

# Modelo genérico del proceso estadístico GSBPM

(Versión 5.1, enero 2019)

## Sobre este documento

Este documento proporciona una descripción del GSBPM y cómo es que se relaciona con otras normas clave para la modernización estadística.



Este documento es una traducción al español de “Generic Statistical Process Model, GSBPM (Version 5.1, January 2019)” que está bajo la licencia “Creative Commons Atribución 4.0”. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by/4.0/>. Si utiliza la totalidad o parte de este trabajo, por favor, atribúyalo a la Comisión Económica de las Naciones Unidas para Europa (CEPE), en nombre de la comunidad estadística internacional.

# Índice

<b>I.</b>	<b>Introducción .....</b>	<b>3</b>
	Antecedentes .....	3
	Principales cambios respecto de la versión 5.0 .....	3
<b>II.</b>	<b>El Modelo.....</b>	<b>4</b>
	Comprensión del GSBPM .....	4
	La estructura .....	5
	Aplicación.....	6
	Utilización del GSBPM .....	6
<b>III.</b>	<b>Relación con Otros Modelos y Normas.....</b>	<b>7</b>
	GAMSO .....	7
	GSIM .....	7
<b>IV.</b>	<b>Niveles 1 y 2 del GSBPM.....</b>	<b>9</b>
<b>V.</b>	<b>Descripción de las fases y subprocesos.....</b>	<b>10</b>
	Fase de especificación de necesidades .....	10
	Fase de diseño.....	12
	Fase de construcción.....	14
	Fase de recolección.....	17
	Fase de procesamiento .....	19
	Fase de análisis .....	22
	Fase de difusión .....	24
	Fase de evaluación .....	26
<b>VI.</b>	<b>Procesos globales.....</b>	<b>27</b>
<b>VII.</b>	<b>Otros usos del GSBPM.....</b>	<b>31</b>
	<b>Anexo – Lista de acrónimos .....</b>	<b>33</b>

# Modelo genérico del proceso estadístico (GSBPM<sup>1</sup>)

## I. Introducción

1. El GSBPM describe y define el conjunto de procesos necesarios para producir estadísticas oficiales. Proporciona un marco estándar y una terminología armonizada para ayudar a las organizaciones estadísticas a modernizar sus procesos de producción estadística, así como a compartir métodos y componentes. Además, el GSBPM puede ser usado para integrar normas en datos y metadatos, como modelo para la documentación de procesos, para la armonización de infraestructuras informáticas estadísticas y para proporcionar un marco de referencia para mejoras y evaluaciones del proceso de calidad. Estos y otros propósitos para los cuales el GSBPM puede ser utilizado, se desarrollan en la Sección VII. Esta versión del GSBPM está alineada con la versión 1.2 del Modelo Genérico de Información Estadística (GSIM) y versión 1.2 del Modelo Genérico de Actividad de los Organismos Estadísticos (GAMSO).

## Antecedentes

2. El GSBPM se desarrolló por primera vez en 2008 por el trabajo conjunto de CEPE/Eurostat/OCDE sobre Metadatos Estadísticos (METIS), basado en el modelo de procesos utilizado por Statistics New Zealand. Tras varios borradores y consultas públicas, en abril de 2009 se publicó la versión 4.0 del GSBPM. Posteriormente, fue ampliamente adoptado por la comunidad mundial de estadísticas oficiales y constituyó una de las bases de la visión y estrategia del Grupo de Alto Nivel para la Modernización de las Estadísticas Oficiales (HLG-MOS)<sup>2</sup> para una modernización basada en normas.
3. La versión anterior del GSBPM (versión 5.0) se publicó en diciembre de 2013 y en 2018 se llevó a cabo una amplia revisión para incorporar la retroalimentación basada en la aplicación práctica y para mejorar la consistencia con los nuevos modelos del HLG-MOS, el GSIM y el GAMSO. Aunque la versión actual del GSBPM (versión 5.1) se considera definitiva en el momento de su publicación, también se espera que en los próximos años puedan ser necesarias futuras actualizaciones, ya sea para reflejar nuevas experiencias de la aplicación del modelo en la práctica, o bien debido a la evolución de la naturaleza de la producción estadística. Por lo tanto, se invita al lector a revisar el wiki de GSBPM<sup>3</sup> para asegurarse de que tiene la última versión.

## Principales cambios respecto de la versión 5.0

4. La revisión de la versión 5.0 de la GSBPM tenía un mandato claro del HLG-MOS de introducir únicamente los cambios que contaran con un caso de negocio sólido y un apoyo generalizado por parte de la comunidad estadística internacional. La razón de este enfoque era que muchas organizaciones habían invertido mucho en la implementación del GSBPM. Los cambios importantes podrían implicar costos significativos y, en última instancia, ser contraproducentes en términos de asegurar la adopción y el uso generalizados del modelo.
5. Los principales cambios en el GSBPM entre las versiones 5.0 y 5.1 son los siguientes:
  - Algunos subprocesos han sido renombrados para mejorar la claridad;

---

<sup>1</sup> Se puede encontrar la forma completa de todos los acrónimos usado en este documento en el anexo al final de este documento.

<sup>2</sup> UNECE (CEPE) Statistics Wikis - HLG-MOS (<https://statswiki.unece.org/display/hlgbas>)

<sup>3</sup> UNECE (CEPE) Statistics Wikis - GSBPM (<https://statswiki.unece.org/display/GSBPM>)

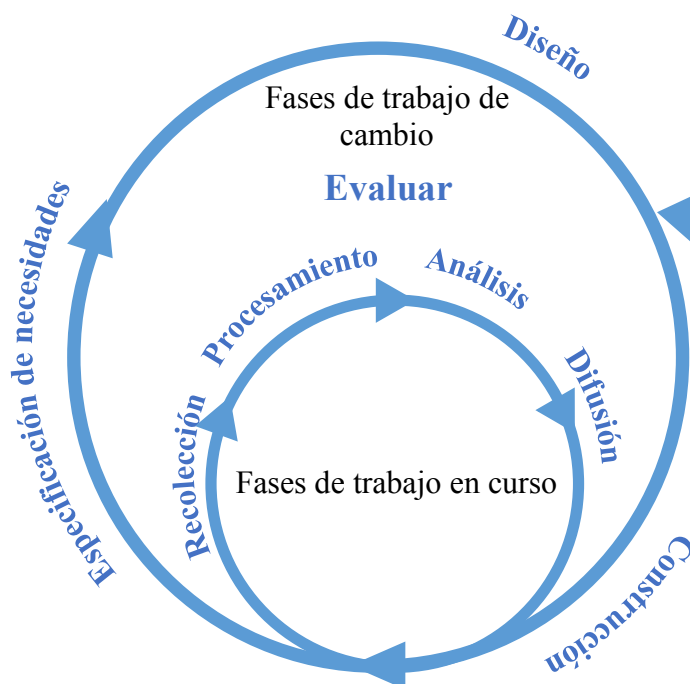
- Se ha resuelto la duplicación entre los procesos globales en la versión 5.0 de GSBPM y la GAMSO;
- Las descripciones de las fases y los subprocesos se han actualizado para que estén menos centrados en la encuesta y se han añadido actividades relacionadas con el trabajo con proveedores de datos no estadísticos cuando ha sido necesario;
- En reconocimiento de la creciente importancia de integrar los datos estadísticos con los datos geoespaciales, las descripciones se han ampliado para incluir las tareas necesarias para utilizar los datos geoespaciales;
- Se ha modificado la terminología cuando ha sido necesario para mejorar la coherencia con la GAMSO y la GSIM;
- Los ejemplos y descripciones se han actualizado y ampliado para mejorar la claridad.

## **II. El Modelo**

### **Comprensión del GSBPM**

6. Un proceso estadístico es un conjunto de actividades y tareas relacionadas y estructuradas para convertir los datos de entrada en información estadística. En el contexto de la GSBPM, las organizaciones o grupos de organizaciones realizan procesos estadísticos para crear estadísticas oficiales que satisfagan las necesidades de los usuarios. El resultado del proceso puede ser un conjunto mixto de productos físicos o digitales que presentan datos y metadatos de diferentes maneras, como publicaciones, mapas, servicios electrónicos, entre otros.
7. El GSBPM debe ser aplicado e interpretado de manera flexible. No es un marco rígido en el que todos los pasos se deben seguir en un orden estricto, sino que identifica los posibles pasos dentro del proceso estadístico y las interdependencias entre ellos.
8. Aunque la presentación del GSBPM sigue una secuencia lógica de pasos en la mayoría de los procesos estadísticos, los elementos del modelo pueden presentarse en diferente orden, de acuerdo a las diferentes circunstancias. Además, algunos subprocesos serán revisados en varias ocasiones, formando ciclos iterativos, particularmente dentro de las fases de Procesamiento y Análisis.
9. El GSBPM puede ser visto como una lista de control para asegurarse de que todos los pasos necesarios han sido considerados o como un “libro de cocina” para identificar todos los “ingredientes” de un proceso estadístico.

**Figura 1.** GSBPM, Fases de trabajo en curso y fases de trabajo de cambio



10. En muchas organizaciones estadísticas, las primeras fases sólo se consideran cuando se crea un nuevo producto o cuando el proceso se revisa como resultado de un proceso de evaluación. Una vez que el producto pasa a formar parte de la actividad continua “normal”, estas fases no se llevan a cabo (por ejemplo, no es necesario crear nuevas herramientas de recolección cada vez que se recolectan datos de la encuesta de población activa). Véase la figura 1.
11. Por lo tanto, el GSBPM debe considerarse más bien como una matriz, a través de la cual hay muchas vías posibles. De esta manera, el GSBPM pretende ser lo suficientemente genérico como para ser ampliamente aplicable y también pretende fomentar una visión estándar del proceso estadístico, sin ser demasiado restrictivo o abstracto y teórico.

### La estructura

12. El GSBPM comprende tres niveles:
  - Nivel 0, el proceso estadístico;
  - Nivel 1, las ocho fases del proceso estadístico;
  - Nivel 2, los subprocesos dentro de cada fase.
13. Un diagrama que muestra las fases (nivel 1) y los subprocesos (nivel 2) está incluido en la Sección IV (véase figura 3). Los subprocesos se describen detalladamente, en la Sección V.
14. El GSBPM reconoce varios procesos globales con un fuerte componente estadístico que se aplican a lo largo de las ocho fases. Estos procesos globales incluyen la siguiente lista. La gestión de la calidad, la gestión de metadatos y la gestión de datos se desarrollan con más detalle en la Sección

VI. Las actividades que se llevan a cabo a nivel de la organización para apoyar la producción estadística se incluyen en la GAMS0 (véase Sección III)<sup>4</sup>.

## **Aplicación**

15. El GSBPM está destinado a aplicarse a todas las actividades emprendidas por los productores de estadísticas oficiales, tanto a nivel nacional como internacional, que dan lugar a la producción de datos.
16. El modelo está diseñado para ser aplicable independientemente de la fuente de datos, por lo que puede utilizarse para la descripción y evaluación de la calidad de los procesos basados en encuestas, censos, registros administrativos y otras fuentes no estadísticas o mixtas.
17. Mientras que los procesos estadísticos típicos incluyen la recolección y el procesamiento de datos para producir resultados estadísticos, el GSBPM también se aplica cuando se revisan los datos existentes o se vuelven a calcular las series temporales, ya sea como resultado de la mejora de las fuentes de datos o de un cambio en la metodología. En estos casos, los datos de entrada pueden ser microdatos originales y/o datos adicionales, que luego se procesan y analizan para producir resultados revisados. En tales casos, es probable que se omitan varios subprocesos y posiblemente algunas fases (especialmente las primeras). Del mismo modo, el GSBPM puede aplicarse a procesos como la compilación de las cuentas nacionales y a los procesos típicos de las organizaciones estadísticas internacionales que utilizan datos secundarios de países u otras organizaciones.
18. Además de ser aplicable a los procesos que dan lugar a estadísticas, el GSBPM también puede aplicarse al desarrollo y mantenimiento de registros estadísticos, en los que los datos son similares a los de la producción estadística (aunque normalmente se centran más en los datos administrativos), y los resultados son típicamente marcos u otras extracciones de datos, que luego se utilizan como insumos para otros procesos<sup>5</sup>.
19. El GSBPM es suficientemente flexible para ser aplicado en todos los escenarios anteriores.

