes de integrar la información.

En referencia al análisis de consistencia, se elaboran reportes sobre distribuciones, mínimos y máximos cuando son variables numéricas. Asimismo, cada variable cuenta con un código que se verifica en este proceso. El análisis de consistencia no solo se realiza con datos internos, sino también con fuentes externas que tienen un peso según el tipo de variable. Las de mayor peso son las encuestas económicas para la actualización de CIIU, direcciones, ventas y personal ocupado.

2.2. Clasificación y codificación

El directorio utiliza clasificadores internacionales como la CIIU revisión 4 para codificar las actividades económicas y el código de ubicación geográfica (UBIGEO). Las variables de tipo texto también tienen una codificación a fin de contar con datos o conceptos que permitan la comparabilidad nacional o internacional.

La codificación puede realizarse de forma automática y también manual. Se asigna por la parte informática y en el proceso de validación de la información.

2.3. Revisión y validación

En el proceso de validación se verifica que las empresas cuenten con número de RUC válido. El código debe tener 11 dígitos y debe empezar con el número 1 (uno) o 2 (dos). También se consideran solo las unidades estadísticas activas y se excluye a las personas naturales sin negocio, las sociedades conyugales sin negocio, las sucesiones indivisas sin negocio y las unidades que están en proceso de liquidación. No debe existir duplicación de datos.

2.4. Edición e imputación

La edición en el DCEE se realiza, por ejemplo, cuando la descripción de las variables viene cortada, o está abreviada. En tal sentido lo que se hace es estandarizar y editar.

Por otro lado, en cuanto a las imputaciones, estas se realizan solo para las variables de ventas y personal ocupado de la empresa, ya que no se cuenta con información de todas las unidades económicas. Por tal motivo, se han creado valores mínimos para asignar a estas variables cuando no se disponga de información.

En el DCEE, generalmente las que no cuentan con este dato son las empresas creadas recientemente o empresas muy pequeñas que no declaran impuestos al ente fiscal.

2.5. Derivación de nuevas variables y unidades

Este subproceso se realiza para obtener nuevas variables según los resultados que se requiera difundir. Generalmente son agregaciones de variables, es decir, agrupaciones de variables que facilitan el análisis de la información, o variables derivadas de otras.

a) Variables derivadas por agrupamiento

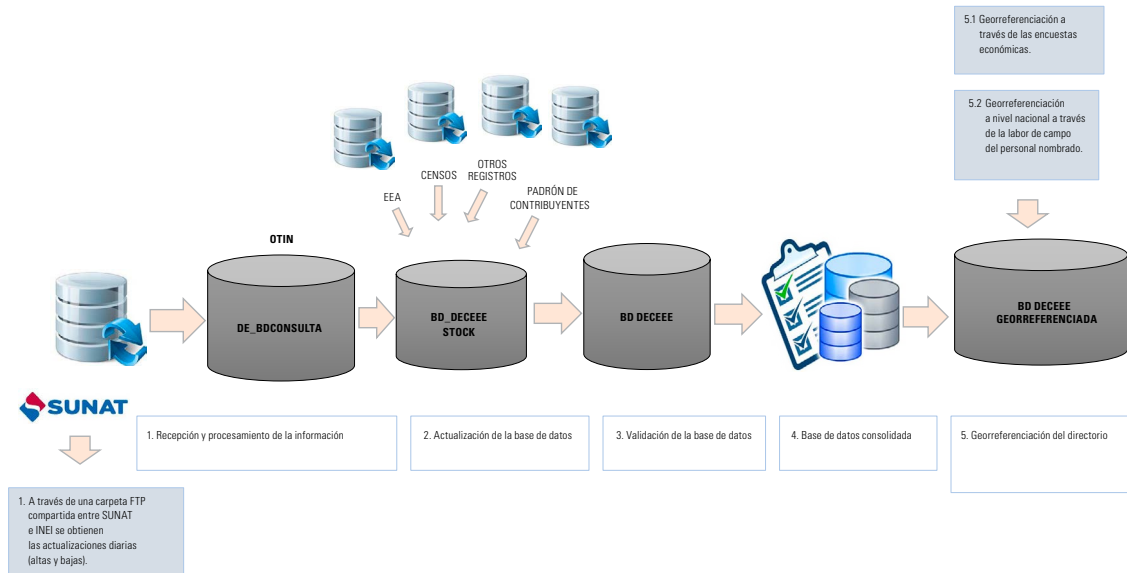
Esta categoría se refiere, por ejemplo, a los rangos de ventas o el rango de personal ocupado. En este caso, las variables numéricas se dividen en intervalos para un mejor análisis.

b) Variables derivadas por agregación

La variable derivada se crea a partir de un registro, usando otro que tiene información más desagregada. Por ejemplo, en el caso del DCEE, se cuenta con la variable CIU y se crean nuevas variables como división, grupo o sección. Incluso pueden crearse variables más agregadas para un mejor análisis (véanse los diagramas A5.1 y A5.2).

Diagrama A5.1

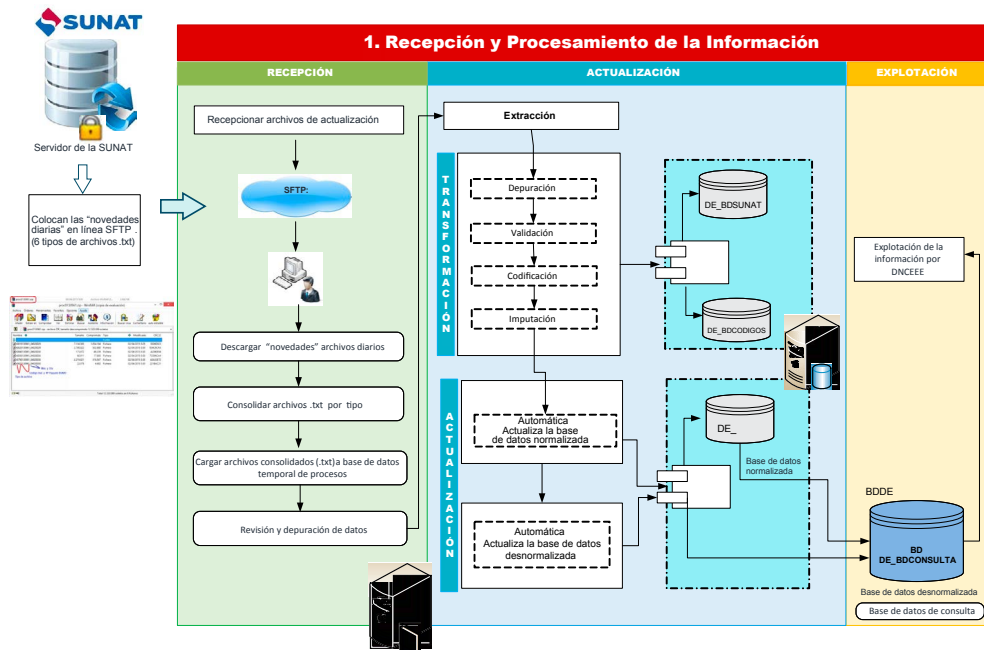
Proceso para la actualización del directorio central de empresas y establecimientos



Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

Diagrama A5.2

Recepción y procesamiento de la información



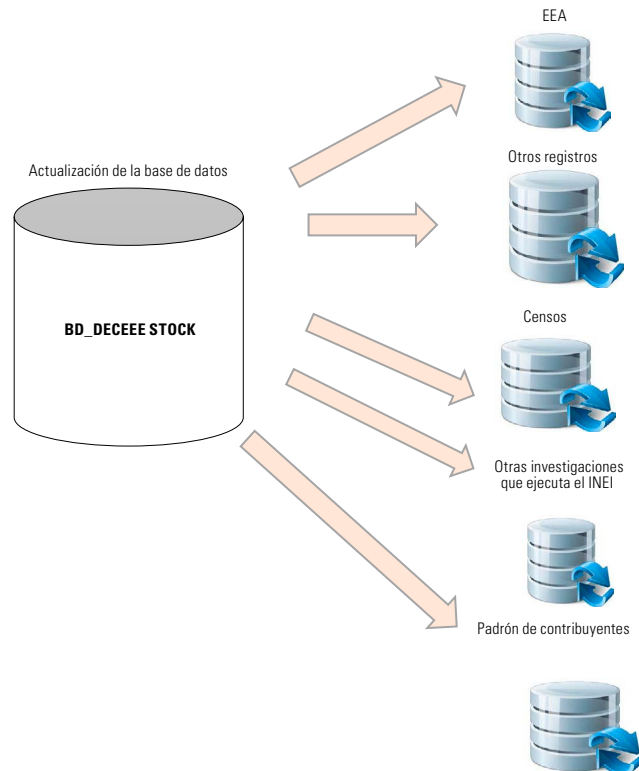
Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

3. Actualización de la base de datos

Luego de la recepción y procesamiento de la información de las novedades, se procede a actualizar la base de datos del DCEE mediante diferentes fuentes de información, tales como encuestas económicas, censos, registros administrativos, otras investigaciones que realiza el INEI y el padrón de contribuyentes que envía la SUNAT cada año. Cabe recalcar que las actualizaciones se realizan cada trimestre (véase el diagrama A5.3).

Diagrama A5.3

Actualización de la base de datos



Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

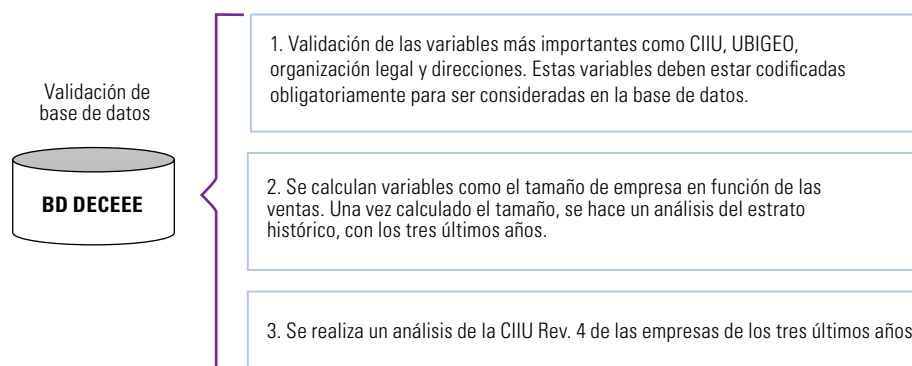
4. Validación de la base de datos

En el proceso de validación se verifica que las empresas cuenten con número de RUC válido. El código debe tener 11 dígitos y debe empezar con el número 1 (uno) o 2 (dos). También se consideran solo las unidades estadísticas activas y se excluye a las personas naturales sin negocio, las sociedades conyugales sin negocio, las sucesiones indivisas sin negocio y las unidades que están en proceso de liquidación.

Luego se validan las variables, tales como CIU, UBIGEO, organización legal y direcciones, entre las más importantes. Sin embargo, la información proporcionada por la SUNAT, que contiene más 8 millones de registros, se almacena en una base de datos que administra el servidor de la OTIN. Del total solo se utilizan, a efectos del DCEE, 2.400 millones de registros (véase el diagrama A5.4).

Diagrama A5.4

Validación de la base de datos



Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

4.1. Validación y actualización mediante las encuestas económicas (EEA, Encuesta Nacional de empleo (ENE), innovación, entre otras)

Este proceso de validación se realiza mediante las muestras seleccionadas que provienen del DCEE. El personal de cada encuesta valida los datos de dirección, CIU, nombre del gerente de la empresa, teléfono, correo electrónico, UBIGEO, entre otras cosas, mediante el teléfono, el correo electrónico o las visitas que realiza el personal de campo durante el período de recogida de información.

5. Base de datos consolidada

Una vez validada la información se obtiene una base de datos consolidada y consistente que se utiliza para la elaboración de publicaciones tales como las de estructura empresarial y demografía empresarial.