## **Utilización del GSBPM**

20. El GSBPM es un modelo de referencia. Se pretende que el GSBPM pueda ser utilizado por organizaciones en diferentes grados. Una organización puede elegir entre implementar el GSBPM directamente o utilizarlo como base para desarrollar una adaptación específica de la organización al modelo. Puede usarlo en algunos casos solamente como un modelo de referencia para la comunicación interna de las organizaciones o con otras para aclarar la discusión. Los diferentes escenarios para el uso del GSBPM son todos válidos.
21. Cuando las organizaciones han desarrollado adaptaciones específicas del GSBPM, pueden hacer algunas modificaciones especiales al modelo para adaptarse a su contexto organizacional. La evidencia hasta ahora sugiere que estas modificaciones especiales no son suficientemente generales como para ser incluidas en el propio GSBPM.

---

<sup>4</sup> Se está perfeccionando la armonización de los procesos generales de la GSBPM con la GAMS0. El trabajo estará disponible (en inglés) en los Wikis de Estadísticas de la CEPE - HLG-MOS Outputs (<https://statswiki.unece.org/display/hlgbas/HLG-MOS+Outputs>; se espera para el tercer trimestre de 2019).

<sup>5</sup> Aplicación del modelo de proceso estadístico genérico al mantenimiento del registro de empresas (documento disponible (en inglés) a <https://www.unece.org/index.php?id=18256>)

22. En algunos casos puede ser apropiado agrupar algunos de los elementos del modelo. Por ejemplo, las tres primeras fases podrían agruparse en una única fase de planificación. En otros casos, particularmente para implementaciones prácticas, puede ser necesario añadir uno o más niveles detallados a la estructura para identificar por separado los diferentes componentes de los subprocesos.
23. También puede haber un requisito de cierre de sesión formal entre fases, donde los productos de una fase están certificados como insumos adecuados para la siguiente. Este tipo de aprobación formal está implícita en el modelo (excepto en el subproceso 1.6), pero puede implementarse de diferentes maneras, dependiendo de las necesidades de la organización.

### **III. Relación con Otros Modelos y Normas**

24. Desde la publicación del GSBPM, se han desarrollado varios modelos bajo los auspicios del HLG-MOS para apoyar la modernización de las estadísticas oficiales. En conjunto, estos modelos se denominan “ModernStats”. Los siguientes párrafos describen los modelos de ModernStats que tienen un fuerte vínculo con el GSBPM.

#### **GAMSO**

25. El GAMSO<sup>6</sup> describe y define las actividades que ocurren dentro de una organización estadística típica. Amplía y complementa la GSBPM añadiendo actividades necesarias para apoyar la producción estadística (es decir, actividades en las áreas de estrategia y liderazgo, desarrollo de capacidades y apoyo corporativo). En el GSBPM v. 5.0, algunas de estas actividades se incluyeron como procesos globales. Las actividades que no están directamente relacionadas con la producción de estadísticas y/o que se gestionan a nivel corporativo o estratégico se incluyen ahora en el GAMSO (por ejemplo, gestión de recursos humanos y actividades de gestión de la calidad que se llevan a cabo a nivel corporativo, como el desarrollo de un marco de calidad).
26. El GAMSO describe las actividades, es decir, lo que hacen las organizaciones estadísticas. Incluye descripciones de alto nivel de estas actividades. Por otra parte, el GSBPM se centra en el proceso de producción y describe con más detalle cómo las organizaciones estadísticas llevan a cabo la actividad de producción estadística.
27. Al igual que el GSBPM, el GAMSO tiene como objetivo proporcionar un vocabulario y un marco común para apoyar las actividades de colaboración internacional. Se obtendrá un mayor valor de la aplicación de la GAMSO si se aplica conjuntamente con la GSBPM.

#### **GSIM**

28. El GSIM<sup>7</sup> es un marco de referencia para la información estadística, diseñado para ayudar a modernizar las estadísticas oficiales tanto a nivel nacional como internacional. Permite descripciones genéricas de la definición, gestión y uso de datos y metadatos a lo largo de todo el proceso de producción estadística. Proporciona un conjunto de objetos de información estandarizados y descritos de forma coherente, que son los insumos y resultados de los subprocesos GSBPM. El GSIM ayuda a explicar las relaciones significativas entre las entidades involucradas en la producción estadística, y puede ser utilizado para guiar el desarrollo y uso de normas o especificaciones de implementación consistentes.

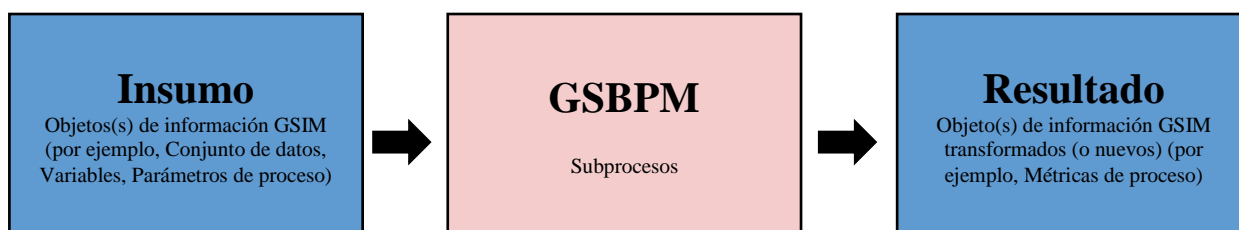
---

<sup>6</sup> UNECE (CEPE) Statistics Wikis - GAMSO (<https://statswiki.unece.org/display/GAMSO>)

<sup>7</sup> UNECE (CEPE) Statistics Wikis - GSIM (<https://statswiki.unece.org/display/GSIM>)

29. Al igual que el GSBPM, el GSIM es una de las claves para modernizar las estadísticas oficiales y para alejarse de las agrupaciones temáticas. Identifica alrededor de 130 objetos de información; ejemplos incluyen conjuntos de datos, variables, clasificaciones estadísticas, unidades, poblaciones, así como las reglas y parámetros necesarios para la ejecución de los procesos de producción (por ejemplo, reglas de edición de datos).
30. El GSIM y el GSBPM son modelos complementarios para la producción y gestión de la información estadística. Como se puede observar en la figura 2, el GSIM ayuda a describir los subprocesos GSBPM con la definición de los objetos de información que fluyen entre ellos, que se crean en ellos y que son utilizados por ellos para producir estadísticas oficiales. Los insumos y resultados pueden definirse en términos de objetos de información y se formalizan en el GSIM.

**Figura 2.** Los objetos de información GSIM como insumo y resultado del subproceso GSBPM



31. Se obtendrá un mayor valor del GSIM si se aplica conjuntamente con el GSBPM. Asimismo, se obtendrá un mayor valor del GSBPM si se aplica conjuntamente con el GSIM. Sin embargo, es posible (aunque no ideal) aplicar uno sin el otro.
32. La aplicación conjunta del GSIM y el GSBPM puede facilitar la creación de sistemas eficientes basados en metadatos y ayudar a armonizar las infraestructuras informáticas estadísticas.



#### IV. Niveles 1 y 2 del GSBPM

**Figura 3.** Las fases (nivel 1) y subprocesos (nivel 2) del GSBPM

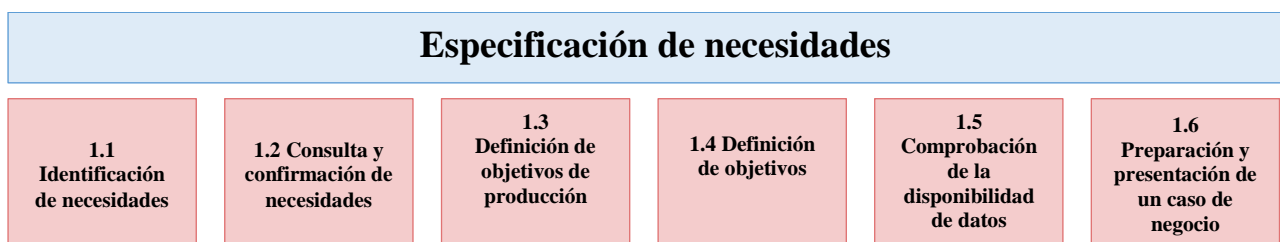
Procesos globales							
Especificación de necesidades	Diseño	Construcción	Recolección	Procesamiento	Análisis	Difusión	Evaluación
1.1 Identificación de necesidades	2.1 Diseño de productos	3.1 Reutilizar o construir instrumentos de recolección	4.1 Conformación del marco muestral y selección de la muestra	5.1 Integración de datos	6.1 Preparación de borradores de resultados	7.1 Actualización de sistemas de salida	8.1 Recolección de insumos para la evaluación
1.2 Consulta y confirmación de necesidades	2.2 Diseño de las descripciones de las variables	3.2 Reutilizar o construir componentes de procesamiento y análisis	4.2 Preparación de la recolección	5.2 Clasificación y codificación	6.2 Validación de los resultados	7.2 Generación de productos de difusión	8.2 Evaluación
1.3 Definición de objetivos de producción	2.3 Diseño de la recolección	3.3 Reutilizar o construir componentes de difusión	4.3 Ejecución de la recolección	5.3 Revisión y validación	6.3 Interpretación y explicación de los resultados	7.3 Gestión de la publicación de productos de difusión	8.3 Determinación de un plan de acción
1.4 Identificación de conceptos	2.4 Diseño del marco muestral y del muestreo	3.4 Configuración de flujos de trabajo	4.4 Cierre de la recolección	5.4 Edición e imputación	6.4 Aplicación del control a la divulgación	7.4 Promoción de productos de difusión	
1.5 Comprobación de la disponibilidad de datos	2.5 Diseño del procesamiento y análisis	3.5 Pruebas al sistema de producción		5.5 Derivación de nuevas variables y unidades	6.5 Finalización de resultados	7.5 Gestión de soporte a usuarios	
1.6 Preparación y presentación de un caso de negocio	2.6 Diseño de los sistemas de producción y de los flujos de trabajo	3.6 Prueba piloto del proceso estadístico		5.6 Cálculo de ponderadores			
		3.7 Finalización del sistema de producción		5.7 Cálculo de agregados			
				5.8 Finalización de los archivos de datos			

## V. Descripción de las fases y subprocesos

33. Esta sección considera cada fase e identifica varios subprocesos dentro de éstas, describiendo su contenido.

### Fase de especificación de necesidades

**Figura 4.** Especificación de necesidades y sus subprocesos



34. Esta fase se inicia cuando se identifica la necesidad de nuevas estadísticas o cuando la retroalimentación sobre las estadísticas actuales inicia una revisión. Incluye todas las actividades asociadas con la participación de las partes interesadas para identificar sus necesidades estadísticas detalladas (actuales o futuras), proponer soluciones de alto nivel y preparar un caso de negocio para satisfacer estas necesidades.

35. La fase de especificación de necesidades se divide en seis subprocesos (véase figura 4). Generalmente, los subprocesos son secuenciales, de izquierda a derecha, pero también pueden ocurrir en paralelo y pueden ser iterativos. Estos subprocesos lo son:

#### 1.1. Identificación de necesidades

36. Este subproceso incluye la investigación inicial y la identificación de qué estadísticas se necesitan y qué se necesita de ellas. Puede ser provocado por una nueva solicitud de información o por un cambio organizacional, como un presupuesto reducido. Los planes de acción de las evaluaciones de iteraciones anteriores del proceso o de otros procesos podrían proporcionar un insumo a este subproceso. También incluye la consideración de la práctica entre otras organizaciones estadísticas (nacionales e internacionales) que producen datos similares y los métodos utilizados por dichas organizaciones.