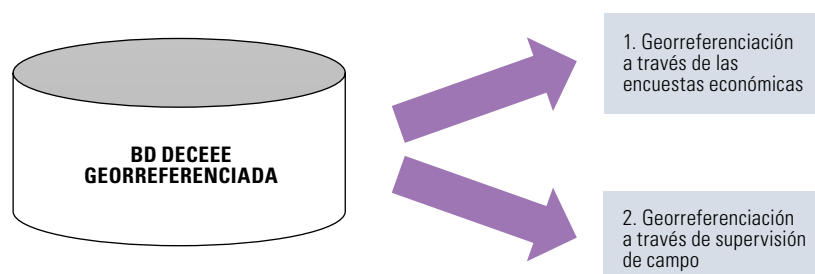
Asimismo, este DCEE actualizado es entregado trimestralmente al equipo que se encarga de realizar la georreferenciación de las empresas.

6. Georreferenciación

La georreferenciación se realiza de dos formas, como se indica en el diagrama A5.5.

Diagrama A5.5

Georreferenciación del directorio



Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

6.1. Georreferenciación mediante las encuestas económicas

a) Mediante la Encuesta Económica Anual

Esta es realizada cada año por el mismo informante, mediante el formulario web. Con este proceso se georreferencian cerca de 10.000 empresas, de las cuales más de 6.000 son grandes empresas. De este modo, gran parte de las grandes empresas están georreferenciadas.

b) Mediante otras encuestas económicas (Encuesta Nacional de Empresas, innovación, entre otras)

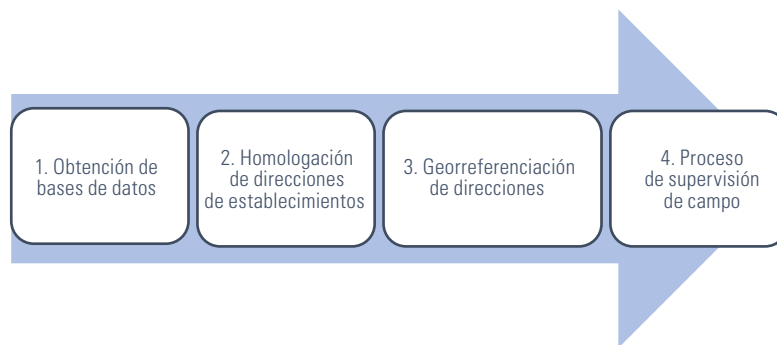
La georreferenciación la realiza el mismo encuestador en campo.

6.2. Georreferenciación mediante supervisiones de campo

La georreferenciación a través de campo se realiza de la manera que se indica en el diagrama A5.6.

Diagrama A5.6

Georreferenciación a través de campo



Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

7. Obtención de bases de datos

7.1. Bases de datos del DCEE

Consiste en la obtención de la base de datos del DCEE, actualizada al último trimestre.

7.2. Datos cartográficos

Se obtiene información cartográfica de las zonas, manzanas y ejes viales y de las direcciones municipales, de otros proyectos que ya hayan sido georreferenciados y cuenten con la base de datos de la Dirección Ejecutiva de Cartografía y Geografía del INEI. También se solicitan y buscan bases cartográficas en formato *shapfile* o *DWG*, de entidades públicas o privadas que sirvan de complemento en la labor de georreferenciación.

7.3. Homologación de las direcciones de los establecimientos

Se extraen del directorio los establecimientos por distrito, utilizando una hoja de cálculo de Excel, donde se detalla la información por empresa como el número de RUC, razón social, tipo de establecimiento, dirección, entre otros tipos de información relevante que ayude a la caracterización de las empresas.

Se inicia el proceso de homologación con el nombre de vías utilizando el eje vial a nivel nacional obtenido de la base cartográfica de la Dirección Ejecutiva de Cartografía y Geografía del INEI, donde se enumeran distintos detalles relacionados con la vía, como la categoría, el nombre de la categoría, el nombre de la vía, el nombre alternativo, el distrito, entre otros tipos de datos, que sirven de ayuda para la estandarización de los nombres de las direcciones de interés obtenidos de la base de datos de la SUNAT, la cual es remitida con múltiples errores o distintas características.

Diagrama A5.7
Información cartográfica

Información recibida		Información homologada	
Descripción	Nombre de la vía	Descripción	Nombre de la vía
Calle	A. Aramburu	Calle	Andres Aramburu

Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

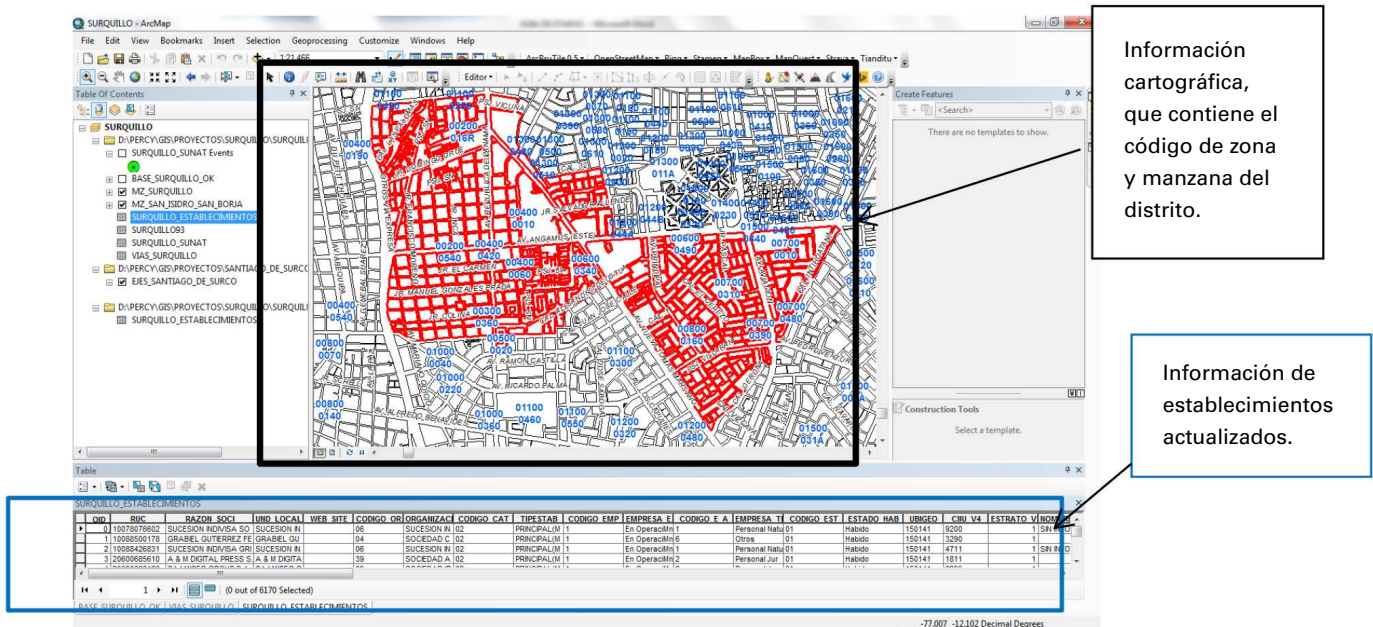
La localización de los establecimientos se realiza a partir de una vía (avenida, calle, jirón, alameda, entre otras) y un número de puerta (número, manzana, lote, departamento, interior, piso, bloque, entre otros), corroborado mediante su RUC con las direcciones publicadas por la SUNAT en su página web.

7.4. Georreferenciación de direcciones

a) Proceso automático

Con información geoespacial proporcionada por la Dirección Ejecutiva de Cartografía y Geografía del INEI, y el archivo Excel de establecimientos actualizados y homologados, se localizan las direcciones de los establecimientos que ya tienen coordenadas.

Imagen A5.1
Información cartográfica con código de zona y manzana

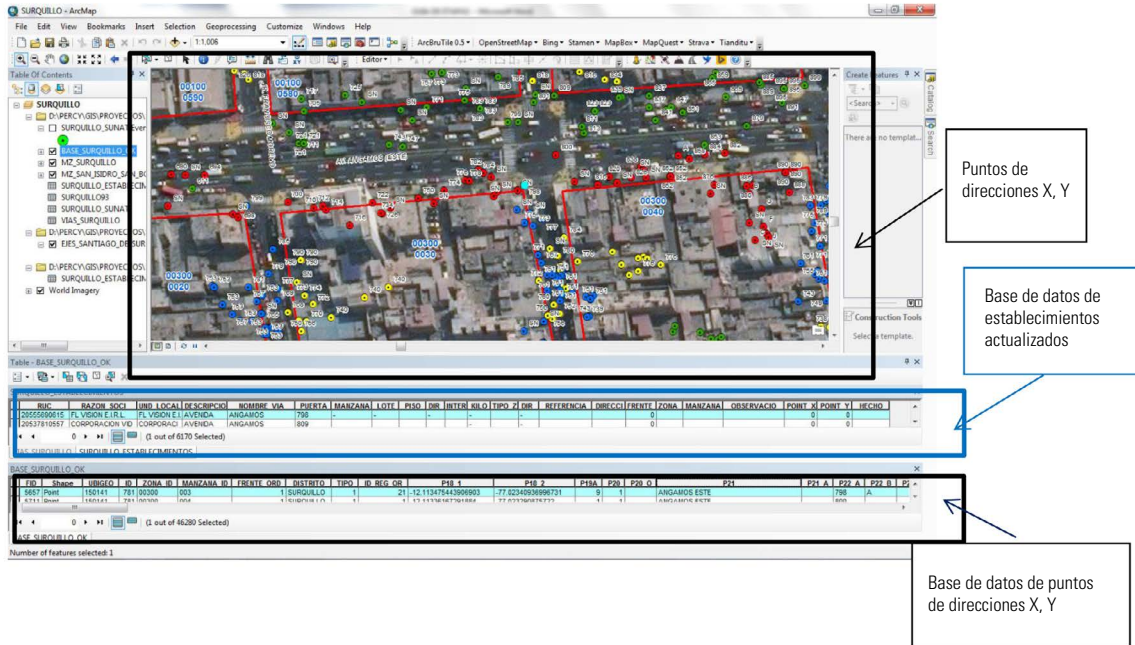


Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

Con la información cartográfica de direcciones (puntos X e Y) proporcionada, de acuerdo con la existencia de información, se realiza una intersección de datos iguales, entre la base de datos de establecimientos actualizados del Directorio Central de Empresas y Establecimientos (DCEE) y la información espacial previamente obtenida de la Dirección Ejecutiva de Cartografía y Geografía del INEI.

Imagen A5.2

Información cartográfica de los establecimientos por coordenadas X e Y

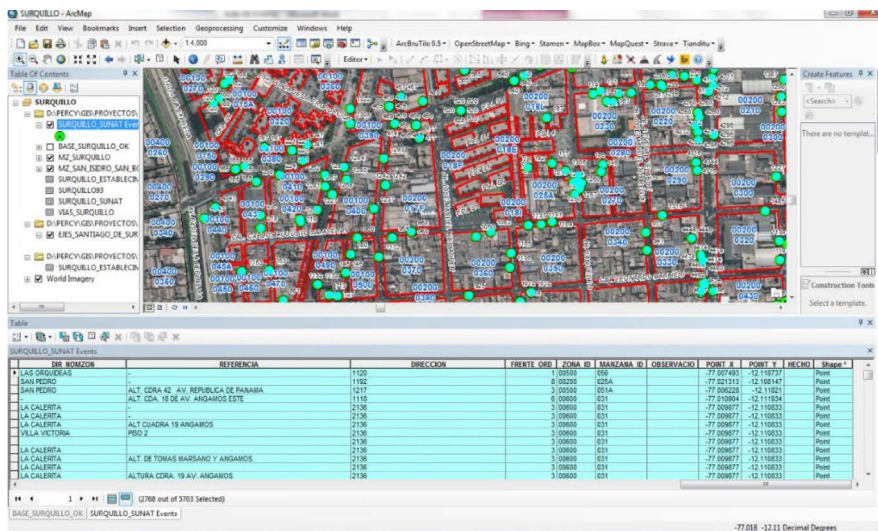


Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

De la intersección de direcciones iguales, se asignan los códigos de frente censal, zona censal, manzana censal y coordenadas X e Y a la base de datos de establecimientos del DCEE.