#### 1.2. Consulta y confirmación de necesidades

37. Este subproceso se centra en la consulta con las partes interesadas internas y externas y en la confirmación detallada de las necesidades de las estadísticas. Se requiere una buena comprensión de las necesidades de los usuarios para que la organización estadística sepa no sólo lo que se espera de ella, sino también cuándo, cómo y, quizás lo más importante, por qué. Para la segunda iteración y las siguientes de esta fase, la atención se centrará principalmente en determinar si han cambiado las necesidades previamente identificadas. Esta comprensión detallada de las necesidades de los usuarios es la parte crítica de este subproceso.

### **1.3. Definición de objetivos de producción**

38. Este subproceso identifica los objetivos de producción estadística que son necesarios para satisfacer las necesidades de los usuarios identificadas en el subproceso 1.2 (consulta y confirmación de necesidades). Incluye acordar con los usuarios la pertinencia de los productos propuestos y sus medidas de calidad. Es probable que los marcos jurídicos (por ejemplo, en relación con la confidencialidad) y los recursos disponibles constituyan limitaciones a la hora de establecer los objetivos de los productos.

### **1.4. Identificación de conceptos**

39. Este subproceso aclara los conceptos necesarios que deben medirse desde el punto de vista de los usuarios. En esta fase, es posible que los conceptos identificados no se ajusten a las normas estadísticas existentes. Esta alineación, y la elección o definición de los conceptos y variables estadísticas y de otro tipo que se utilizarán, se realiza en el subproceso 2.2 (descripciones de variables de diseño).

### **1.5. Comprobación de la disponibilidad de datos**

40. Este subproceso comprueba si las fuentes actuales de datos podrían cumplir los requisitos de los usuarios y las condiciones en las que estarían disponibles, incluidas las restricciones sobre su uso. Una evaluación de las posibles alternativas incluiría normalmente la investigación de posibles fuentes de datos administrativas u otras fuentes no estadísticas para:

- Determinar si serían adecuados para su uso con fines estadísticos (por ejemplo, en qué medida los conceptos administrativos se ajustan a los requisitos de los datos, la puntualidad y la calidad de los datos, la seguridad y la continuidad del suministro de datos);
- Evaluar la división de responsabilidades entre los proveedores de datos y la organización estadística;
- Comprobar los recursos de TIC necesarios (por ejemplo, almacenamiento de datos, tecnología necesaria para manejar los datos entrantes y para el procesamiento de datos), así como cualquier acuerdo formal con los proveedores de datos para acceder a los datos y compartirlos (por ejemplo, formatos, entrega, metadatos de acompañamiento y control de calidad).

41. Una vez evaluadas las fuentes existentes, se prepara una estrategia para llenar los vacíos que subsistan en la exigencia de datos. Esto puede incluir la identificación de posibles asociaciones con los encargados de los datos. Este subproceso también incluye una evaluación más general del marco jurídico en el que se recolectan y utilizarán los datos y, por lo tanto, puede identificar propuestas para modificar la legislación existente o para introducir introducción un nuevo marco jurídico.

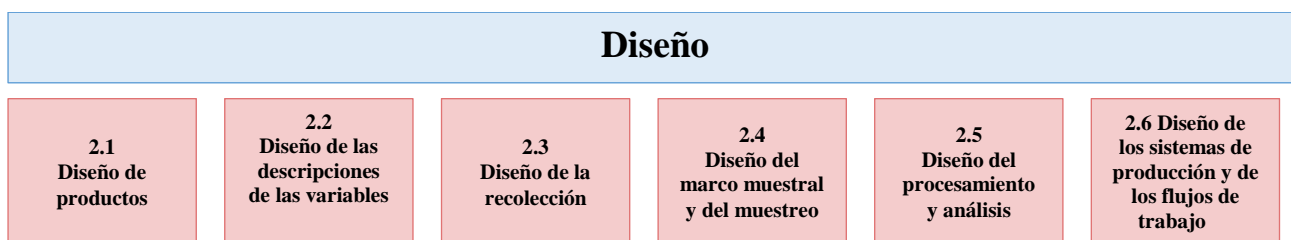
### **1.6. Preparación y presentación de un caso de negocio**

42. Este subproceso documenta los resultados de los otros subprocesos en esta fase en forma de un caso de negocio para obtener la aprobación para implementar el proceso estadístico nuevo o modificado. Dicho caso de negocio tendría que ajustarse a los requisitos del organismo de aprobación, pero normalmente incluiría elementos tales como:

- Una descripción del proceso “tal cual” (si ya existe), con información sobre cómo se elaboran las estadísticas actuales, destacando las ineficiencias y las dificultades que deben abordarse;
  - La solución “futura” propuesta, que detalla cómo se desarrollará el proceso estadístico para producir las estadísticas nuevas o revisadas;
  - Una evaluación de los costos y beneficios, así como de cualquier restricción externa.
43. El caso de negocio describe las opciones y hace recomendaciones. Puede incluir los beneficios, los costos, los resultados, el calendario, el presupuesto, los recursos técnicos y humanos necesarios, la evaluación de riesgos y el impacto en las partes interesadas para cada opción.
44. Después de que se prepara el caso de negocio, se somete a aprobación para pasar a la siguiente fase del proceso. En este subproceso, se toma una decisión de “ir/no ir”. Típicamente, el caso de negocio es revisado y aprobado o desaprobado formalmente por los patrocinadores apropiados y los comités de administración.

## Fase de diseño

**Figura 5.** Fase de diseño y sus subprocesos



45. En esta fase se describen las actividades de diseño y desarrollo, así como cualquier trabajo práctico de investigación asociado que se necesite para definir los productos, los conceptos, metodologías, instrumentos recolección y los procesos operativos de las estadísticas. Incluye todos los elementos de diseño necesarios para definir o refinar los productos estadísticos o los servicios identificados en el caso de negocios, así como en los procedimientos de aseguramiento de la calidad. Para los productos estadísticos elaborados de forma regular, esta fase ocurre normalmente durante la primera iteración y siempre que las acciones de mejora sean identificadas en la fase de evaluación de una iteración anterior.
46. Las actividades de diseño hacen uso sustancial de normas internacionales y nacionales con el fin de reducir la duración y el costo del proceso de diseño y mejorar la comparabilidad y uso de los resultados. También se alienta a las organizaciones a reutilizar o adaptar los elementos de diseño de procesos existentes, y considerar los aspectos geoespaciales de los datos en el diseño para mejorar la utilidad y el valor de la información estadística. Adicionalmente, los productos de los procesos de diseño pueden formar la base para las futuras normas a nivel organizacional, nacional o internacional.
47. Esta fase se divide en seis subprocesos que generalmente son secuenciales, de izquierda a derecha, pero también pueden darse en paralelo y pueden ser iterativos. Estos subprocesos son:

### 2.1. Diseño de productos

48. Este subproceso contiene el diseño detallado de los resultados, productos y servicios estadísticos que se producirán, incluyendo el trabajo de desarrollo relacionado y la preparación de los sistemas y herramientas utilizados en la fase de difusión. También se diseñan aquí los procesos que rigen el acceso a cualquier resultado confidencial. Los resultados deben ser diseñados para seguir las normas existentes siempre que sea posible, de modo que los insumos a este proceso puedan incluir metadatos de recolecciones similares o anteriores (incluyendo extracciones de registros y bases de datos estadísticos, administrativos, geospaciales y otros registros y bases de datos no estadísticos), normas internacionales, e información sobre las prácticas de otras organizaciones estadísticas del subproceso 1.1 (identificación de necesidades). Los productos también pueden diseñarse en colaboración con otros organismos interesados, en particular si se consideran productos conjuntos, o si serán difundidos por otra organización.

## **2.2. Diseño de las descripciones de las variables**

49. Este subproceso define las variables que se recolectarán a través del instrumento de recolección, así como cualquier otra variable que se derive de ellas en el subproceso 5.5 (derivación de nuevas variables y unidades), y cualquier clasificación estadística o geoespacial que se vaya a utilizar. Se espera que, siempre que sea posible, se cumplan las normas nacionales e internacionales vigentes. Es posible que este subproceso deba ejecutarse en paralelo con el subproceso 2.3 (diseño de la recolección), ya que la definición de las variables que deben recolectar y la elección de los instrumentos de recolección pueden ser hasta cierto punto interdependientes. La preparación de descripciones de metadatos de las variables recolectadas y derivadas y la clasificación estadística y geoespacial es una condición previa necesaria para las fases siguientes.

## **2.3. Diseño de la recolección**

50. Este subproceso determina los instrumentos y métodos de recolección más apropiados que pueden depender del tipo de recolección de datos (censo, encuesta por muestreo u otro), el tipo de unidad de recolección (empresa, persona u otro) y las fuentes de datos disponibles. Las actividades reales de este subproceso variarán en función del tipo de instrumento de recolección necesario, que puede incluir entrevistas asistidas por computadora, cuestionarios en papel, registros administrativos (por ejemplo, utilizando las interfaces de servicio existentes), métodos de transferencia de datos, tecnologías de raspado en la web, así como tecnología para datos geospaciales. El uso directo o indirecto de los datos administrativos puede introducirse en el modo de recolección de datos, ya sea para controlar los datos de la encuesta o para ayudarla a capturar la información de la misma.
51. Este subproceso incluye el diseño de los instrumentos de recolección, las plantillas de preguntas y respuestas (junto con las variables y clasificaciones estadísticas diseñadas en el subproceso 2.2 (diseño de las descripciones de las variables)). También incluye la confirmación de cualquier acuerdo formal. Este subproceso es posible gracias a herramientas como las bibliotecas de preguntas (para facilitar la reutilización de las preguntas y los atributos relacionados), las herramientas de cuestionarios (para permitir la rápida y fácil compilación de preguntas en formatos adecuados para las pruebas cognitivas) y las plantillas de acuerdos (para ayudar a estandarizar los términos y condiciones). Este subproceso también incluye el diseño de sistemas de gestión de proveedores que son específicos para este proceso.
52. Cuando las organizaciones estadísticas no recolectan datos directamente (es decir, cuando un tercero controla la recolección y el tratamiento de los datos), este subproceso puede incluir el diseño de mecanismos para supervisar los datos y los metadatos para evaluar el impacto de cualquier cambio realizado por el tercero.

## 2.4. Diseño del marco muestral y del muestreo

53. Este subproceso sólo se aplica a los procesos que implican la recolección de datos basados en el muestreo, por ejemplo, a través de encuestas estadísticas. Identifica y especifica la población de interés, define un marco muestral (y, cuando sea necesario, el registro del que se deriva) y determina los criterios y la metodología de muestreo más apropiados (que podrían incluir la enumeración completa). Las fuentes comunes para un marco muestral son los registros administrativos y estadísticos, los censos y la información de otras encuestas por muestreo. Puede incluir datos geospaciales y clasificaciones. Este subproceso describe cómo se pueden combinar estas fuentes si es necesario. Se debe analizar si el marco cubre la población objetivo. Debe elaborarse un plan de muestreo. La muestra real se crea en el subproceso 4.1 (conformación del marco muestral y selección de la muestra), utilizando la metodología especificada en este subproceso.