Imagen A5.3

Información cartográfica por código de frente censal, zona censal, manzana censal y coordenadas X e Y



Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

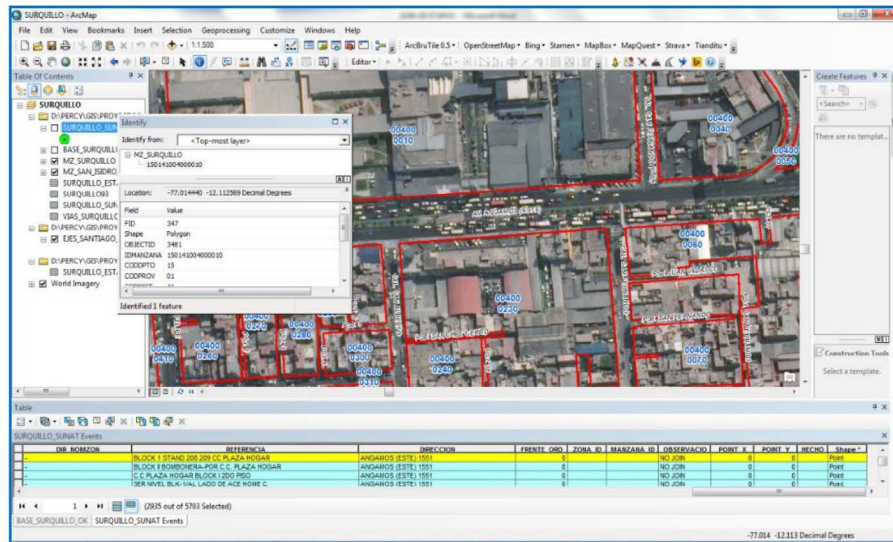
b) Proceso supervisado

Cuando no exista coincidencia entre los campos de direcciones, se completan los códigos de frente censal, zona censal, manzana censal y coordenadas X e Y. Para su localización se emplea la base cartográfica de ciudades del Sistema de Focalización de Hogares (SISFOH), que contiene los códigos de zona y manzana censal. Cuando

se trata de direcciones en que no se precisa el nombre de la calle o la numeración (solo número de manzana, lote o referencia), la ubicación de la dirección se realiza con apoyo de páginas web de direcciones, como Google Maps, Geodir Maps y guías de calles, entre otras.

Imagen A5.4

Información cartográfica de establecimientos con empleo de la base cartográfica de ciudades del Sistema de Focalización de Hogares (SISFOH)

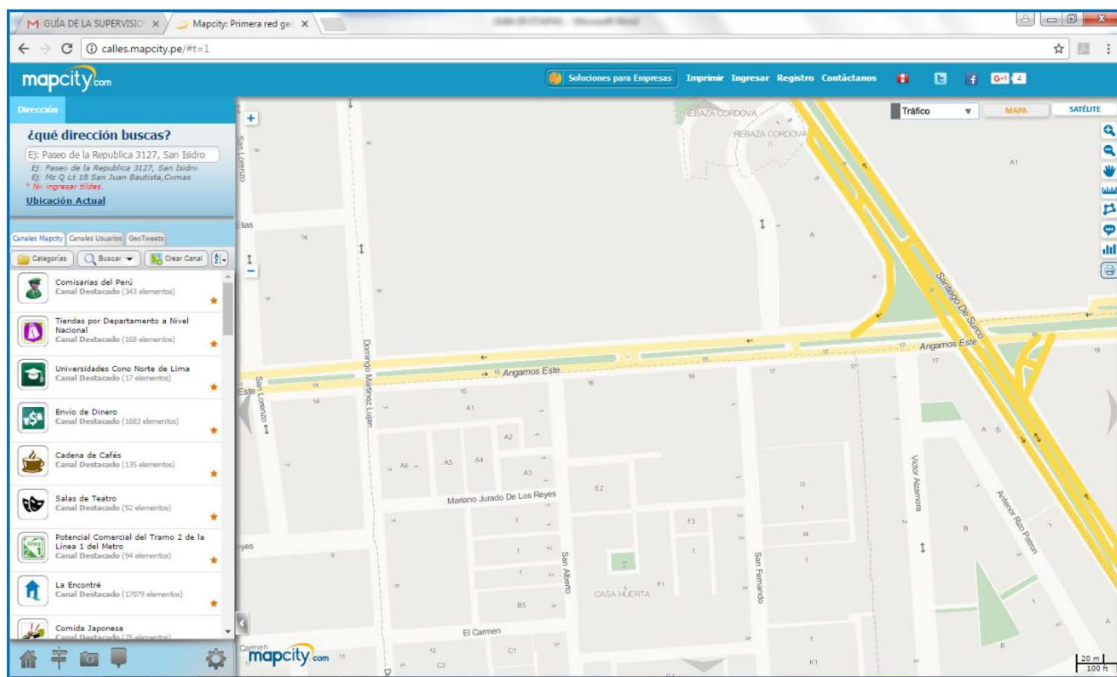


Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

En la imagen A5.5 se muestra el empleo de Mapcity para direcciones con número de lote y manzana.

Imagen A5.5

Empleo de Mapcity para direcciones con número de lote y manzana



Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

7.5. Proceso de supervisión de campo

a) Impresión de cargas de trabajo

Una vez georreferenciado el DCEE, se estandarizan los encabezados, mediante los pasos siguientes:

- Se trasladan los campos de interés de la base de datos georreferenciados según las especificaciones para la supervisión en campo de las empresas.
- Se ubica en las primeras columnas la información de zona censal, manzana censal y el nombre del personal encargado de verificar las direcciones georreferencias del DCEE.
- Se completa la columna de la actividad económica de acuerdo con los códigos de la CIU rev. 4 de cada empresa para su verificación en campo.
- Se detalla un orden de asignación a personal de campo, que obtendrá su carga correspondiente de la zona a verificar.
- Finalmente, mediante el programa Access se generan los PDF de Lima y de oficinas departamentales de estadística e informática (ODEI) para su impresión y traslado físico a cartografía, donde se adjuntan el mapa geográfico y la distribución al personal de verificación de campo de Lima y provincia.

b) Registro de la información brindada por la supervisión en campo

Verificación de la carga, que estará en función del número de empresas a verificar, que en promedio debería ser 25 por cada uno de ellos. Se reciben las cargas, se verifica que hayan llegado completas y que estén debidamente llenadas.

Imagen A5.6

Registro de supervisión de campo

SEDE	DIRECCIÓN_EJECUTIVA	DEPARTAMENTO	PROVINCIA	DISTRITO	VERIFICADOR	N_FICHAS
LIMA	OTAJ	LIMA	LIMA	PUEBLO LIBRE	CARMEN ELISA ALE DUEÑAS	4
LIMA	OTAJ	LIMA	LIMA	PUEBLO LIBRE	JULIA ELENA AYALA CUYA	3
LIMA	DNCE	LIMA	LIMA	PUEBLO LIBRE	FELIX RUIZ FLORES	1
LIMA	DNCE	LIMA	LIMA	PUEBLO LIBRE	HECTOR SAAVEDRA MORA	1
LIMA	DNCE	LIMA	LIMA	PUEBLO LIBRE	MARIA I. VALLADARES ALCANTARA	2
LIMA	DNCE	LIMA	LIMA	PUEBLO LIBRE	LEONEL VILCHEZ SANCHEZ	2
LIMA	DNCE	LIMA	LIMA	PUEBLO LIBRE	SIMEON QUISPE MUÑOZ	4
LIMA	DNCE	LIMA	LIMA	PUEBLO LIBRE	WALTER ROMERO CASTILLO	3
LIMA	DNCE	LIMA	LIMA	PUEBLO LIBRE	SILVIA ANGULO VERA	1
LIMA	DNCE	LIMA	LIMA	PUEBLO LIBRE	VICTOR LEU GUTIERREZ	1
LIMA	DNCE	LIMA	LIMA	PUEBLO LIBRE	EVA AGUADO MOLINA	1
LIMA	DNCE	LIMA	LIMA	PUEBLO LIBRE	LUIS ARIAS MANRIQUE	3
LIMA	DNCE	LIMA	LIMA	PUEBLO LIBRE	MAXIMO ASURZA OLAECHEA	1
LIMA	DNCE	LIMA	LIMA	PUEBLO LIBRE	MARIO BELLIDO CASTILLO	3
LIMA	DNCE	LIMA	LIMA	PUEBLO LIBRE	MARLENY BLANCAS DE LA CRUZ	2
LIMA	DNCE	LIMA	LIMA	PUEBLO LIBRE	JUAN ABAD ROMERO	6
LIMA	DNCE	LIMA	LIMA	PUEBLO LIBRE	CARMEN ASSANTE DIAZ	3

Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

Digitación de la información; el registro se desarrolla sobre la base de un archivo en formato Excel, con el encabezado proporcionado por la persona responsable de actualizar el DCEE. Simultáneamente, se ingresa la información recogida.

Imagen A5.7

Registro de digitación

RUC	RAZON_SOCIO	NOMB_COMER	ACTIV_ECON
20544437063	INMOBILIARIA & CONSTRUCTORA JOSANED S.R.L.		ACTIVIDADES INMOBILIARIAS REALIZADAS A CAMBIO DE U
20557466992	CARDIO-LAC PRESTACIÒN DE SERVICIOS MEDICO DE ALTA COMPLEJIDAD E.I.R.L.		ACTIVIDADES DE MÉDICOS Y ODONTÓLOGOS
20562792881	EMPRESAS & NEGOCIOS CORP EI.R.L.		CONSULTORÍA DE INFORMÁTICA Y DE GESTIÓN DE INSTALA
20600518870	KORY WASI ENGINEERS SAC		CONSTRUCCIÓN DE EDIFICIOS
20600609981	T-REGISTRO S.A.C.		ACTIVIDADES JURÍDICAS
20600757467	CENTRO DE CONCILIACION, ARBITRAJE Y ESTUDIO	CCAES VEREDICTUM	ACTIVIDADES JURÍDICAS
20601256011	CONTROLE S.A.C. - CONTROLE		ACTIVIDADES DE CONSULTORÍA DE GESTIÓN
20509383368	GRUPO MOLINERO VILLA EL SALVADOR S.A.C.		OTRAS ACTIVIDADES DE SERVICIOS PERSONALES N.C.P.
20510879920	CLAUDI SERVICIOS E INVERSIONES GENERALES S.A.C.		OTRAS ACTIVIDADES DE SERVICIOS PERSONALES N.C.P.
20535642321	365 CONSULTORIA Y COBRANZAS S.A.C. - 365 COBRANZAS S.A.C		ACTIVIDADES DE CONSULTORÍA DE GESTIÓN
20477894861	CONSORCIO Y SUMINISTROS LOGISTICOS NERALM	CONSULNER SAC	OTRAS ACTIVIDADES DE SERVICIOS PERSONALES N.C.P.
20508698659	AMFARCO SOCIEDAD ANONIMA CERRADA		ACTIVIDADES DE CONSULTORÍA DE GESTIÓN
20512503897	CORPORACION CIRILO ANTONIO MAURO ESPERIL	COPORACION CAMERV SAC	VENTAS DE PARTES, PIEZAS Y ACCESORIOS PARA VEHÍCULO
20478197331	AUTO SANCHEZ ASESORES SOCIEDAD ANONIMA CERRADA - AUTOSAN S.A.C.		OTRAS ACTIVIDADES DE SERVICIOS PERSONALES N.C.P.
20392969773	TALLER TORETTO SOCIEDAD ANONIMA CERRADA		MANTENIMIENTO Y REPARACIÓN DE VEHÍCULOS AUTOMO
20532366268	IMPORTACIONES EXPORTACIONES COMERCIAL JACKELYN EMPRESA INDIVIDUAL DE RESP		VENTA AL POR MAYOR NO ESPECIALIZADA
20545492963	IMPORTACIONES LJM SOCIEDAD ANONIMA CERRADA - IMPORTACIONES LJM S.A.C.		VENTA AL POR MENOR DE PRENDAS DE VESTIR, CALZADO Y
20545922305	GESTOCAR SOCIEDAD ANONIMA CERRADA-GESTOCAR S.A.C.		OTRAS ACTIVIDADES DE SERVICIOS DE APOYO A LAS EMPRE
20552670494	GER MEDICAL ESTUDIOS Y SERVICIOS MEDICOS ESPECIALIZADOS E.I.R.L.		ACTIVIDADES DE MÉDICOS Y ODONTÓLOGOS
20254302041	ASOC PARA EL DESARRO INTEG DE LA COMUNID	ADINCO	ACTIVIDADES DE OTRAS ASOCIACIONES N.C.P.

Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

Asimismo, se asignan códigos a las variables “estado de la empresa” y “resultado del registro”. Ambas deben estar enlazadas y tener relación entre sí.

Imagen A5.8

Resultado de estado y visita

Estado		Resultado de Visita	
Ubicada.....1	Otra Empresa.....5	Dirección Verificada.....1	
No Ubicada	Vivienda	Ubicación Censal No	
Especifique.....2	Particular.....6	Corresponde.....2	
Ausente.....3	Rechazo.....7	Dirección Inexacta.....3	
Desocupado.....4	Liquidación.....8	Otros	
	Otros.....9	Especifique.....4	

Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

- Homologación de la base de datos; se procede a hacer esta etapa para las direcciones actualizadas con los siguientes criterios: “cod_tipo”, “nombre_vía”, “puerta”, “cuadra”, “piso”, “interior”, “manzana”, “lote”, “kilometro”, “tipo_zona”, “nomb_zona” y “referencia”.
- Control de calidad; se revisa que la base de datos terminada tenga todas las variables y que la información sea consistente.

Posteriormente, se realiza un control de calidad de la información obtenida por la supervisión en campo, tanto para Lima como para las oficinas departamentales de estadística e informática (ODEI). Sobre todo, se comprueban aquellas fichas que no estén completas con la información brindada o no hayan sido llenadas correctamente. Esto se realiza mediante llamadas a los verificadores a sus anexos, o también mediante búsquedas en la SUNAT.

Finalmente, a la información verificada se le asignan las coordenadas X e Y para su posterior incorporación al servidor de la base de datos del DCEE.

7.6. Mejoras realizadas en la verificación y actualización del Directorio Central de Empresas y Establecimientos

- Actualización y verificación del DCEE obligatoria del último trimestre del año, para cada base de datos obtenida previa a la labor de homologación de direcciones a realizarse por cada distrito y ODEI a trabajarse.
- Homologación de direcciones de la base cartográfica de censos de establecimientos actualizada al 2016, obtenida de la Dirección Ejecutiva de Cartografía y Geografía del INEI, con la base SISFOH de ejes viales. Ello permite un mayor número de intersecciones de datos iguales con las direcciones de los establecimientos que ya tienen coordenadas.
- Página web del Organismo de Formalización de la Propiedad Informal (COFOPRI) (GEO LLAQTA): utilización de la plataforma geográfica virtual que sirve para la georreferenciación de direcciones mediante el proceso supervisado. Este permite la ubicación a nivel nacional de aquellas direcciones más complicadas de localizar, que solo muestran información como: manzana, lote, nombre de centro poblado, asentamiento humano, entre otros.

8. Software utilizado en la actualización

Las aplicaciones de *software* utilizadas en el procesamiento, actualización y georreferenciación son las que se presentan en el cuadro A5.1.

Cuadro A5.1

Descripción de los *softwares* utilizados en la actualización

Software	Usos
SQL Server	Se utiliza para analizar, revisar, fusionar y actualizar la base de datos del DCEE con los insumos proporcionados por las investigaciones estadísticas y el Padrón de Contribuyentes de la SUNAT. También se usan para generar tabulados y reportes solicitados por la Dirección Ejecutiva.
SPSS	Se utiliza para procesar, analizar, revisar y generar los tabulados para las publicaciones trimestrales, anuales y pedidos de Infoinei de instituciones públicas, privadas y del público en general, referidos a empresas del DCEE.
Access	Se realizan plantillas de impresión para la generación de listados de empresas para la supervisión en campo.
Excel	Se utiliza para la revisión y análisis de información del DCEE.
ArcGIS	Principal <i>software</i> utilizado para la actualización y georreferenciación del directorio de empresas. Se realizan actividades como: migración de la base de datos y otros archivos vectoriales, verificación y análisis de cobertura geográfica, cruce de base de datos, vectorización geolocalizada de empresas, elaboración de mapas temáticos, entre otras labores.
AutoCAD	Se utiliza para visualizar los archivos en formato DWG, tales como planos municipales disponibles y otros, que se complementan con la labor de georreferenciación y actualización del DCEE.
QGIS	<i>Software</i> complementario que ayuda en la labor de la actualización y georreferenciación.

Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

9. Determinación de variables

9.1. Determinación del tamaño de empresa

Para la determinación del tamaño de las empresas, se toma el rango de ventas que proporciona la SUNAT con base en un convenio de cooperación. En otros casos, se obtienen las ventas puntuales por las investigaciones estadísticas que realiza el INEI.

Otra variable que sirve para determinar el tamaño de la empresa es el rango de personal ocupado.

Cabe recalcar que también se realizan algunas imputaciones principalmente en microempresas, dado que no se registran o declaran ventas por regímenes especiales con el ente tributario (véase el cuadro A5.2).

Cuadro A5.2

Determinación del segmento empresarial

Segmento empresarial	Unidades impositivas tributarias (UIT)
Microempresa	Hasta 150 UIT (hasta 630 000)
Pequeña empresa	Más de 150 y hasta 1 700 UIT (más de 630 000 hasta 7 140 000)
Mediana empresa	Más de 1 700 y hasta 2 300 UIT (más de 7 140 000 hasta 9 660 000)
Gran empresa	Más de 2300 UIT (más de 9 660 000)

Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

Nota: UIT (2019) = 4.200 (soles).

9.2. Determinación del ámbito geográfico

En la determinación de los ámbitos político-administrativos y geográficos se utiliza el código de ubicación geográfica (UBIGEO) aprobado por resolución jefatural núm. 149-2001-INEI, que establece la obligatoriedad para los órganos que integran el SEN.

9.3. Determinación de la actividad económica

La determinación de la actividad económica se basa en la cuarta revisión de la Clasificación Industrial Internacional Uniforme de Todas las Actividades Económicas (CIIU), que fue adoptada en el Perú según resolución jefatural núm. 024-2010-INEI, en enero de 2010. A efectos de presentación se agrupan en grandes categorías de actividad económica.

Cabe recalcar que el registro de la SUNAT usa la CIIU rev. 3 para clasificar las actividades económicas y el DCEE utiliza la CIIU rev. 4. Para realizar el cambio de CIIU se usó una tabla de correspondencia forzosa elaborada por el INEI (véase el cuadro A5.3). Desde julio de 2015, la SUNAT viene implementando gradualmente la adopción de la CIIU rev. 4.

Cuadro A5.3

Determinación de la organización jurídica

Organización jurídica
Persona natural
Sociedad anónima
Sociedad civil
Sociedad comercial de responsabilidad limitada
Empresa individual de responsabilidad limitada
Asociaciones
Otros ^a

Fuente: Instituto Nacional de Estadística e Informática (INEI) del Perú.

^a Incluye sociedad en comandita simple, sociedad en comandita por acciones, sociedad colectiva, fundaciones, cooperativas y otros.

Anexo A6

Brasil: la experiencia de la Coordinación de Servicios y Comercio con las encuestas estructurales por empresa

El MGPIE fue adoptado como referencia metodológica en el ámbito del proyecto de modernización de la producción estadística económica en el Instituto Brasileño de Geografía y Estadística (IBGE). Entre 2018 y 2019, la Coordinación de Servicios y Comercio implementó este modelo en las encuestas estructurales por empresa:

- Encuesta Anual de Comercio;
- Encuesta Anual de Servicios;
- Encuesta Industrial Anual - Producto;
- Encuesta Industrial Anual - Empresa, y
- Encuesta Anual de la Industria de la Construcción.

El proceso contó con gerentes y técnicos de las encuestas utilizadas en las etapas y subprocesos con el fin de analizar y describir los procesos responsables de la producción de estadísticas oficiales.

En este proceso se enumeran todas las actividades asociadas a la construcción y a la difusión de los productos finales mediante diversos canales. En el IBGE los principales canales son:

- Sitio web: portal www.ibge.gov.br¹⁴, en el cual la sociedad puede encontrar la descripción, metadatos y resultados finales de todas las encuestas del Instituto. Las publicaciones con los análisis de los resultados pueden ser consultadas a través de la biblioteca en línea del Instituto (publicación oficial y notas técnicas) o adquiridas en la tienda virtual del portal.

¹⁴ Véase [en línea] <https://www.ibge.gov.br/>.

- SIDRA (Sistema IBGE de Recuperación Automática): el SIDRA constituye el depósito central de tablas estadísticas del IBGE, que posibilita la reproducción y personalización de tablas de resultados de las encuestas a lo largo de la serie histórica.
- Canales de atención: atención telefónica o por correo electrónico para dudas sobre encuestas o publicaciones de IBGE, así como para la solicitud de tablas de resultados personalizadas. Cada requerimiento es protocolado y encaminado para el equipo de diseminación, que evalúa la viabilidad y emite una resolución final. Por este canal también ingresan los pedidos concentrados en la plataforma Fala.BR (Presidencia de la República del Brasil, 2011)¹⁵. El más frecuente es el pedido de informaciones adicionales sobre las encuestas, en conformidad con la Ley de Acceso a la Información. Finalmente, los pedidos de acceso a los microdatos para uso en la Sala de Acceso Restringido también son registrados en los canales de atención del IBGE para la evaluación del equipo de diseminación.

En el caso de los datos estadísticos producidos regularmente, esta fase ocurre en cada iteración. Al haber una periodicidad anual, los productos estadísticos de las Encuestas Estructurales por Empresa se producen regularmente (anualmente). O sea, la etapa “Difusión” comienza una vez por año para cada una de las cinco encuestas de la Coordinación de Servicios y Comercio.

Para la actualización de sistemas de salida, el IBGE busca utilizar las mejores prácticas en la producción de estadísticas y adecuar sus productos a estándares internacionalmente reconocidos. Siguiendo la tendencia internacional al uso creciente de sistemas de metadatos en el papel de orientadores de la producción estadística, la implementación de un sistema de metadatos de la institución ha recibido especial atención. Se desarrolló un sistema único y centralizado para almacenar y gestionar de manera integrada todos los metadatos estadísticos. Abarca, más allá de los conceptos y metodología, la descripción de los procesos de trabajo, así como las informaciones sobre la calidad de la producción y del producto.

Con ese sistema se espera mejorar la calidad de la documentación estadística suministrada a las personas usuarias de los datos, por medio de una base estructurada para organizar y divulgar informaciones sobre el contenido y la estructura de la información estadística producida, elaborados con base en estándares y modelos internacionales y reflejados en las mejores prácticas de modernización de la producción estadística.