## 2.5. Diseño del procesamiento y análisis

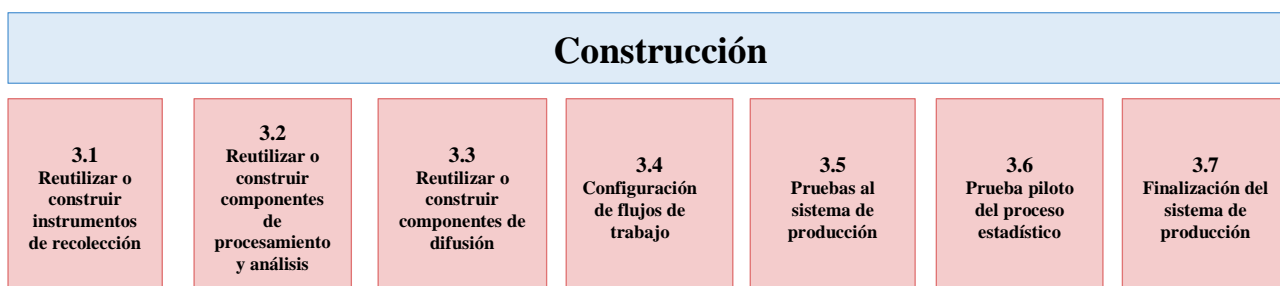
54. Este subproceso diseña la metodología de procesamiento estadístico que se aplicará durante las fases de procesamiento y análisis. Esto puede incluir, entre otras cosas, la especificación de rutinas y reglas para la codificación, edición e imputación, que pueden variar en función del modo de recolección de datos y de la fuente de los mismos. Este subproceso también incluye el diseño de especificaciones para la integración de datos de múltiples fuentes de datos, la validación de datos y la estimación. Los métodos de control de divulgación de información estadística también se diseñan aquí si son específicos de este proceso.

## 2.6. Diseño de los sistemas de producción y de los flujos de trabajo

55. Este subproceso determina el flujo de trabajo desde la recolección de datos hasta la difusión, tomando una visión general de todos los procesos requeridos dentro de todo el proceso de producción y asegurándose de que encajan eficientemente sin vacíos ni redundancias. Se necesitan varios sistemas y bases de datos a lo largo del proceso. El GSBPM puede utilizarse como base de la capa de arquitectura cuando una organización estadística dispone de una arquitectura existente. El diseño puede ser ajustado para adaptarse a la organización. Un principio general es reutilizar los procesos y la tecnología en muchos procesos estadísticos, por lo que las soluciones de producción existentes (por ejemplo, servicios, sistemas y bases de datos) deben examinarse en primer lugar para determinar si son adecuadas para este proceso de producción específico y, a continuación, si se detectan vacíos, deben diseñarse nuevas soluciones. Este subproceso también considera cómo el personal interactuará con los sistemas y quién será responsable de qué y cuándo.

## Fase de construcción

**Figura 6.** Fase de construcción y sus subprocesos



56. Esta fase construye y prueba la solución de producción hasta el punto de que esté lista para su uso en el entorno “vivo”. Los productos de la fase de diseño se ensamblan y configuran en esta fase para crear el entorno operativo completo para ejecutar el proceso. Los nuevos servicios se construyen por excepción, creados en respuesta a los vacíos en el catálogo existente de servicios que provienen de dentro de la organización y de fuera de ella. Estos nuevos servicios están contruidos para ser ampliamente reutilizables en alineación con la arquitectura de la organización siempre que sea posible.
57. En el caso de los resultados estadísticos producidos de forma regular, esta fase suele producirse para la primera iteración, tras una revisión o un cambio de metodología o tecnología, y no para cada iteración.
58. La fase de construcción se divide en siete subprocesos (véase figura 6), que generalmente son secuenciales, de izquierda a derecha, pero también pueden ocurrir en paralelo, y pueden ser iterativos. Los tres primeros subprocesos se refieren al desarrollo y la mejora de los sistemas utilizados en la recolección, el tratamiento, el análisis y la difusión de datos. Los últimos cuatro subprocesos se centran en el proceso de principio a fin. Estos subprocesos son:

### **3.1. Reutilizar o construir de instrumentos de recolección**

59. Este subproceso describe las actividades de construcción y reutilización de los instrumentos de recolección que se utilizarán durante la fase de recolección. Los instrumentos de la recolección se construyen en base a las especificaciones de diseño creadas durante la fase de diseño. Una recolección puede utilizar uno o más modos para recibir los datos (por ejemplo, entrevistas personales o telefónicas; cuestionarios en papel, electrónicos o en línea; servicios web de SDMX). Los instrumentos de recolección pueden ser también rutinas de extracción de datos utilizadas para recolectar datos de registros estadísticos o administrativos existentes (por ejemplo, utilizando interfaces de servicio existentes). Este subproceso también incluye la preparación y comprobación del contenido y el funcionamiento de ese instrumento de recolección (por ejemplo, la comprobación cognitiva de las preguntas de un cuestionario). Se recomienda considerar la conexión directa de los instrumentos de recolección a un sistema de metadatos, de modo que los metadatos puedan capturarse más fácilmente en la fase de recolección. La conexión de metadatos y datos en el punto de captura puede ahorrar trabajo en fases posteriores. La captura de las métricas de la recolección de datos (paradatos) es también una consideración importante en este subproceso para calcular y analizar los indicadores de calidad del proceso.

### **3.2. Reutilizar o construir componentes de procesamiento y análisis**

60. Este subproceso describe las actividades de reutilización de componentes existentes o de construcción de nuevos componentes necesarios para las fases de procesamiento y análisis, tal como se diseñaron en la fase de diseño. Los servicios pueden incluir funciones y características de cuadros de mando, servicios de información, funciones de transformación, servicios de datos geoespaciales, marcos de flujo de trabajo y servicios de gestión de proveedores y metadatos.

### **3.3. Reutilizar o construir componentes de difusión**

61. Este subproceso describe las actividades para construir nuevos componentes o reutilizar componentes existentes necesarios para la difusión de productos estadísticos según lo diseñado en el subproceso 2.1 (diseño de productos). Se incluyen todos los tipos de componentes de difusión, desde los que producen publicaciones tradicionales en papel hasta los que proporcionan servicios

web, resultados de datos abiertos (vinculados), estadísticas geoespaciales, mapas o acceso a microdatos.

### **3.4. Configuración de flujos de trabajo**

62. Este subproceso configura el flujo de trabajo, los sistemas y las transformaciones utilizadas en los procesos, desde la recolección de datos hasta su difusión. En este subproceso, el flujo de trabajo se configura en base al diseño creado en el subproceso 2.6 (diseño de los sistemas de producción y de los flujos de trabajo). Esto podría incluir la modificación de un flujo de trabajo estandarizado para un propósito específico, el ensamblaje de los flujos de trabajo para las diferentes fases (posiblemente con un sistema de gestión de flujos de trabajo/procesos) y la configuración de los sistemas en consecuencia.

### **3.5. Pruebas al sistema de producción**

63. Este subproceso se refiere a la prueba de servicios ensamblados y configurados y a flujos de trabajo relacionados. Incluye pruebas técnicas y la aprobación de nuevos programas y rutinas, así como la confirmación de que las rutinas existentes de otros procesos estadísticos son adecuadas para su uso en este caso. Si bien parte de esta actividad relativa a la prueba de componentes y servicios individuales podría lógicamente vincularse con los subprocesos 3.1, 3.2 y 3.3, este subproceso también incluye la prueba de las interacciones entre servicios ensamblados y configurados, y la garantía de que toda la solución de producción funciona de manera coherente.

### **3.6. Prueba piloto del proceso estadístico**

64. Este subproceso describe las actividades para gestionar una prueba de campo o piloto del proceso estadístico. Típicamente, incluye una recolección de datos a pequeña escala, para probar los instrumentos de recolección, seguida por el procesamiento y análisis de los datos recolectados, para asegurar que el proceso estadístico funcione como se espera. Después de la prueba piloto, puede ser necesario volver a un paso anterior y hacer ajustes en los instrumentos, sistemas o componentes de recolección. Para un proceso estadístico importante, por ejemplo, un censo de población, puede haber varias iteraciones hasta que el proceso funcione satisfactoriamente.

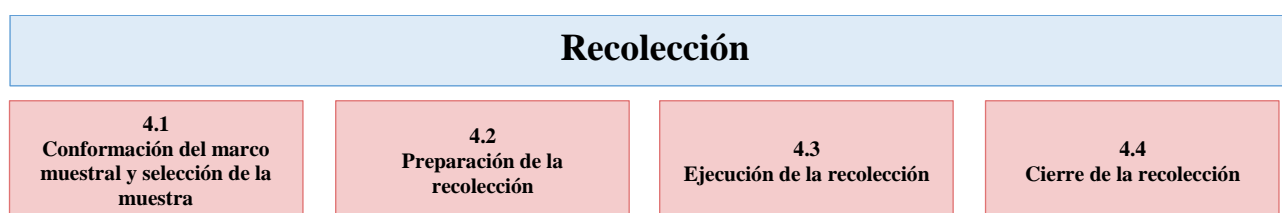
### **3.7. Finalización del sistema de producción**

65. Este subproceso incluye las actividades para poner en producción los procesos y servicios ensamblados y configurados, incluidos los servicios modificados y de nueva creación, listos para su uso. Las actividades incluyen:
- Producir documentación sobre los componentes del proceso, incluyendo documentación técnica y manuales de usuario;
  - Capacitar a los usuarios sobre cómo operar el proceso;
  - Trasladar los componentes del proceso al entorno de producción y asegurarse de que funcionan como se espera en ese entorno (esta actividad también puede formar parte del subproceso 3.5 (pruebas al sistema de producción)).



## Fase de recolección

**Figura 7.** Fase de recolección y sus subprocesos



66. Esta fase recolecta toda la información necesaria (por ejemplo, datos, metadatos y parados), utilizando diferentes modos de recogida (por ejemplo, adquisición, recolección, extracción y transferencia), y los carga en el entorno adecuado para su posterior procesamiento. Aunque puede incluir la validación de formatos de conjuntos de datos, no incluye ninguna transformación de los datos en sí, ya que todos ellos se realizan en la fase de procesamiento. Para los resultados estadísticos producidos regularmente, esta fase se produce en cada iteración.
67. La fase de recolección está conformada por cuatro subprocesos (véase figura 7) que suelen ser secuenciales, de izquierda a derecha, pero que también pueden ocurrir en paralelo y pueden ser iterativos. Estos subprocesos son:

### 4.1. Conformación del marco muestral y selección de la muestra

68. Este subproceso establece el marco y selecciona la muestra para esta iteración de la recolección, como se especifica en el subproceso 2.4 (diseño del marco muestral y del muestreo). También incluye la coordinación de muestras entre instancias del mismo proceso (por ejemplo, para gestionar el solapamiento o la rotación), y entre diferentes procesos utilizando un marco o registro común (por ejemplo, para gestionar el solapamiento o repartir la carga de la respuesta). El aseguramiento de la calidad y la aprobación del marco y de la muestra seleccionada también se llevan a cabo en este subproceso, mediante el mantenimiento de los registros subyacentes, de los que se extraen los marcos de varios procesos estadísticos, se trata como un proceso separado. El aspecto de muestreo de este subproceso no suele ser relevante para los procesos basados exclusivamente en el uso de fuentes preexistentes (por ejemplo, registros administrativos, sitios web), ya que tales procesos generalmente crean marcos a partir de los datos disponibles y luego siguen un enfoque de censo. Las variables de fuentes de datos administrativas y otras fuentes no estadísticas pueden utilizarse como variables auxiliares en la construcción del diseño del muestreo.