El Sistema de Metadatos Estadísticos del IBGE utiliza el estándar de la Iniciativa de Documentación de Datos (DDI) como base para la documentación. Este estándar internacional permite la documentación de todos los aspectos de una operación estadística, tanto de encuestas y censos como de registros administrativos.

La generación de productos de difusión crea los productos necesarios para satisfacer las necesidades de las personas usuarias de cada una de las encuestas. La Coordinación de Servicios y Comercio optó por un modelo flexible que contempla las especificidades de los principales tipos de las personas usuarias de las encuestas estructurales por empresa. Con esto, además de la consulta directa de los resultados en los canales de diseminación del IBGE, se crearon productos en forma de texto que contemplan tanto el análisis genérico de los resultados como el detallado de los resultados específicos.

Por lo tanto, está asociado a este subproceso el conjunto de tareas que tienen como objetivo la preparación de los componentes de este producto (elaboración de las tablas, gráficos y textos explicativos), así como la acción de agregarlos en la forma de productos específicos para cada canal de comunicación y editarlos conforme a las normas del Instituto. Esta etapa también abarca la interlocución con las áreas técnicas de cada encuesta y tiene como objetivo la construcción de los textos analíticos de los resultados.

- Informativo de las encuestas: este producto es elaborado individualmente para cuatro de las cinco encuestas estructurales por empresa, a saber, la Encuesta Anual de Comercio, la Encuesta Anual de Servicios, la Encuesta Industrial Anual - Empresa y la Encuesta Anual de la Industria de la Construcción, y contienen entre cuatro y ocho páginas. Este producto trae algunos análisis agregados por sección del CNAE 2.0, además de

¹⁵ En esta dirección el ciudadano brasileño puede registrar un pedido de acceso a información, denuncia, elogio, reclamación, sugerencia o solicitud de servicios del gobierno federal, entre los que figuran aquellos conectados a la producción estadística de IBGE. Véase Instituto Brasileño de Geografía y Estadística (IBGE) [en línea] : <https://sistema.ouvidorias.gov.br/publico/Manifestacao/SelecionarTipoManifestacao.aspx>; “Sala de Acceso a Datos Restringidos” [en línea] <https://www.ibge.gov.br/acesso-informacao/sala-de-acesso-a-dados-restringidos.html>.

los principales destaques en el nivel de la división del CNAE 2.0 (o agrupamientos de actividades). Debido a la restricción de espacio, los análisis son genéricos y sus destinatarios son las personas usuarias interesadas en una visión sistematizada de los principales cambios estructurales verificados en algunos de los sectores que engloba cada encuesta, a partir de la comparación de los resultados a lo largo de una década. El texto es editado y diagramado por diseñadores y permite la inserción de infografías para ilustrar los resultados. Al final de cada lanzamiento de la encuesta, este informativo se envía por correo electrónico a la lista de contactos de las empresas informantes de aquel ciclo (boletín informativo o *newsletter*).

- **Destaques y aclaraciones:** este producto es elaborado individualmente para todas las encuestas estructurales por empresa y posee cerca de 20 páginas. Cada texto trae los principales resultados coyunturales con el fin de situar al lector y aclarar el escenario macroeconómico que permeó los resultados de la encuesta, los principales resultados del ciclo de divulgación de la encuesta y los años más recientes, los principales cambios estructurales en diez años y los principales cambios estructurales regionales en diez años. El texto es producido por el propio equipo de diseminación de la encuesta y posee un formato más libre.
- **Tablas de apoyo para producción de lanzamientos:** Este producto, elaborado en cada una de las encuestas, constituye un conjunto de resultados personalizados para cada una de las 27 unidades estatales del IBGE, correspondientes a las 27 unidades de la federación del país. Su objetivo es dar subsidios para que sus técnicos puedan escribir los lanzamientos y componer el material local de apoyo a la diseminación.
- **Presentación de resultados:** este producto sintetiza los resultados de la encuesta en formato de diapositivas.
- **La gestión de la diseminación es centralizada en un equipo específico.** Este equipo es responsable de garantizar que todos los elementos para la divulgación estén correctos y realizar la articulación con las demás áreas del Instituto que participan en las preparaciones y divulgación de los resultados. De esta forma, el equipo organiza las secciones de información con periodistas responsables de la comunicación institucional del IBGE. Los acuerdos de embargo junto a la prensa y el prelanzamiento de las encuestas también están insertos en este ítem.
- **La gestión de soporte a las personas usuarias ocurre durante todos los meses del año y comprende la atención directa a las personas usuarias de las encuestas estructurales por empresa.** Las personas usuarias representan una gama diversificada que va desde las agencias públicas en diversas esferas del Gobierno hasta investigadores académicos. Los pedidos de acceso a datos confidenciales son monitoreados por técnicos que son responsables de elaborar resoluciones de viabilidad técnica y articular con otras áreas del Instituto la viabilidad operacional y las condiciones de acceso a los datos. Este subproceso garantiza que las consultas, pedidos de tabulaciones especiales y acceso a microdatos sean registrados y atendidos dentro de los plazos acordados. En promedio, la Coordinación de Servicios y Comercio ejecutó 40 tabulaciones especiales por año en 2019 y 2020. Estos pedidos comprenden tablas que no podrían ser construidas mediante los otros canales oficiales de depósito de datos de la institución debido a las especificidades en la desagregación de las actividades económicas, la cobertura geográfica o el recorte en alguna de las variables de la encuesta. A lo largo del año, estos registros son consultados a fin de proporcionar una revisión en el proceso de diseminación de las encuestas e identificar nuevas necesidades de las personas usuarias.

Anexo A7

Brasil: las iniciativas subnacionales de utilización de registros fiscales y administrativos en la producción estadística

Con la intención de identificar iniciativas de utilización de registros administrativos para fines estadísticos, el IBGE realizó, entre mayo y junio de 2020 y con apoyo y financiamiento del Banco Mundial, el análisis del uso de registros fiscales y administrativos en cuentas regionales en los órganos de estadística o en las Secretarías de Hacienda de las 27 unidades de la federación, buscando esfuerzos de producción estadística relacionados, principalmente, con el uso de los registros fiscales de impuestos sobre el valor agregado (IVA).

De las 27 unidades de la federación, 23 respondieron al análisis del IBGE. De estas, 14 utilizaban registros fiscales de manera estructurada y regular para fines estadísticos o producían algún tipo de información para el público a partir de esos registros. De las nueve restantes que respondieron al análisis y todavía no utilizaban los registros fiscales de manera estructurada, se determinó que dos tenían estudios en curso con esos fines.

La utilización más frecuente identificada (18 unidades de la federación) se refería a las herramientas de consulta en línea del tipo “menor precio”, en las cuales los fiscos locales, a partir de datos extraídos de las facturas de venta al consumidor final, informan al público el menor precio de determinado producto en determinado período. Además de esas herramientas de consulta, también se identificaron otras iniciativas, tales como:

- acompañamiento de la evolución de la actividad de la construcción civil a partir de las facturas de ventas de materiales de construcción;
- acompañamiento del impacto de festividades locales en la economía del turismo de los municipios, a partir de la variación del volumen de documentos fiscales emitidos, y
- producción y disponibilidad de indicadores de salud a partir de las facturas de ventas de medicamentos.

La experiencia del IBGE con el uso de registros fiscales muestra que el mayor énfasis en la utilización de los registros fiscales del tipo IVA se debe, en el caso del Brasil, a los factores siguientes:

- la naturaleza electrónica del registro;
- la estandarización del registro y codificación de las operaciones en el ámbito nacional;
- la tempestividad del registro, que normalmente está disponible pocos días después de la conclusión de la operación, y
- la estandarización de las clasificaciones utilizadas para el registro de las actividades y productos.

En el Brasil, los registros fiscales del tipo IVA relacionados con los bienes utilizan para la codificación de productos la Nomenclatura Común del MERCOSUR (NCM). Se trata de una clasificación de uso regional, inicialmente desarrollada para su uso en el comercio exterior y basada en el Sistema Armonizado (SA). Para la codificación de la actividad, todos los establecimientos y empresas se clasifican según el Clasificador Nacional de Actividades Económicas (CNAE) que, a su vez, se basa en la CIIU rev. 4.

Anexo A8

Argentina: cuestionarios de evaluación

Considerando que se trata de un documento que describe el proceso en términos generales e ideales, y que cada oficina nacional de estadística reviste particularidades, en los recuadros A8.1 a A8.9 se presentan las preguntas orientadoras para cada dimensión expuesta que cada oficina nacional de estadística adaptará a su realidad efectiva, suprimiendo o adicionando preguntas o ítems a esta herramienta.

Recuadro A8.1

Cuestionario referido a convenios o acuerdos

Considerando que el acceso a la fuente de registros administrativos de organismos públicos resulta irregular debido a los cambios de gobierno, así como a los cambios de gestión de un organismo en el marco de un mismo gobierno, la vigencia deberá considerarse en tiempo transcurrido, pero también considerando los cambios de gestión de la entidad proveedora como en periodos de gobierno. Es esperable que estos acuerdos favorezcan la institucionalidad, esto es, que su redacción contemple las necesidades a largo plazo de las estadísticas oficiales, evitando cláusulas que permitan la interrupción de este a criterio de funcionarios de turno.

1. ¿Cuál es la vigencia del convenio? Medido en años y en gestiones de gobierno.
2. ¿Participó el personal técnico de la ONE de la redacción de contenidos?
3. ¿Permite planificar a la ONE en el mediano o largo plazo sobre la disponibilidad de la fuente?
4. ¿Se cumplió con todas las cláusulas acordadas?
5. ¿Contiene cláusulas sobre la plataforma de intercambio de datos?
6. ¿Cuáles no fueron cumplidas? Exprese de ser posible los motivos.
7. ¿Tuvo el convenio modificaciones posteriores? Exprese de ser posible los motivos.
8. ¿Tuvo difusión entre el personal técnico del organismo? ONE y el organismo poseedor de registros administrativos.
9. ¿Significó algún cambio en la dinámica de trabajo en el organismo poseedor de registros administrativos?
10. ¿Qué percepción tiene el personal técnico sobre este tipo de acuerdos? Organismos poseedores de registros administrativos.

Notas de campo del evaluador(a).

Fuente: Instituto Nacional de Estadística y Censos (INDEC) de la Argentina.

Recuadro A8.2

Cuestionario referido a necesidades

1. ¿Qué grado de cumplimiento han tenido los planteos de necesidades anteriores?
2. ¿Están disponibles las variables requeridas?
3. ¿Existe un inventario de operaciones estadísticas en la ONE?
4. ¿Qué alcance geográfico tiene? Nacional y subnacional.
5. ¿Existe un inventario de fuentes en la ONE?

Notas de campo del evaluador(a).

Fuente: Instituto Nacional de Estadística y Censos (INDEC) de la Argentina.

Recuadro A8.3

Cuestionario referido a diseño conceptual y analítico

1. ¿Existe inventario de fuentes de registros administrativos?
2. ¿Cumple con las necesidades planteadas por las personas usuarias? (internas, SEN, externas).
3. ¿El identificador utilizado por la oficina nacional de estadística es el mismo que el utilizado por el organismo poseedor de registros administrativos?
4. De no ser iguales, ¿cómo se resolvió?
5. ¿Fue posible la explotación estadística de esos registros administrativos?
6. ¿Se pudo proyectar la utilización de un único identificador en futuras entregas?
7. ¿Existe armonía conceptual entre el diseño de la ONE y la fuente del organismo poseedor de registros administrativos?
8. ¿Cómo se resolvió en caso de que no existiera?
9. ¿Cómo funcionó el control de calidad de entrada?
10. ¿Significó merma en las necesidades de la ONE?
11. ¿Se pudo avanzar en el proceso de armonización en caso de no existir?
12. Entre los equipos de la ONE y el organismo poseedor de registros administrativos, ¿existe un ambiente de trabajo que permita proyectar en el mediano o largo plazo?
13. ¿Hubo cambios en la frecuencia de producción y difusión?

Notas de campo del evaluador(a).

Fuente: Instituto Nacional de Estadística y Censos (INDEC) de la Argentina.

Recuadro A8.4

Cuestionario referido a la captura de los registros administrativos

1. ¿El almacenamiento elegido soportó el volumen de captura?
2. ¿La captura con entidades proveedoras fue homogénea en la tecnología utilizada o fue diferente?
3. ¿Resulta versátil el diseño de captura de datos, considerando las diferencias entre entidades proveedoras y soportes?
4. ¿Se pudieron integrar estas diferentes captaciones en el sistema de registros administrativos económicos?
5. ¿El inventario de registros administrativos y entidades proveedoras resultó suficiente?
6. ¿Los datos recibidos fueron estructurados o no?
7. En el caso de los no estructurados, ¿cómo se abordó su tratamiento?
8. ¿El diseño y la construcción los contemplaban?
9. ¿La vinculación de bases de datos se dio de acuerdo con lo previsto?
10. ¿Hubo un solo identificador?
11. En caso de que no lo hubiera, ¿cómo se solucionó?
12. ¿Existe capacidad en la ONE para la captación y procesamiento de datos no estructurados?
13. En caso de requerir cambios la plataforma de captura, ¿resultó ágil la adaptación?
14. ¿Resultaron versátiles o rígidos los estándares definidos en el diseño?
15. ¿El módulo de documentación considera las actualizaciones o cambios en la fuente y su impacto en el sistema de registros administrativos económicos?
16. ¿Es posible identificar los cambios o actualizaciones en una línea de tiempo?
17. ¿Cómo se trabajó con las clasificaciones?
18. ¿El registro de los metadatos siguió el ritmo de la captura de datos?
19. ¿El equipo de trabajo resultó suficiente en cantidad y calidad?
20. ¿Fueron necesarias capacitaciones?
21. ¿La incorporación de los registros administrativos significó mayor participación de otras áreas de la ONE?
22. ¿Fue ágil la gestión del sistema de registros administrativos económicos a partir de los diferentes módulos?

Notas de campo del evaluador(a).

Fuente: Instituto Nacional de Estadística y Censos (INDEC) de la Argentina.

Recuadro A8.5

Cuestionario referido a integración de los registros administrativos

1. ¿Cómo fue el proceso de conversión estadística de los registros administrativos según los soportes de origen? Digital, papel.
2. ¿Se pudo avanzar en la informatización de los no digitales?
3. Del análisis conceptual propuesto en la fase de diseño, ¿qué proporción de los registros administrativos resultó utilizable estadísticamente?
4. ¿Mermó la cobertura de las necesidades planteadas?
5. ¿Cómo se trataron los registros administrativos no utilizables?
6. ¿Fue posible avanzar con los organismos poseedores de registros administrativos en la armonización?
7. ¿Qué ítems del análisis conceptual resultaron más conflictivos?
8. ¿Se pudieron realizar correspondencias de clasificaciones?
9. ¿Existen en el área estimaciones de lo no registrado?
10. ¿Fue posible la creación de variables agregadas?
11. ¿Cómo se detectaron los duplicados?
12. ¿Qué porcentaje de duplicados se detectó?
13. ¿Las actualizaciones del sistema de registros administrativos económicos se dan de manera automática y oportuna?
14. ¿Los tiempos de trabajo variaron con relación a los habituales?
15. ¿Fue posible un ensamblado o vinculación de los registros administrativos con datos censales o de encuestas?
16. ¿La cobertura geográfica varió, fue mayor, menor, igual?
17. ¿Las actualizaciones de datos de fuente de registros administrativos permitieron mejoras en la oportunidad de los productos estadísticos?

Notas de campo del evaluador(a).

Fuente: Instituto Nacional de Estadística y Censos (INDEC) de la Argentina.

Recuadro A8.6

Cuestionario referido a confidencialidad

Uno de los limitantes a la hora de disponer de los datos de fuente de registros administrativos tiene que ver con los marcos legales que establecen diferentes tipos de secretos que existen según el fenómeno en cuestión: fiscal, administrativo, médico, estadístico, entre otros.

1. ¿Hubo confidencialidades o secretos en conflicto?
2. ¿Se pudo salvar?
3. ¿Se pudo acceder a los microdatos no anonimizados?
4. ¿Significó una merma en el acceso a la fuente?
5. ¿Se crearon mecanismos para solucionarlo en el futuro?
6. ¿Se mantuvo la confidencialidad?

Notas de campo del evaluador(a).

Fuente: Instituto Nacional de Estadística y Censos (INDEC) de la Argentina.

Recuadro A8.7

Cuestionario referido a costos y beneficios

El costo deberá considerarse desde el punto de vista del costo-beneficio global para el SEN y no solo como una ecuación monetaria unitaria por operación estadística.

El costo-beneficio se entiende como una mayor y mejor oferta estadística al mismo tiempo que las y los informantes de la ONE ven disminuidos o estables sus aportes.

El acceso a fuentes como registros administrativos puede resultar de la disponibilidad de un caudal de datos mayor, tanto cuantitativamente como en términos de una mayor variedad de fenómenos y también de alcances geográficos.

1. ¿Cuántas veces es consultada una unidad económica por las distintas áreas de la ONE?
2. ¿Existe un repositorio central de datos en la ONE?
3. ¿La incorporación de los registros administrativos generó más productos estadísticos?
4. ¿Los registros administrativos pudieron suplir consultas a respondientes?
5. ¿Hubo cambios en la frecuencia de consulta a respondientes?
6. ¿Se detectaron nuevas personas usuarias de la ONE?
7. ¿Existen registros de consultas no satisfechas de las personas usuarias?

Notas de campo del evaluador(a).

Fuente: Instituto Nacional de Estadística y Censos (INDEC) de la Argentina.

Recuadro A8.8

Cuestionario referido a análisis y difusión

1. Considerando la existencia de lo no registrado y el subregistro, ¿cómo impacta en el análisis de los productos generados sobre esta base?
2. Si existen diferencias, ¿se pueden identificar y explicar?
3. ¿Generan cambios en las series?
4. ¿Son significativos?
5. ¿Implican discontinuidades?
6. ¿Exige explicaciones o notas metodológicas adicionales?
7. ¿El equipo de trabajo tiene experiencia en el análisis e interpretación de productos con base en registros administrativos?
8. La actividad económica en registros administrativos impositivos en algunos países es autodeclarativa, a diferencia de las ONE, donde un equipo de expertos la asigna a las unidades económicas. ¿Cómo impacta esta diferencia en los productos estadísticos analizados?
9. ¿Y en las vinculaciones con censos y encuestas?
10. En relación con el diseño, ¿en qué grado se cumplió?
11. ¿Hubo desvíos?
12. ¿Cuáles fueron los motivos?
13. ¿Requirió explicaciones adicionales las publicaciones con base en registros administrativos?
14. ¿Hubo cambios en los metadatos?
15. ¿Tuvo impacto en las personas usuarias?
16. ¿Hubo requerimientos de explicaciones adicionales?
17. ¿Generó nuevas personas usuarias?
18. ¿La promoción de productos con base en registros administrativos significó cambios en la estrategia?
19. ¿Existe una estrategia de promoción con base en la fuente y sus ventajas en general, o solo de productos específicos?
20. ¿Cómo reaccionó la prensa con relación a los productos con fuente en registros administrativos oficiales?
21. ¿Hubo cuestionamientos o requerimientos sobre la confidencialidad?

Notas de campo del evaluador(a).

Fuente: Instituto Nacional de Estadística y Censos (INDEC) de la Argentina.

Recuadro A8.9

Cuestionario referido a asistencia técnica y registración

1. ¿Fue posible avanzar en los organismos poseedores de registros administrativos en los cambios propuestos de los registros administrativos?
2. ¿Se transfirió el conjunto de metadatos de la ONE?
3. ¿Se pudieron realizar capacitaciones, encuentros, reuniones para avanzar en el proceso de armonización conceptual?
4. ¿Es dinámica y colaborativa la relación entre los equipos de la ONE y de los organismos poseedores de registros administrativos?
5. ¿Las consultas y observaciones realizadas desde la ONE se resuelven en tiempo y forma?
6. ¿Se detectan aspectos del convenio que deberían modificarse? Agregar, modificar, excluir, entre otras cosas.

Notas de campo del evaluador(a).

Fuente: Instituto Nacional de Estadística y Censos (INDEC) de la Argentina.

Glosario

Accesibilidad: conjunto de condiciones bajo las cuales las personas usuarias pueden obtener información estadística o geográfica.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario del SNIEG*, Aguascalientes, 2021.

Anidamiento: presentación de una clasificación en cada una de las clases de otra.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Presentación de datos Estadísticos en Cuadros y Gráficas: actualización 2017*, Aguascalientes, 2017.

Aprovechamiento de registros administrativos: método de generación de información estadística mediante el uso de los registros sobre hechos o elementos existentes, que realizan las unidades del Estado como parte de su función.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Norma Técnica para la Elaboración de Metadatos para Proyectos de Generación de Información de Estadística Básica y de los Componentes Estadísticos Derivados de Proyectos Geográficos*, Aguascalientes, 2015.

Bases de datos: conjunto de datos relacionados, estructurados y almacenados sistemáticamente y que pueden ser consultados rápidamente de acuerdo con las características selectivas que se deseen.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Procesamiento de la información*, Aguascalientes, 2011.

Big data o macrodatos o grandes datos: el término inglés *big data* (macrodatos o grandes datos) se utiliza para describir conjuntos de datos de volumen, velocidad y variedad, es decir, las tres V. Las fuentes que se consideran macrodatos a menudo no están estructuradas, lo que significa que no tienen un modelo de datos predefinido o no se ajustan bien a las bases de datos relacionales.
Fuente: Comisión Económica para Europa (CEPE), "In-depth review of big data. Prepared by the temporary Task Team on Big Data and the Secretariat", Ginebra, 2014.

Calidad: el grado en que un conjunto de características inherentes de los procesos y productos cumple con determinados atributos.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Principios y Buenas Prácticas para las Actividades Estadísticas y Geográficas del SNIEG*, Aguascalientes, 2015.

Calidad estadística: cumplimiento de las propiedades que deben tener el proceso y el producto estadístico para satisfacer las necesidades de información de las personas usuarias.
Fuente: Organización de Cooperación y Desarrollo Económicos (OCDE), "Glossary of Statistical Terms" [en línea] <https://stats.oecd.org/glossary/index.htm>; Naciones Unidas, "Glossary", 3 de febrero de 2012 [en línea] <https://unstats.un.org/unsd/dnss/docs-nqaf/NOAF%20GLOSSARY.pdf>.

Categoría: conjunto objeto de cuantificación y caracterización.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Presentación de datos Estadísticos en Cuadros y Gráficas: actualización 2017*, Aguascalientes, 2017.

Celda: intersección de una columna y un renglón.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Presentación de datos Estadísticos en Cuadros y Gráficas: actualización 2017*, Aguascalientes, 2017.

Clase: cada una de las modalidades nominales o intervalos numéricos admitidos por una variable.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Presentación de datos Estadísticos en Cuadros y Gráficas: actualización 2017*, Aguascalientes, 2017.

Clasificación: conjunto de observaciones homogéneas, exhaustivas y mutuamente excluyentes, que pueden ser asignadas a una o más variables que van a ser medidas en la recopilación o la presentación de los datos. Una clasificación estadística es la que se elabora para la recolección y presentación de datos numéricos recolectados sistemáticamente (es decir, estadísticas). Las clasificaciones se dividen en dos:

- i) Clasificaciones adoptadas: son clasificaciones que se toman del referente internacional sin ningún tipo de adecuación al contexto de aplicación particular.
- ii) Clasificaciones adaptadas: son clasificaciones que se adecúan a la estructura económica y social del país, otorgando mayor importancia a los fenómenos que son más relevantes y propios de la realidad nacional.