### 4.2. Preparación de la recolección

69. Este subproceso garantiza que las personas, los procesos y la tecnología (por ejemplo, las aplicaciones basadas en la web, el sistema GPS) estén preparados para recolectar datos y metadatos, en todos los modos previstos. Se desarrolla a lo largo de un período de tiempo, ya que incluye la estrategia, la planificación y las actividades de capacitación en preparación para el caso específico del proceso estadístico. Cuando el proceso se repite regularmente, algunas (o todas) de estas actividades pueden no ser explícitamente requeridas para cada iteración. En el caso de procesos únicos y nuevos, estas actividades pueden ser muy largas. Para los datos de la encuesta, este subproceso incluye:

- Preparar una estrategia de recolección;

- Capacitar personal de recolección;
- Usar un sistema de capacitación mediante técnicas de aprendizaje automático supervisado;
- Asegurar de que los recursos de recolección estén disponibles (por ejemplo, computadoras portátiles, aplicaciones de recolección, API);
- Acordar las condiciones con cualquier organismo intermedio de recolección (por ejemplo, subcontratistas para entrevistas telefónicas asistidas por computadora, servicios web);
- Configurar sistemas de recolección para solicitar y recibir los datos;
- Garantizar la seguridad de los datos que se van a recolectar;
- Preparar instrumentos de recolección (por ejemplo, imprimir cuestionarios, rellenarlos previamente con los datos existentes, cargar cuestionarios y datos en las computadoras de los entrevistadores, API, herramientas de raspado web);
- Proporcionar información a los encuestados (por ejemplo, redactar cartas o folletos que expliquen el propósito de la encuesta, notificar a los encuestados cuándo se pondrán a su disposición los instrumentos de información en línea);
- Traducir materiales (por ejemplo, a los diferentes idiomas hablados o utilizados en el país).

70. En el caso de las fuentes no encuestadas, este subproceso garantiza la existencia de los procesos, sistemas y procedimientos de confidencialidad necesarios para recibir o extraer la información necesaria de la fuente. Esto incluye:

- Evaluar las solicitudes de adquisición de datos y registro de la solicitud en un inventario centralizado;
- Iniciar contactos con las organizaciones que proporcionan los datos y enviar un paquete introductorio con detalles sobre el proceso de adquisición de los datos;
- Comprobación de información detallada sobre archivos y metadatos con el proveedor de datos y recepción de un archivo de prueba para evaluar si los datos son aptos para su uso;
- Organizar canales seguros para la transmisión de los datos.

### **4.3. Ejecución de la recolección**

71. Este subproceso es donde se implementa la recolección. Los diferentes instrumentos de recolección utilizan para recoger la información que puede incluir microdatos brutos o agregados producidos en la fuente, así como cualquier metadato asociado. Puede incluir el contacto inicial con los proveedores y cualquier acción de seguimiento o recordatorio posterior. Puede incluir la introducción manual de datos en el punto de contacto o la gestión del trabajo en terreno, dependiendo de la fuente y el modo de recolección. Registra cuándo y cómo se contactó a los proveedores, y si han respondido. Dependiendo del marco geográfico y de la tecnología utilizada, puede ser necesario realizar la geocodificación<sup>8</sup> al mismo tiempo que se recolectan los datos, utilizando insumos de los sistemas GPS, colocando una marca en un mapa, etc. Este subproceso también incluye la gestión de los proveedores que participan en la recolección actual, garantizando que la relación entre la organización estadística y los proveedores de datos siga siendo positiva, y registrando y respondiendo a los comentarios, consultas y reclamos. Una comunicación adecuada con las unidades de información y la minimización del número de no encuestados contribuyen significativamente a mejorar la calidad de los datos recolectados.

---

<sup>8</sup> En este documento, la geocodificación se define como “asignación de códigos relacionados con lugares geográficos”, mientras que la georeferenciación se define de forma más general como “datos relacionados con el lugar donde se produce el hecho”.

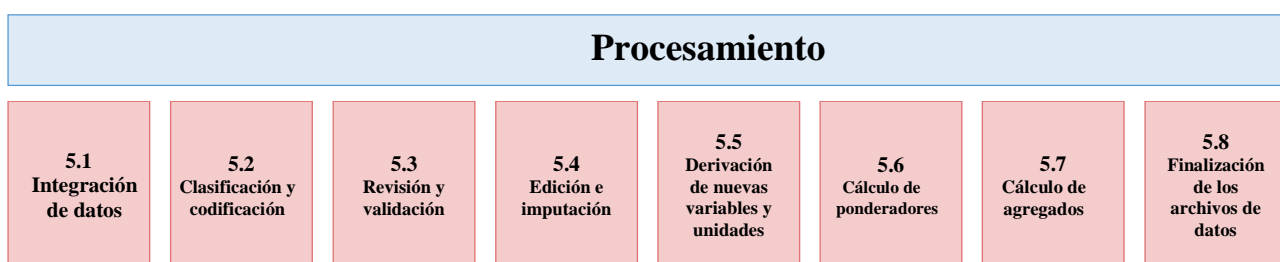
72. En el caso de datos administrativos, geográficos u otros datos no estadísticos, se contacta con el proveedor para que envíe la información o la envía según lo previsto. Este proceso puede llevar mucho tiempo y puede requerir un seguimiento para asegurar que los datos se proporcionen de acuerdo con los acuerdos. En el caso de que los datos se publiquen bajo una licencia Open Data y existan en forma legible por máquinas, se puede acceder a ellos y utilizarlos libremente.
73. Este subproceso también puede incluir el seguimiento de la recolección de datos y la realización de los cambios necesarios para mejorar la calidad de los datos. Esto incluye la generación de informes, la visualización y el ajuste del proceso de adquisición para garantizar que los datos son adecuados para su uso. Cuando la colección cumple sus objetivos, se cierra y se elabora un informe sobre la recolección. En este subproceso pueden tener lugar algunas comprobaciones básicas de la estructura e integridad de la información recibida (por ejemplo, comprobar que los archivos tienen el formato correcto y contienen los campos esperados).

#### 4.4. Cierre de la recolección

74. Este subproceso incluye la carga de los datos y metadatos recolectados en un entorno electrónico adecuado para su posterior procesamiento. Puede incluir la captura manual o automática de datos, por ejemplo, utilizando personal de oficina o herramientas de reconocimiento óptico de caracteres para extraer información de cuestionarios en papel, o convirtiendo los formatos de los archivos o codificando las variables recibidas de otras organizaciones. También puede incluir el análisis de los metadatos y parámetros asociados con la recolección para asegurar que las actividades de recolección han cumplido con los requisitos. En los casos en que existe un instrumento físico de recolección, como un cuestionario en papel, que no es necesario para su procesamiento posterior, este subproceso gestiona el almacenamiento de ese material. Cuando el instrumento de recolección utiliza software como una API o una aplicación, este subproceso también incluye la creación de versiones y el almacenamiento de las mismas.

### Fase de procesamiento

**Figura 8.** Fase de procesamiento y sus subprocesos



75. Esta fase describe el procesamiento de los datos y la preparación para su análisis. Se compone de subprocesos que integran, clasifican, verifican, limpian y transforman los datos, para que puedan ser analizados y difundidos como resultados estadísticos. Para los resultados estadísticos producidos regularmente, esta fase se realiza en cada iteración. Los subprocesos de esta fase pueden aplicarse a los datos de fuentes estadísticas y no estadísticas (con la posible excepción del subproceso 5.6 (cálculo de ponderadores), que suele ser específico de los datos de la encuesta).
76. Las fases de procesamiento y análisis pueden ser iterativas y paralelas. El análisis puede revelar una comprensión más amplia de los datos, lo que podría hacer evidente la necesidad de un

procesamiento adicional. A veces, las estimaciones que se están procesando pueden ser agregados ya publicados (realizados de acuerdo con una política de revisión).

77. Las actividades dentro de las fases de procesamiento y análisis también pueden comenzar antes de que se complete la fase de recolección. Esto permite la compilación de resultados provisionales cuando la puntualidad es una preocupación importante para los usuarios, y aumenta el tiempo disponible para el análisis.
78. La fase de procesamiento se divide en ocho subprocesos (véase figura 8), que pueden ser secuenciales, de izquierda a derecha, pero también pueden ocurrir en paralelo, y pueden ser iterativos. Estos subprocesos lo son:

### **5.1. Integración de datos**

79. Este subproceso integra datos de una o más fuentes. Es donde se combinan los resultados de los subprocesos de la fase de recolección. Los datos pueden provenir de una mezcla de fuentes externas o internas y de una variedad de instrumentos de recolección, incluidos extractos de fuentes de datos administrativas y otras fuentes no estadísticas. Los datos administrativos u otras fuentes de datos no estadísticas pueden sustituir a todas o algunas de las variables recolectadas directamente de la encuesta. Este subproceso también incluye la armonización o la creación de nuevas cifras que coincidan entre las fuentes de datos. El resultado es un conjunto de datos vinculados. La integración de datos puede incluir:

- Combinar datos de múltiples fuentes, como parte de la creación de estadísticas integradas como las cuentas nacionales;
- Combinar datos geospaciales y datos estadísticos u otros datos no estadísticos;
- Poner en común de datos, con el objetivo de aumentar el número efectivo de observaciones de algunos fenómenos;
- Establecer rutinas de enlace de emparejamiento o registro, con el objetivo de enlazar datos micro o macro procedentes de diferentes fuentes;
- Fusión de datos — integración seguida de reducción o sustitución;
- Priorizar, cuando dos o más fuentes contienen datos para la misma variable, con valores potencialmente diferentes.

80. La integración de datos puede realizarse en cualquier momento de esta fase, antes o después de cualquiera de los otros subprocesos. También puede haber varios casos de integración de datos en cualquier proceso estadístico. Tras la integración, dependiendo de los requisitos de protección de datos, los datos pueden ser “des-identificados”, es decir, desprovistos de identificadores como el nombre y la dirección, para ayudar a proteger la confidencialidad.

### **5.2. Clasificación y codificación**

81. Este subproceso clasifica y codifica los datos. Por ejemplo, las rutinas de codificación automática pueden asignar códigos numéricos a las respuestas de texto de acuerdo con una clasificación estadística predeterminada para facilitar la captura y el procesamiento de datos. Algunas preguntas han codificado categorías de respuesta en los cuestionarios o en la fuente administrativa de datos, otras se codifican después de la recolección utilizando un proceso automatizado (que puede aplicar técnicas de aprendizaje automático) o un proceso interactivo y manual.

### 5.3. Revisión y validación

82. Este subproceso examina los datos para identificar problemas potenciales, errores y discrepancias tales como valores atípicos, falta de respuesta de los ítems y codificación errónea. También puede denominarse validación de datos. Se puede ejecutar de forma iterativa, validando los datos contra reglas de edición predefinidas, normalmente en un orden establecido. Puede marcar datos para la inspección o edición automática o manual. La revisión y validación puede aplicarse a los datos de cualquier tipo de fuente, antes y después de la integración, así como a los datos imputados del subproceso 5.4 (edición e imputación). Aunque la validación se trata como parte de la fase de procesamiento, en la práctica, algunos elementos de validación pueden producirse junto con las actividades de recolección, especialmente en el caso de modos como la recolección asistida por computadora. Mientras que este subproceso se ocupa de la detección y localización de errores reales o potenciales, cualquier actividad de corrección que realmente cambie los datos se realiza en el subproceso 5.4 (edición e imputación).

### 5.4. Edición e imputación<sup>9</sup>

83. Cuando los datos se consideren incorrectos, faltantes, poco fiables o anticuados, podrán insertarse nuevos valores o suprimirse datos anticuados en este subproceso. Los términos edición e imputación cubren una variedad de métodos para hacer esto, a menudo utilizando un enfoque basado en reglas. Los pasos específicos típicamente incluyen:

- Determinar si se deben añadir o modificar datos;
- Seleccionar el método a utilizar;
- Añadir/modificar valores de datos;
- Escribir los nuevos valores de datos de vuelta al conjunto de datos y marcarlos como modificados;
- Producir metadatos sobre el proceso de edición e imputación.

### 5.5. Derivación de nuevas variables y unidades

84. Este subproceso deriva los datos de las variables y unidades que no se proporcionan explícitamente en la recolección, pero que son necesarios para obtener los resultados requeridos. El subproceso deriva nuevas variables aplicando fórmulas aritméticas a una o más de las variables que ya están presentes en el conjunto de datos, o aplicando diferentes suposiciones del modelo. Esta actividad puede necesitar ser iterativa, ya que algunas variables derivadas pueden estar basadas en otras variables derivadas. Por lo tanto, es importante asegurarse de que las variables se deriven en el orden correcto. Las nuevas unidades pueden derivarse de la agregación o división de los datos de las unidades de recolección, o de otros métodos de estimación. Los ejemplos incluyen los hogares que derivan, en los que las unidades de recolección son personas, así como las empresas, en las que las unidades de recolección son unidades legales.