Fuente: Naciones Unidas, "UN Glossary of Classification Terms" [en línea] https://unstats.un.org/unsd/classifications/bestpractices/glossary_short.pdf.

Cobertura geográfica: territorio al que se refiere la captación de datos en un proyecto estadístico.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario de Estadística Básica*, Aguascalientes, 2012.

Codificación: procedimiento para asignar identificadores numéricos o alfanuméricos a conceptos en un orden establecido.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario de Estadística Básica*, Aguascalientes, 2012.

Coherencia: se refiere al grado en que están lógicamente conectados los conceptos utilizados, las metodologías aplicadas y los resultados producidos por la operación.

Fuente: Organización de Cooperación y Desarrollo Económicos (OCDE), "Glossary of Statistical Terms" [en línea] <https://stats.oecd.org/glossary/index.htm>

Comparabilidad: grado en que son equivalentes las definiciones y clasificaciones de conceptos comunes y de datos referentes a distintas fuentes, momentos o unidades geográficas.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario del SNIEG 2021*, Aguascalientes, 2021.

Confiability: la cercanía del valor inicial estimado al valor estimado posterior.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario del SNIEG 2021*, Aguascalientes, 2021.

Principio de veracidad o calidad de los registros o datos: La información contenida en los bancos de datos debe ser veraz, completa, exacta, actualizada, comprobable y comprensible. Se prohíbe el registro y divulgación de datos parciales, incompletos, fraccionados o que induzcan a error.

Fuente: Congreso de la República de Colombia, *Ley 1266 de 2008*, Bogotá, 2008.

Confidencialidad: atributo de seguridad de la información que indica que la información solo es revelada a individuos o procesos autorizados.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Política para la Seguridad de la Información Estadística y Geográfica del Instituto Nacional de Estadística y Geografía*, Aguascalientes, 2020.

Congruencia: una medida de la calidad de los datos relativa al grado en que estos pueden combinarse en diferentes conjuntos y para diversos propósitos con información estadística de otras fuentes.

Fuente: Elaboración propia, sobre la base de Instituto Nacional de Estadística y Geografía (INEGI), *Proceso estándar para el aprovechamiento de registros administrativos*, Aguascalientes, 2012.

Conjunto de datos: datos organizados. Incluyen metadatos estructurales a través de los cuales estos datos se vinculan entre sí y se agrupan dentro de una misma unidad temática.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Norma Técnica del Proceso de Producción de Información Estadística y Geográfica para el Instituto Nacional de Estadística y Geografía*, Aguascalientes, 2020.

Consistencia: solidez o estabilidad de los datos en tanto muestran coherencia con otros períodos y con la información relacionada.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario de Estadística Básica*, Aguascalientes, 2012.

Criterios de validación: conjunto de reglas de naturaleza conceptual, que sirven de base para la identificación y solución de los problemas que se presentan en los datos estadísticos.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario de Estadística Básica*, Aguascalientes, 2012.

Cruce de variables: combinación de cada una de las clases de una clasificación con cada una de las clases de otra, respecto a las variables involucradas.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Presentación de datos Estadísticos en Cuadros y Gráficas: actualización 2017*, Aguascalientes, 2017.

Cuadro estadístico: ordenamiento matricial de valores cuantitativos y su descripción conceptual.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Presentación de datos Estadísticos en Cuadros y Gráficas: actualización 2017*, Aguascalientes, 2017.

Cuestionario electrónico: documento elaborado por medio de programas en equipos informáticos, con preguntas organizadas de manera secuencial y estructurada a partir de una determinada necesidad de información.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Proceso estándar para encuestas por muestreo*, Aguascalientes, 2010.

Dato: el valor que toma una variable estadística o un objeto espacial geográfico en una unidad de observación.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario del SNIEG*, Aguascalientes, 2021.

Dato administrativo: valor de una variable correspondiente a una acción, hecho o evento que forma parte de un registro administrativo.

Fuente: F. Seguí, "Guía de la herramienta para la evaluación de la calidad de registros administrativos a ser usados con fines estadísticos", Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

Dato espacial: el registro digital con combinaciones de valores de atributos que lo hacen único e inconfundible respecto a otros datos espaciales, y que debe estar georreferenciado y contener un atributo que lo relacione con el tiempo para fines de comparación. Un dato espacial corresponde a una ocurrencia de un objeto espacial.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario del SNIEG 2021*, Aguascalientes, 2021.

Dato estadístico: dato que es recolectado o generado por estadísticas, ya sea en el proceso de las observaciones estadísticas o por procesamiento de datos estadísticos.

Fuente: Oficina Estadística de la Unión Europea (Eurostat), "Eurostat's Concepts and Definitions Database" [en línea] http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=DSP_GLOSSARY_NOM_DTL_VIEW&StrNom=CODED2&StrLanguageCode=EN&IntKey=24031736&RdoSearch=CONTAIN&TxtSearch=DATA&CboTheme=&IsTer=&IntCurrentPage=6&ter_valid=0; Comisión de Estadística/Comisión Económica para Europa (CEPE), *Terminology on Statistical Metadata*, Naciones Unidas, Ginebra, 2000 [en línea] http://ec.europa.eu/eurostat/ramon/coded_files/UNECE_TERMINOLOGY_STAT_METADATA_2000_EN.pdf.

Datos primarios: son los datos obtenidos de las y los informantes del sistema sobre hechos que son relevantes para el conocimiento de los fenómenos económicos, de gobierno, seguridad pública e impartición de justicia, demográficos y sociales, así como sus relaciones con el medio ambiente y el espacio territorial.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Norma Técnica del Proceso de Producción de Información Estadística y Geográfica para el Instituto Nacional de Estadística y Geografía*, Aguascalientes, 2020.

Desagregación clasificatoria: nivel de detalle en el listado de una clasificación.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Presentación de datos Estadísticos en Cuadros y Gráficas: actualización 2017*, Aguascalientes, 2017.

Disponibilidad: atributo de seguridad de la información que consiste en que la información permanece accesible para su uso cuando así lo requieran las personas, servidores públicos o procesos autorizados.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Política para la Seguridad de la Información Estadística y Geográfica del Instituto Nacional de Estadística y Geografía*, Aguascalientes, 2020.

Edición de datos: procedimiento para detectar y corregir datos que están en blanco o no cumplen con ciertas reglas de consistencia de la información.

Fuente: F. Seguí, "Guía de la herramienta para la evaluación de la calidad de registros administrativos a ser usados con fines estadísticos", Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

Fuente de datos administrativa: dependencia u organismo público, privado o mixto responsable del registro administrativo y sus oficinas donde se llevan a cabo los procesos de captación y mantenimiento de los datos del registro.

Fuente: F. Seguí, "Guía de la herramienta para la evaluación de la calidad de registros administrativos a ser usados con fines estadísticos", Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

Fuente informante: entidad o persona que informa sobre la acción, evento o hecho objeto de registro.

Fuente: F. Seguí, "Guía de la herramienta para la evaluación de la calidad de registros administrativos a ser usados con fines estadísticos", Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

Gestión de la información: el desarrollo, ejecución y supervisión de planes, políticas, programas y procedimientos que aseguran, controlan, protegen y mejoran el valor de la información a lo largo de sus ciclos de vida.

Fuente: International Data Management Association (DAMA International), "Resumen Webinar: la importancia del gobierno del dato en el framework de DAMA", Madrid, 27 de mayo de 2020 [en línea] <https://anjanadata.com/resumen-la-importancia-del-gobierno-del-dato-en-el-framework-de-dama/>.

Gráfica: ilustración que representa datos estadísticos por medio de puntos, líneas y figuras, asociadas a escalas de medición.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario de Estadística Básica*, Aguascalientes, 2012.

Imputación de datos: procedimiento estadístico para asignar valores a un dato faltante o en sustitución de valores de respuesta no válidos o inconsistentes. El objetivo es reemplazar los valores faltantes, erróneos o inconsistentes utilizando variables auxiliares, por medio de procedimientos estadísticos estandarizados.

Fuente: F. Seguí, "Guía de la herramienta para la evaluación de la calidad de registros administrativos a ser usados con fines estadísticos", Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

Indicador: medición que relaciona uno o más conceptos mediante la aplicación de una metodología sobre fenómenos de interés, que permite su análisis y sirve de base para el establecimiento de objetivos y metas, así como para su seguimiento, desde el punto de vista de su magnitud, distribución o comportamiento en el tiempo y el espacio.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Reglas para la integración, difusión y administración del Catálogo Nacional de Indicadores*, Aguascalientes, 2015.

Indicador estadístico: variable cuantitativa cuyos valores son susceptibles de interpretación en un campo de conocimiento, respecto a determinados valores de referencia, establecidos en forma teórica o empírica.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Presentación de datos Estadísticos en Cuadros y Gráficas: actualización 2017*, Aguascalientes, 2017.

Información estadística: el conjunto de resultados cuantitativos o datos que se obtienen de las actividades estadísticas y geográficas en materia estadística, tomando como base los datos primarios obtenidos de los informantes del sistema sobre hechos que son relevantes para el conocimiento de los fenómenos económicos, demográficos y sociales, así como sus relaciones con el medio ambiente y el espacio territorial.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Norma para el Aseguramiento de la Calidad de la Información del Instituto Nacional de Estadística y Geografía*, Aguascalientes, 2020.

Información geográfica: conjunto de datos que posee un componente geométrico o espacial, que describe la localización de los objetos en el espacio y las relaciones espaciales entre ellos. También se entiende como información geográfica el producto de la georreferenciación de bases de datos temáticas que posean atributos geográficos.

Fuente: Departamento Administrativo de Estadística (DANE)/Instituto Geográfico Agustín Codazzi (IGAC), "Consolidación de la política nacional de información geográfica y la infraestructura colombiana de datos espaciales – ICDE", *Documento Conpes*, N° 3585, 2009.

Instrumento de captación: formato, en medio impreso o electrónico, diseñado para el registro de los datos que han de obtenerse de las unidades de observación, en un proyecto de generación de estadística básica.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario de Estadística Básica*, Aguascalientes, 2012.

Intervalo de clase: segmento, en una escala de medición, acotado convencionalmente por dos puntos distintos de dicha escala.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Presentación de datos estadísticos en cuadros y gráficas*, Aguascalientes, 2011.

Marco legal: conjunto de leyes, reglamentos, políticas y normas que fundamentan jurídicamente los registros que las dependencias y organismos de la administración pública realizan como parte de su función.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario de Estadística Básica*, Aguascalientes, 2012.

Medio del formato en que se hace el registro: tipo de instrumento físico utilizado para contener los formatos, puede ser de dos tipos: impreso y electrónico.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario de Estadística Básica*, Aguascalientes, 2010.

Metadatos: datos estructurados que describen las características del contenido, captura, procesamiento, calidad, condición, acceso y distribución de la información estadística o geográfica.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Norma Técnica del Proceso de Producción de Información Estadística y Geográfica para el Instituto Nacional de Estadística y Geografía*, Aguascalientes, 2020.

Metadatos estadísticos: información acerca de los datos estadísticos. Los metadatos comprenden datos y otra documentación que describe objetos de una manera formalizada.

Fuente: Naciones Unidas, "Terminology on statistical metadata", *Conference of European Statisticians Statistical Standards and Studies*, N° 53, Ginebra, 2000.

Microdato: corresponde a los datos sobre las características asociadas a las unidades de observación que se encuentran consolidadas en una base de datos.

Fuente: Departamento Administrativo Nacional de Estadística (DANE), *Norma técnica de la calidad del proceso estadístico: requisitos de calidad para la generación de estadísticas*, 2020 [en línea] <https://www.dane.gov.co/files/sen/normatividad/NTC-Proceso-Estadistico-PE-1000-2020.pdf>.