### 5.6. Cálculo de ponderadores

85. Este subproceso crea ponderaciones para los registros de datos de las unidades de acuerdo con la metodología desarrollada en el subproceso 2.5 (diseño del procesamiento y análisis). Por ejemplo, las ponderaciones pueden utilizarse para obtener datos “más brutos” para hacerlos representativos

---

<sup>9</sup> Para una consideración más detallada de los diferentes pasos y flujos implicados en la edición de datos, véase UNECE (CEPE) Statistics Wikis - Generic Statistical Data Editing Models (<https://statswiki.unece.org/display/sde/GSDEMs>).

de la población objetivo (por ejemplo, para encuestas por muestreo o extractos de datos de escáneres), o para ajustar la falta de respuesta en las enumeraciones totales. En otras situaciones, puede ser necesario ponderar las variables a efectos de normalización. También puede incluir la corrección de peso para los indicadores de evaluación comparativa (por ejemplo, los totales de población conocidos).

## 5.7. Cálculo de agregados

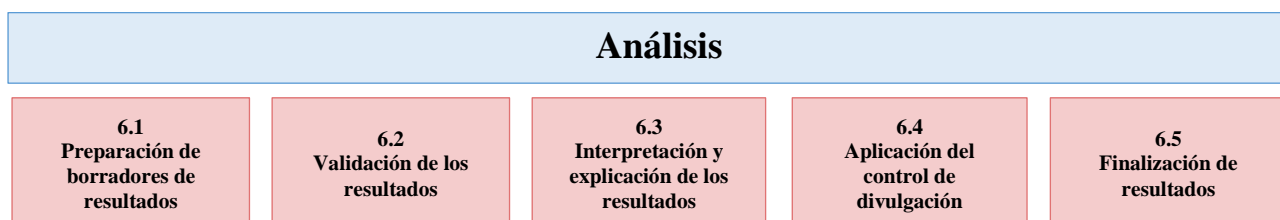
86. Este subproceso crea datos agregados y totales de población a partir de microdatos o agregados de nivel inferior. Incluye la suma de datos para registros que comparten ciertas características (por ejemplo, la agregación de datos por clasificaciones demográficas o geográficas), la determinación de medidas de promedio y dispersión y la aplicación de ponderaciones del subproceso 5.6 (cálculo de ponderadores) para derivar los totales apropiados. En el caso de la producción estadística que utilice encuestas por muestreo, también podrán calcularse en este subproceso los errores de muestreo correspondientes a los agregados pertinentes.

## 5.8. Finalización de los archivos de datos

87. Este subproceso reúne los resultados de los otros subprocesos de esta fase en un archivo de datos (normalmente macro-datos), que se utiliza como insumo para la fase de análisis. A veces se trata de un archivo intermedio en lugar de un archivo final, especialmente en el caso de los procesos en los que existe una fuerte presión de tiempo y la necesidad de elaborar estimaciones tanto preliminares como finales.

## Fase de análisis

**Figura 9.** Fase de análisis y sus subprocesos



88. En esta fase, se elaboran y se examinan detalladamente los resultados estadísticos. Incluye la preparación del contenido estadístico (comentarios, notas técnicas, etc.) y verificación de que los resultados sean “adecuados para el propósito” antes de su difusión a los usuarios. Esta fase también incluye los subprocesos y actividades que permiten a los analistas estadísticos comprender los datos y las estadísticas producidas. Los resultados de esta fase también podrían utilizarse como insumos para otros subprocesos (por ejemplo, el análisis de nuevas fuentes como insumos para la fase de diseño). Para los resultados estadísticos producidos regularmente, esta fase se produce en cada iteración. La fase de análisis y los subprocesos son genéricos para todos los resultados estadísticos, independientemente de cómo se obtuvieron los datos.

89. La fase de análisis se divide en cinco subprocesos (véase figura 9), que generalmente son secuenciales, de izquierda a derecha, pero también pueden ocurrir en paralelo y pueden ser iterativos. Estos subprocesos lo son:

## **6.1. Preparación de borradores de resultados**

90. En este subproceso se transforman los datos de los subprocesos 5.7 (cálculo de agregados) y 5.8 (finalización de los archivos de datos) en resultados estadísticos tales como índices, estadísticas desestacionalizadas (por ejemplo, tendencias, ciclos, componentes estacionales e irregulares, medidas de accesibilidad, etc.), así como en el registro de características de calidad tales como coeficientes de variación. Se puede incluir la preparación de mapas, resultados de GIS y servicios geoestadísticos para maximizar el valor y la capacidad de análisis de la información estadística.

## **6.2. Validación de los resultados**

91. En este subproceso, los estadísticos validan la calidad de los resultados producidos, de acuerdo con un marco general de calidad y con las expectativas. Este subproceso incluye actividades relacionadas con la recopilación de información, con el efecto acumulativo de crear un cuerpo de conocimientos sobre un ámbito estadístico específico. Este conocimiento se aplica luego a la recolección actual, en el entorno actual, para identificar cualquier divergencia con respecto a las expectativas y permitir análisis informados. Las actividades de validación pueden incluir:

- Comprobar que la cobertura de la población y las tasas de respuesta son las necesarias;
- Comparar las estadísticas con ciclos anteriores (si corresponde);
- Comprobar que los metadatos, parámetros e indicadores de calidad asociados están presentes y en línea con las expectativas;
- Comprobar la coherencia geoespacial de los datos;
- Comparar las estadísticas con otros datos relevantes (tanto internos como externos);
- Investigar las incoherencias en las estadísticas;
- Realizar edición de macros;
- Validar las estadísticas frente a las expectativas y la información del entorno.

## **6.3. Interpretación y explicación de los resultados**

92. En este subproceso, los estadísticos adquieren un conocimiento profundo de los resultados. Utilizan esa comprensión para interpretar y explicar las estadísticas evaluando qué tanto reflejan sus expectativas iniciales. Observan las estadísticas desde todas las perspectivas utilizando diferentes herramientas y medios, y llevan a cabo análisis estadísticos rigurosos tales como análisis de series de tiempo, análisis de consistencia y comparabilidad, análisis de revisión (análisis de las diferencias entre las estimaciones preliminares y revisadas), análisis de asimetrías (discrepancias en las estadísticas espejo), etc.

## **6.4. Aplicación del control de divulgación**

93. Este subproceso garantiza que los datos (y metadatos) que se difundan no infrinjan las normas de confidencialidad correspondientes, de acuerdo con las políticas y normas de la organización, o con la metodología específica del proceso creada en el subproceso 2.5 (diseño del procesamiento y análisis). Esto puede incluir comprobaciones de la divulgación primaria y secundaria, así como la aplicación de técnicas de supresión o perturbación de datos y comprobación de los resultados. El grado y el método de control de la divulgación de información estadística pueden variar según los distintos tipos de resultados. Por ejemplo, el enfoque utilizado para los conjuntos de microdatos con fines de investigación será diferente del utilizado para los cuadros publicados, los resultados finalizados de las estadísticas geoespaciales o las visualizaciones en los mapas.

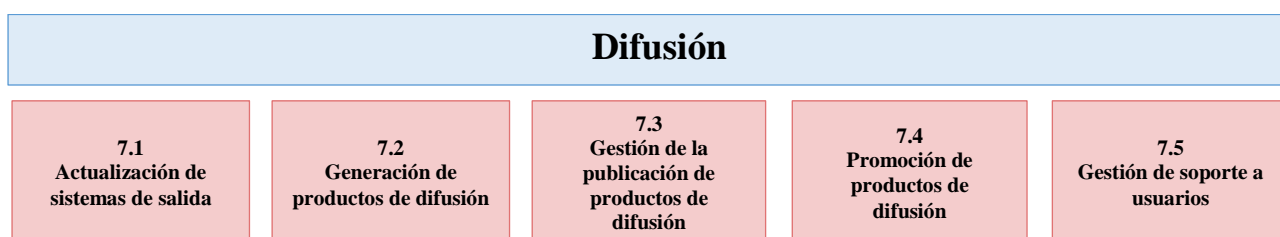
## 6.5. Finalización de resultados

94. Este subproceso garantiza que las estadísticas y la información asociada son adecuadas para el fin previsto y alcanzan el nivel de calidad requerido, por lo que están listas para su uso. Incluye:

- Completar verificaciones de consistencia;
- Determinar el nivel de difusión y aplicación de advertencias;
- Recopilar información de apoyo, incluida la interpretación, los comentarios, las notas técnicas, las sesiones informativas, las medidas de incertidumbre y cualquier otro metadato necesario;
- Elaborar de los documentos internos de apoyo;
- Llevar a cabo una discusión previa a la publicación con los expertos en la materia internos apropiados;
- Traducir los resultados estadísticos de los países con difusión multilingüe;
- Aprobar el contenido estadístico para su publicación.

## Fase de difusión

**Figura 10.** Fase de difusión y sus subprocesos



95. Esta fase gestiona la publicación de los productos estadísticos a los usuarios. Incluye todas las actividades relacionadas con el ensamblaje y la publicación de una gama de productos estáticos y dinámicos a través de una serie de canales. Estas actividades ayudan a los usuarios a acceder a los productos publicados por la organización estadística y a utilizarlos. En el caso de los productos estadísticos elaborados regularmente, esta fase se produce en cada iteración.

96. La fase de difusión se divide en cinco subprocesos (véase figura 10), que generalmente son secuenciales, de izquierda a derecha, pero también pueden ocurrir en paralelo y pueden ser iterativos. Estos subprocesos lo son:

### 7.1. Actualización de sistemas de salida

97. Este subproceso gestiona la actualización de los sistemas (por ejemplo, bases de datos) en los que se almacenan los datos y metadatos listos para su difusión, incluyendo:

- Formatear los datos y metadatos listos para ser introducidos en sistemas de salida;
- Cargar los datos y metadatos en sistemas de salida;
- Asegurarse de que los datos están vinculados a los metadatos relevantes.

98. El formateo, la carga y la vinculación de los metadatos deberían realizarse preferiblemente en fases anteriores, pero este subproceso incluye una comprobación final de que todos los metadatos necesarios están listos para su difusión.



## **7.2. Generación de productos de difusión**

99. Este subproceso elabora los productos de difusión, tal como se diseñó previamente en el subproceso 2.1 (diseño de productos), para satisfacer las necesidades de los usuarios. Podrían incluir publicaciones impresas, comunicados de prensa y sitios web. Los productos pueden tomar muchas formas, incluyendo gráficos interactivos, tablas, mapas, conjuntos de microdatos de uso público, datos abiertos enlazados y archivos descargables. Los pasos típicos incluyen:
- Preparación de los componentes del producto (textos explicativos, tablas, gráficos, mapas, declaraciones de calidad, etc.);
  - Ensamblaje de los componentes en productos;
  - Edición de los productos y comprobación de que cumplen con las normas de publicación.

## **7.3. Gestión de la publicación de productos de difusión**

100. Este subproceso asegura que todos los elementos para la publicación estén en su lugar, incluyendo la gestión de la sincronización de la publicación. Incluye reuniones informativas para grupos específicos, como la prensa o los ministros, así como los arreglos para cualquier embargo previo a la publicación. También incluye el suministro de productos a los usuarios y la gestión del acceso a datos confidenciales por parte de grupos de usuarios autorizados, como los investigadores. A veces una organización puede necesitar retirar un producto, por ejemplo, si se descubre un error; esto también se incluye en este subproceso.

## **7.4. Promoción de productos de difusión**

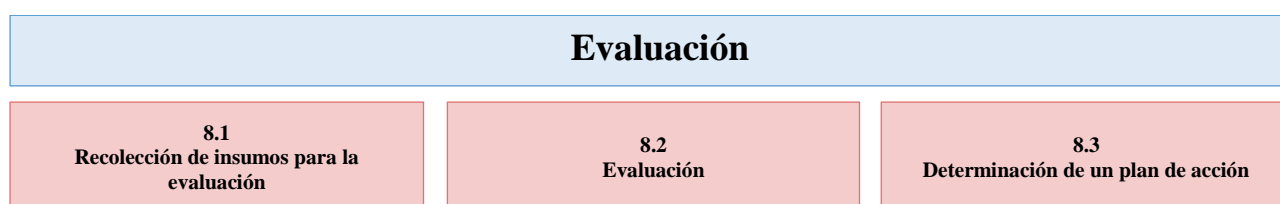
101. Aunque la comercialización en general puede considerarse un proceso global, este subproceso se refiere a la promoción activa de los productos estadísticos elaborados en un proceso estadístico específico, para así ayudarles a llegar a un público más amplio. Incluye el uso de herramientas de gestión de las relaciones con los clientes, para enfocar mejor a los usuarios potenciales de los productos, también incluye el uso de herramientas como sitios web, wikis y blogs para facilitar el proceso de comunicación de información estadística a los usuarios.

## **7.5. Gestión de soporte a usuarios**

102. Este subproceso garantiza que las consultas de los usuarios y las solicitudes de servicios, como el acceso a los microdatos, se registren y que las respuestas se proporcionen dentro de los plazos acordados. Estas consultas y solicitudes deben ser revisadas periódicamente para proporcionar un insumo al proceso general de gestión de calidad, ya que pueden indicar necesidades nuevas o cambios en necesidades de los usuarios. Las respuestas a las solicitudes de los usuarios también pueden utilizarse para completar una base de datos de conocimientos o una página de “preguntas más frecuentes”, que se pone a disposición del público, lo que reduce la carga de responder a solicitudes repetidas y/o similares de usuarios externos. Este subproceso también incluye la gestión del apoyo a las organizaciones asociadas que participan en la difusión de los productos.

## Fase de evaluación

**Figura 11.** Fase de evaluación y sus subprocesos



103. Esta fase gestiona la evaluación de un caso específico de un proceso estadístico, en contraposición con el proceso general de gestión de la calidad estadística descrito en la Sección VI (Procesos Globales). Puede ocurrir al final de la instancia del proceso, pero también puede realizarse de forma continua durante el proceso de producción estadística. Se basa en los datos recolectados a lo largo de las diferentes fases. Incluye la evaluación del éxito de un ejemplo específico del proceso estadístico, la utilización de una serie de datos cuantitativos y cualitativos, y la identificación y priorización de las posibles mejoras.
104. En el caso de los resultados estadísticos producidos regularmente, la evaluación debe realizarse, al menos en teoría, para cada iteración, determinando si se deben realizar iteraciones futuras y, en caso afirmativo, si se debe implementar alguna mejora. Sin embargo, en algunos casos, en particular en el caso de los procesos operativos estadísticos regulares y bien establecidos, es posible que la evaluación no se lleve a cabo formalmente para cada iteración. En tales casos, se puede considerar que esta fase proporciona la decisión de si la siguiente iteración debe comenzar desde la fase de especificación de necesidades o desde alguna fase posterior (a menudo la fase de recolección).
105. La fase de evaluación se divide en tres subprocesos (véase figura 11), que generalmente son secuenciales, de izquierda a derecha, pero también pueden ocurrir en paralelo y pueden ser iterativos. Estos subprocesos lo son:

### 8.1. Recolección de insumos para la evaluación

106. El material de evaluación puede producirse en cualquier otra fase o subproceso. Puede tomar muchas formas, incluyendo la retroalimentación de los usuarios, los metadatos del proceso (paradatos), las métricas del sistema y las sugerencias del personal. Los informes sobre los progresos realizados en relación con un plan de acción acordado durante una iteración anterior también pueden constituir un insumo a las evaluaciones de las iteraciones subsiguientes. Este subproceso reúne todos estos insumos, recopila indicadores de calidad y los pone a disposición de la persona o equipo que realiza la evaluación. La recopilación de algunos de estos materiales de evaluación puede automatizarse y llevarse a cabo de forma continua a lo largo de todo el proceso, tal y como se define en el marco de calidad (véase Gestión de Calidad en la Sección VI). Por otra parte, para la evaluación de determinados procesos puede ser necesario realizar actividades específicas, como pequeñas encuestas (por ejemplo, encuestas de post-enumeración, estudios de re-entrevistas, encuestas sobre la eficacia de la difusión).

## 8.2. Evaluación

107. Este subproceso analiza los insumos de la evaluación, los compara con los resultados de referencia esperados/objetivos (cuando estén disponibles) y los sintetiza en un informe de evaluación o tablero de control. La evaluación puede ocurrir al final de todo el proceso (evaluación ex post) de las actividades seleccionadas, durante su ejecución de forma continua o a lo largo de todo el proceso, permitiendo así soluciones rápidas o mejoras continuas. El informe resultante debe anotar cualquier problema de calidad específico de esta iteración del proceso de estadístico, así como destacar cualquier desviación de las métricas de rendimiento con respecto a los valores esperados, y debe hacer recomendaciones de cambios si procede. Estas recomendaciones pueden abarcar cambios en cualquier fase o subproceso para futuras iteraciones del proceso, o pueden sugerir que el proceso no se repite.

## 8.3. Determinación de un plan de acción

108. Este subproceso reúne el poder de decisión necesario para elaborar y acordar un plan de acción basado en el informe de evaluación. También debe incluir el examen de un mecanismo para supervisar los efectos de esas medidas, que a su vez puede servir de insumo para las evaluaciones de las futuras iteraciones del proceso.

## VI. Procesos globales

109. El GSBPM reconoce varios procesos globales que se aplican en todas las fases de producción y en todos los procesos estadísticos. Algunos de estos procesos generales se enumeran en la Sección II. Los procesos de gestión de calidad, gestión de metadatos y gestión de datos se detallan en esta sección.

### Gestión de Calidad

110. La calidad afecta a organizaciones, productos, fuentes y procesos. En el marco actual, el proceso global de gestión de la calidad se refiere a la calidad del producto y del proceso. La calidad a nivel institucional (por ejemplo, la adopción de una política de calidad o de un marco de garantía de calidad) se considera en la GAMSQ.

111. El objetivo principal de la gestión de la calidad dentro del proceso estadístico es comprender y gestionar la calidad de las fuentes, procesos y productos estadísticos. Los organismos estadísticos coinciden en que la calidad debe definirse de acuerdo con la norma ISO 9000-2015: “El grado en el que un conjunto de características inherentes de un objeto cumple con los requisitos”<sup>10</sup>. Por lo tanto, la calidad es un concepto complejo y multifacético, generalmente definido en términos de varias dimensiones de calidad. Las dimensiones de la calidad que se consideran más importantes dependen de las perspectivas, necesidades y prioridades de los usuarios, que varían entre procesos y entre grupos de usuarios.

112. Para mejorar la calidad, la gestión de la calidad debe estar presente a lo largo de todo el modelo de proceso. Está estrechamente ligada a la fase de evaluación. Sin embargo, la gestión de calidad tiene un alcance más profundo y más amplio. Además de evaluar las iteraciones de un proceso,

---

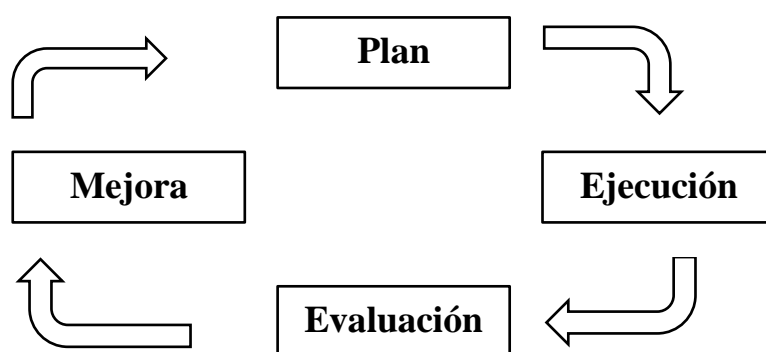
<sup>10</sup> ISO 9000:2015, Sistemas de gestión de la calidad — Fundamentos y vocabulario. Organización Internacional de Normalización (ISO).

también es necesario evaluar fases y subprocesos separados, idealmente cada vez que se aplican, pero al menos según un calendario acordado. Los metadatos generados por los diferentes subprocesos también son de interés como insumo para la gestión de la calidad de los procesos. Estas evaluaciones pueden aplicarse dentro de un proceso específico, o a través de varios procesos que utilizan componentes comunes.

Además, un rol fundamental en la gestión de calidad lo desempeña el conjunto de acciones de control de calidad que se deben implementar dentro de los subprocesos para prevenir y monitorear errores y fuentes de riesgo. Éstos deben estar documentados y pueden utilizarse para la elaboración de informes de calidad.

113. Dentro de una organización, la gestión de calidad suele referirse a un marco de calidad específico y, por lo tanto, puede adoptar diferentes formas y producir diferentes resultados dentro de las distintas organizaciones. La multiplicidad actual de marcos de calidad aumenta la importancia de los enfoques de evaluación comparativa y de revisión por pares, y aunque es poco probable que estos enfoques sean viables para cada iteración de cada parte de cada proceso estadístico de la empresa, deberían utilizarse de manera sistemática de acuerdo con un calendario predeterminado que permita la revisión de todas las partes principales del proceso dentro de un período de tiempo específico<sup>11</sup>.
114. Ampliando el campo de aplicación del proceso global de gestión de calidad, también se puede considerar la evaluación de grupos de procesos estadísticos, con el fin de identificar posibles duplicaciones o vacíos.
115. Todas las evaluaciones dan como resultado retroalimentación, la cual debe ser utilizada para mejorar el proceso, fase o subproceso relevante, creando un círculo de calidad que refuerce el enfoque de las mejoras continuas y el aprendizaje organizacional.

**Figura 12.** Círculo de calidad



116. Algunos ejemplos de actividades de gestión de calidad incluyen:

- Evaluar los riesgos e implementar tratamientos de riesgo para asegurar una calidad adecuada para el propósito;
- Establecer los criterios de calidad a utilizar en el proceso;

<sup>11</sup> Un marco mundial adecuado es el marco nacional de control de calidad elaborado por un grupo mundial de expertos de la Comisión de Estadística de las Naciones Unidas. (<http://unstats.un.org/unsd/dnss/QualityNOAF/nqaf.aspx>)

- Establecer los objetivos de calidad de los procesos y monitorear el cumplimiento de los mismos;
- Buscar y analizar la retroalimentación de los usuarios;
- Revisar las operaciones y documentar las lecciones aprendidas;
- Examinar los metadatos del proceso y los indicadores de calidad;
- Usar auditoría interna o externa sobre el proceso.

117. Los indicadores de calidad apoyan una gestión de calidad orientada al proceso. En cuanto a las estadísticas derivadas de las encuestas y de las fuentes de datos administrativos, una lista sugerida de indicadores de calidad para las fases y subprocesos del GSBPM, así como para los procesos generales de gestión de la calidad y los metadatos, puede encontrarse en el sitio web de Indicadores de Calidad para el GSBPM<sup>12</sup>. Entre otras cosas, pueden utilizarse como lista de control para identificar vacíos y/o duplicación de trabajo en la organización.

### Gestión de Metadatos

118. Los metadatos tienen un rol importante y deben ser gestionados a nivel operativo dentro del proceso de producción estadística. Cuando los aspectos de la gestión de metadatos se consideran a nivel corporativo o estratégico (por ejemplo, existen sistemas de metadatos que afectan a grandes partes del sistema de producción), deben considerarse en el marco de la GAMSO.