Operación estadística: conjunto de procesos y actividades que comprende la identificación de necesidades, el diseño, la construcción, la recolección o acopio, el procesamiento, el análisis, la difusión y la evaluación, que conduce a la producción de información estadística sobre un tema de interés nacional o territorial.

Fuente: Departamento Administrativo Nacional de Estadística (DANE), *Norma técnica de la calidad del proceso estadístico: requisitos de calidad para la generación de estadísticas*, 2020 [en línea] <https://www.dane.gov.co/files/sen/normatividad/NTC-Proceso-Estadistico-PE-1000-2020.pdf>.

Oportunidad: se refiere al tiempo que transcurre entre la ocurrencia del fenómeno de estudio y la publicación de la información estadística, de tal manera que sea útil para la toma de decisiones.

Fuente: Comisión Europea, *Guía de la estadística en la cooperación al desarrollo de la Comisión Europea*, Luxemburgo, 2013 [en línea] <http://ec.europa.eu/eurostat/documents/3859598/5926561/KS-RA-13-002-ES.PDF/c0174f3d-c08c-4f26-aa80-bf5ee6678eb3>.

Persona física o personas físicas: individuo(s) con capacidad para contraer obligaciones y ejercer derechos.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario del SNIEG 2021*, Aguascalientes, 2021.

Persona moral o personas morales: sujetos con derechos y obligaciones. Suelen ser creadas por un grupo de personas que se unen con un fin determinado y lícito.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario del SNIEG 2021*, Aguascalientes, 2021.

Pertinencia: se refiere a una evaluación cualitativa del valor aportado por los datos, el que se caracteriza por el grado en que los datos sirven para abordar los fines para los que son buscado por los usuarios. Depende tanto de la cobertura de los temas requeridos como del uso de conceptos apropiados.
Fuente: Organización de Cooperación y Desarrollo Económicos (OCDE), *Quality Framework and Guidelines for OECD Statistical Activities* (STD/QFS(2011)1), París, 2012.

Población de estudio: es el conjunto de unidades de observación, seres humanos, establecimientos, u otro tipo de elementos que se van a estudiar, en cierta ubicación de tiempo y espacio. En un esquema de tipo probabilístico y, a fines de calcular el tamaño de la muestra, se requiere conocer de ese conjunto una estimación de su varianza, el coeficiente de variación, el efecto del diseño, obtenidos de encuestas similares previas o de un estudio piloto.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario de Estadística Básica*, Aguascalientes, 2012.

Predio o inmueble: el terreno urbano o rústico que contiene o no construcción, que está sujeto en su caso a un régimen de propiedad con extensión y límites físicos reconocidos, en posesión y administrados por una sola entidad, ya sea de manera particular, colectiva, social o pública y es avalada por la autoridad competente.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario del SNIEG*, Aguascalientes, 2021.

Procesamiento: serie de actividades mediante las cuales se ordenan, almacenan y preparan los archivos con la información captada, asegurando su congruencia a fin de proceder a su explotación para la presentación de resultados estadísticos.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario de Estadística Básica*, Aguascalientes, 2012.

Proceso: conjunto de actividades, recursos humanos, datos e infraestructura relacionadas lógicamente para producir un resultado.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Norma Técnica del Proceso de Producción de Información Estadística y Geográfica para el Instituto Nacional de Estadística y Geografía*, Aguascalientes, 2020.

Producto estadístico: uno o varios indicadores, cuadros estadísticos y archivos de microdatos preparados para ser aprovechados con fines estadísticos, ya sea para la toma de decisiones, la definición de políticas o el análisis de series históricas.
Fuente: F. Seguí, "Guía de la herramienta para la evaluación de la calidad de registros administrativos a ser usados con fines estadísticos," Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

Proyecto estadístico o geográfico: conjunto de actividades ordenadas que se realiza para producir, integrar, analizar y difundir información estadística o geográfica, que permita cuantificar y caracterizar un universo de estudio o aspecto específico.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Reglas para la Prestación del Servicio Público de Información Estadística y Geográfica*, Aguascalientes, 2012.

Puntualidad: tiempo entre la entrega real de los datos y la fecha establecida en el calendario de publicación.
Fuente: Departamento Administrativo Nacional de Estadística (DANE), *Norma técnica de la calidad del proceso estadístico: requisitos de calidad para la generación de estadísticas*, 2020 [en línea] <https://www.dane.gov.co/files/sen/normatividad/NTC-Proceso-Estadistico-PE-1000-2020.pdf>.

Raspado web: proceso de recopilar información automáticamente de la World Wide Web, basado en herramientas (llamadas *scrapers*, robots de Internet, rastreadores, arañas, entre otros) que navegan, extraen el contenido de sitios web y almacenan datos extraídos en bases de datos locales para un posterior fin de elaboración.
Fuente: G. Barcaroli y D. Summa, "Using Internet as a data source for official statistics: a comparative analysis of web scraping technologies" [en línea] https://ec.europa.eu/eurostat/cros/system/files/Barcaroli-etal_WebScraping_Final_unblinded.pdf.

Referencia geográfica: área territorial a la que corresponde un conjunto de datos.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Presentación de datos Estadísticos en Cuadros y Gráficas: actualización 2017*, Aguascalientes, 2017.

Referencia temporal: momento, fecha o período al que corresponden los datos.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Presentación de datos Estadísticos en Cuadros y Gráficas: actualización 2017*, Aguascalientes, 2017.

Registro: aquel en el que se contienen los nombres geográficos de los objetos espaciales de todo el país.
Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario del SNIEG 2021*, Aguascalientes, 2021.

Registro administrativo: conjunto de datos que fueron generados con fines operacionales o como parte de las funciones de una institución pública o privada sobre un tipo de objeto, sujeto, acción, hecho o evento, y obtenidos sistemáticamente con base en un formato específico ya sea impreso, digital u otro y bajo un marco de funciones y facultades formalmente establecidas en instrumentos jurídicos o reglamentarios.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Norma Técnica del Proceso de Producción de Información Estadística y Geográfica para el Instituto Nacional de Estadística y Geografía*, Aguascalientes, 2020.

Registro base: registro estadístico de gran importancia para todo el sistema de registro. Debe definir tipos de objeto importantes, conjuntos de objetos importantes o poblaciones estandarizadas, y contiene vínculos con objetos en otros registros base.

Fuente: A. Wallgren y B. Wallgren, *Estadísticas basadas en registros: aprovechamiento estadístico de datos administrativos*, Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

Registro estadístico o geográfico: lista estructurada, completa y exhaustiva de las unidades de una población cuyo mantenimiento, evaluación de su calidad y actualización se lleva a cabo con fines estadísticos o geográficos.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Norma Técnica del Proceso de Producción de Información Estadística y Geográfica para el Instituto Nacional de Estadística y Geografía*, Aguascalientes, 2020.

Regulación administrativa: son los trámites y las formalidades administrativas con que los gobiernos recogen información e intervienen en decisiones económicas individuales.

Fuente: F. Segui, "Guía de la herramienta para la evaluación de la calidad de registros administrativos a ser usados con fines estadísticos," Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

Responsable del registro administrativo: unidad o unidades administrativas responsables de la gestión, captación, mantenimiento y actualización de los datos del registro administrativo.

Fuente: F. Segui, "Guía de la herramienta para la evaluación de la calidad de registros administrativos a ser usados con fines estadísticos," Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

Serie de tiempo: conjunto de datos estadísticos referentes a distintos momentos o intervalos de un período y a un mismo indicador, presentado en secuencia cronológica.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Presentación de datos Estadísticos en Cuadros y Gráficas: actualización 2017*, Aguascalientes, 2017.

Sistema de registros administrativos: conjunto de registros administrativos, así como sus metadatos, metodologías o especificaciones concretas utilizadas en su generación, que han sido captados por el responsable del registro administrativo y que la oficina nacional de estadística, garantizando su calidad, ha integrado en un sistema para su aprovechamiento.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI).

Transferencia de datos: es el envío de los datos del registro administrativo al usuario primario, por parte de la fuente administrativa.

Fuente: F. Segui, "Guía de la herramienta para la evaluación de la calidad de registros administrativos a ser usados con fines estadísticos," Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

Unidad de observación: elemento unitario del cual se obtienen datos con propósitos estadísticos sobre el conjunto al que pertenece.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario del SNIEG 2021*, Aguascalientes, 2021.

Unidad económica: unidad de observación sobre la cual se solicita y se publica información de carácter económico. De acuerdo con la actividad que realiza, puede ser una unidad agropecuaria o, una unidad no agropecuaria o establecimiento, pudiendo esta última ser establecimiento único, matriz o sucursal, de acuerdo con su tipo de organización; fijo o semifijo, de acuerdo con su tipo de instalación.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario del SNIEG 2021*, Aguascalientes, 2021.

Usuario final: es quien recibe finalmente el producto estadístico. Se trata de quien consume la información estadística contenida en la publicación final sobre el fenómeno de estudio basado en el registro administrativo. En algunos casos podrá tratarse de un usuario de la misma institución responsable del registro administrativo y en otros casos podrán ser las personas usuarias externas a esta.

Fuente: F. Segui, "Guía de la herramienta para la evaluación de la calidad de registros administrativos a ser usados con fines estadísticos," Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

Validación: conjunto de actividades para identificar, en la información captada, los datos que cumplen con los requisitos de congruencia lógica y aritmética, completez e integridad, a fin de aplicar, a los que no los cumplen, una solución bajo criterios específicos, que aseguren la eliminación de inconsistencias sin afectar los datos válidos originales.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario de Estadística Básica*, Aguascalientes, 2012.

Valor: cada uno de los números o modalidades nominales que puede admitir una variable.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Presentación de datos Estadísticos en Cuadros y Gráficas: actualización 2017*, Aguascalientes, 2017.

Variable: característica de las unidades estadísticas que puede asumir un conjunto de valores, los cuales pueden ser cuantitativos o cualitativos.

Fuente: Departamento Administrativo Nacional de Estadística (DANE), *Norma técnica de la calidad del proceso estadístico: requisitos de calidad para la generación de estadísticas*, 2020 [en línea] <https://www.dane.gov.co/files/sen/normatividad/NTC-Proceso-Estadistico-PE-1000-2020.pdf>.

Variable de clasificación: caso particular de una variable utilizada para clasificar las unidades de la población de interés, de acuerdo con ciertos criterios de clasificación.

Fuente: F. Seguí, "Guía de la herramienta para la evaluación de la calidad de registros administrativos a ser usados con fines estadísticos", Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

Variable de identificación: caso particular de una variable que permite hacer una identificación única de cada unidad o caso del registro administrativo.

Fuente: F. Seguí, "Guía de la herramienta para la evaluación de la calidad de registros administrativos a ser usados con fines estadísticos", Aguascalientes, Instituto Nacional de Estadística y Geografía (INEGI), 2012.

Veracidad: grado de aproximación de los cálculos o estimaciones a los valores exactos de la información estadística o geográfica que está destinada a representar.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario del SNIEG 2021*, Aguascalientes, 2021.

Verificación de datos: revisión detallada que hace el analista, quien se apoya en los criterios de validación establecidos en el diseño conceptual, de acuerdo con la temática abordada en el cuestionario. Tiene como objetivo identificar las inconsistencias inmediatamente después de levantada la información y poder corregirlas mediante reconsultas directas, dada la posibilidad de regresar con el informante. En el caso de los registros administrativos, la reconsulta se refiere a la fuente informante.

Fuente: Instituto Nacional de Estadística y Geografía (INEGI), *Glosario de Estadística Básica*, Aguascalientes, 2012.



Comisión Económica para América Latina y el Caribe (CEPAL)
Economic Commission for Latin America and the Caribbean (ECLAC)
www.cepal.org



LC/CEA.11/18/Rev.1