119. Una buena gestión de los metadatos es esencial para el funcionamiento eficaz de los procesos estadísticos. Los metadatos están presentes en todas las fases, ya sean creados, actualizados o arrastrados desde una fase anterior o reutilizados desde otro proceso estadístico. En el contexto de este modelo, el énfasis del proceso global de gestión de metadatos se centra en la creación/revisión, actualización, uso y almacenamiento de metadatos estadísticos, aunque también son de interés los metadatos sobre los diferentes subprocesos, incluso como insumos a la gestión de la calidad. El reto clave es garantizar que estos metadatos se capturen lo antes posible, se almacenen y se transfieran de fase a fase junto con los datos a los que se refieren. Por lo tanto, la estrategia y los sistemas de gestión de metadatos son vitales para el funcionamiento de este modelo, y pueden ser facilitados por el GSIM.

120. El GSIM es un marco de referencia de objetos de información, que permite descripciones genéricas de la definición, gestión y uso de datos y metadatos a lo largo del proceso de producción estadística. El GSIM apoya un enfoque coherente de los metadatos, facilitando la función principal de los metadatos, es decir, que los metadatos deben definir de manera única y formalmente el contenido y los vínculos entre los objetos y procesos de información en el sistema de información estadística.

121. El Marco Común de Metadatos de METIS identifica los siguientes dieciséis principios básicos para la gestión de metadatos, todos los cuales están destinados a ser abordados en el proceso general de gestión de metadatos, y tomados en consideración al diseñar e implementar un sistema de metadatos estadísticos. Los principios se presentan en cuatro grupos:

---

<sup>12</sup> UNECE (CEPE) Statistics Wikis - Quality Indicators for the GSBPM  
(<https://statswiki.unece.org/display/GSBPM/Quality+Indicators>)

Manejo de los Metadatos	<ul style="list-style-type: none"> <li>i. Modelo genérico de proceso estadístico: Gestione los metadatos enfocado al modelo genérico de proceso estadístico.</li> <li>ii. Activo no pasivo: Active los metadatos lo más que se pueda. Los metadatos activos son metadatos que dirigen otros procesos y acciones. El manejar los metadatos de esta forma asegura que sean precisos y se mantengan actualizados.</li> <li>iii. Reutilización: Reutilice los metadatos donde sea posible para llevar a cabo la integración estadística, así como por motivos de eficiencia.</li> <li>iv. Versiones: Conserve la historia de metadatos (versiones anteriores).</li> </ul>
Autoridades de los Metadatos	<ul style="list-style-type: none"> <li>i. Registro: Asegúrese de documentar el proceso de registro (flujo de trabajo) asociado con cada elemento de los metadatos para que se pueda identificar claramente la propiedad, estatus de aprobación, fecha de operación, etc.</li> <li>ii. Fuente única: Asegúrese de que exista una sola fuente de autoridad (“autoridad de registro”) para cada uno de los elementos de los metadatos existentes.</li> <li>iii. Un insumo/actualización: Reduzca errores entrando una sola vez y actualice en un solo lugar.</li> <li>iv. Variación de las normas: Asegúrese de que la variación de las normas esté firmemente gestionadas o aprobadas, documentadas y visibles.</li> </ul>
Relación con el Ciclo Estadístico / Proceso	<ul style="list-style-type: none"> <li>i. Integridad: Haga el trabajo relacionado con los metadatos una parte integral del proceso transversal a lo largo de la organización.</li> <li>ii. Emparejamiento de los metadatos: Asegúrese que los metadatos presentados a los usuarios finales coincidan con los metadatos que dirigieron el proceso transversal o que fueron creados en el proceso.</li> <li>iii. Describir el flujo: Describa el flujo de los metadatos con los procesos estadísticos y globales (de la mano con el flujo de datos y la lógica de negocio).</li> <li>iv. Capturar en la fuente: Capture los metadatos en su fuente, preferentemente de forma automática como un subproducto de otros procesos.</li> <li>v. Intercambio y uso: Intercambie metadatos y utilícelos para informar a los procesos basados en computadora y la interpretación humana. La infraestructura para el intercambio de datos y los metadatos que se asocian a ella deben estar basados en componentes ligeramente acoplados con opciones estándar para el cambio de lenguaje como XML.</li> </ul>
Usuarios	<ul style="list-style-type: none"> <li>i. Identificar al usuario: Asegúrese que los usuarios estén claramente identificados para todos los procesos de metadatos.</li> <li>ii. Diferentes formatos: La diversidad de los metadatos se reconoce y hay diferentes enfoques correspondientes a los diferentes usos que se le puedan dar a los datos. Diferentes usuarios requieren diferentes niveles de detalle. Los metadatos aparecen en diferentes formatos dependiendo del proceso y objetivos para los que se generan y utilizan.</li> </ul>

	iii. Disponibilidad: Asegúrese que los metadatos estén fácilmente disponibles y sean utilizables en el contexto de las necesidades de los usuarios (ya sea un usuario interno o externo).
--	---

## Gestión de Datos

122. La gestión de datos es esencial, ya que los datos se producen dentro de muchas de las actividades del proceso estadístico y son los resultados clave. El objetivo principal de la gestión de datos es garantizar que los datos se utilicen de forma adecuada y sean utilizables a lo largo de su ciclo de vida. La gestión de datos a lo largo de su ciclo de vida abarca actividades como la planificación y evaluación de los procesos de gestión de datos, así como el establecimiento y la implementación de procesos relacionados con la recolección, organización, uso, protección, conservación y eliminación de los datos.
123. La forma en que se gestionan los datos estará estrechamente vinculada al uso de los datos, que a su vez está vinculado al proceso estadístico en el que se crean los datos. Tanto los datos como los procesos en los que se crean deben estar bien definidos para garantizar una gestión adecuada de los datos.
124. Algunos ejemplos de actividades de gestión de datos incluyen:
- Establecer una estructura de administración y asignar responsabilidades de gestión de datos;
  - Diseñar estructuras de datos y conjuntos de datos asociados, y el flujo de datos a través del proceso estadístico;
  - Identificar la base de datos (repositorios) para almacenar los datos y administrar la base de datos;
  - Documentar los datos (por ejemplo, registro e inventario de datos, clasificación de los datos según su contenido, retención u otra clasificación requerida);
  - Determinar los períodos de retención de datos;
  - Proteger los datos contra el acceso y uso no autorizados;
  - Proteger los datos contra el cambio tecnológico, la degradación de los medios físicos y la corrupción de los datos;
  - Realizar comprobaciones de la integridad de los datos (por ejemplo, comprobaciones periódicas que garanticen la precisión y coherencia de los datos a lo largo de todo su ciclo de vida);
  - Realizar actividades de disposición una vez que el período de retención de los datos haya expirado.

## VII. Otros usos del GSBPM

125. El objetivo original de la GSBPM era proporcionar una base para que las organizaciones estadísticas acordaran una terminología normalizada para los debates sobre el desarrollo de sistemas y procesos de metadatos estadísticos. Sin embargo, a medida que el modelo ha evolucionado, su uso ha aumentado para muchos otros fines, en particular, relacionados con la modernización de las estadísticas oficiales. En la plataforma wiki de la CEPE se encuentran

disponibles varios documentos que describen los usos reales y potenciales de la GSBPM<sup>13</sup>. La lista que figura a continuación tiene por objeto destacar algunos de los usos actuales e inspirar nuevas ideas sobre la forma en que se puede utilizar el GSBPM en la práctica.

- Proporcionar una estructura para la documentación de los procesos estadísticos — El GSBPM puede proporcionar una estructura para organizar y almacenar la documentación dentro de una organización, promoviendo la normalización y la identificación de buenas prácticas;
- Facilitar la difusión de métodos y programas informáticos estadísticos — El GSBPM define los componentes de los procesos estadísticos de manera que no sólo fomente la difusión de métodos y herramientas informáticas entre los procesos estadísticos, sino que también facilite la difusión entre las diferentes organizaciones estadísticas que aplican el modelo;
- Describir qué normas se utilizan o podrían utilizarse en las diferentes fases del proceso de producción estadística. Por ejemplo, el Anexo 2 de la Guía del Usuario de SDMX 2.1<sup>14</sup> explora cómo se aplica SDMX al trabajo estadístico en el contexto de un modelo de procesos;
- Proporcionar un marco para la evaluación y mejora de la calidad de los procesos — Para que un enfoque de evaluación comparativa de la calidad de los procesos tenga éxito, es necesario normalizar los procesos en la medida de lo posible. El GSBPM proporciona un mecanismo para facilitar esto;
- En relación con el punto anterior, el marco común proporcionado por el GSBPM puede ayudar a integrar el trabajo internacional sobre metadatos estadísticos con el de calidad de los datos, proporcionando un marco común y una terminología común para describir el proceso estadístico;
- Proporcionar el modelo subyacente para los marcos de normas metodológicas — Las normas metodológicas pueden vincularse a la(s) fase(s) o subproceso(s) con los que se relacionan y luego pueden clasificarse y almacenarse en una estructura basada en el GSBPM;
- Desarrollar un repositorio de modelos de procesos para almacenar los resultados del modelado de procesos y permitir su vinculación con el modelo estadístico de procesos;
- Proporcionar un modelo subyacente para diseñar un marco de capacitación y un conjunto de terminología normalizada para describir las habilidades y conocimientos necesarios en el proceso de producción estadística;
- Medir los costes operativos — El GSBPM puede utilizarse como base para medir los costes de las diferentes partes del proceso estadístico. Esto ayuda a orientar las actividades de modernización para mejorar la eficiencia de las partes del proceso que son más costosas;
- Medir el rendimiento del sistema — En relación con el punto anterior sobre costes, el GSBPM también puede utilizarse para identificar los componentes que no funcionan de forma eficiente, que se están duplicando innecesariamente o que requieren ser reemplazados. Del mismo modo, puede identificar los vacíos para las que deben desarrollarse nuevos componentes;
- Proporcionar una herramienta para alinear los procesos de los proveedores de datos no estadísticos (por ejemplo, datos administrativos y datos geoespaciales), facilitar la comunicación entre estadísticos y expertos de otros ámbitos y armonizar la terminología relacionada;
- Proporcionar una herramienta para construir capacidades y conocimientos técnicos metódicamente, haciendo referencia a los detalles de cada fase;
- Proporcionar una herramienta para el desarrollo/revisión de las clasificaciones estadísticas

---

<sup>13</sup> UNECE (CEPE) Statistics Wikis - Uses of the GSBPM  
(<https://statswiki.unece.org/display/GSBPM/Uses+of+GSBPM>).

<sup>14</sup> SDMX community ([http://sdmx.org/index.php?page\\_id=38](http://sdmx.org/index.php?page_id=38))



## Anexo – Lista de acrónimos

**API** – Interfaz de programación de aplicaciones (inglés: *Application Programming Interface*)

**CEPE** – Comisión Económica de las Naciones Unidas para Europa (inglés: UNECE: United Nations Economic Commission for Europe).

**GAMSO** – Modelo genérico de actividad de los organismos estadísticos (inglés: *Generic Activity Model for Statistical Organizations*): Un marco de referencia que describa y defina las actividades que tienen lugar dentro de una organización estadística típica.

**GIS** – Sistema de Información Geográfica (inglés: *Geographic Information System*)

**GPS** – Sistema de Posicionamiento Global (inglés: *Global Positioning System*)

**GSBPM** – Modelo Genérico del Proceso Estadístico (inglés: *Generic Statistical Business Process Model*): Una herramienta flexible para describir y determinar un conjunto de procesos necesarios para elaborar estadísticas oficiales.

**GSIM** – Modelo Genérico de Información Estadística (inglés: *Generic Statistical Information Model*): Un marco de referencia de objetos de información, que permite descripciones genéricas de la definición, gestión y uso de datos y metadatos a lo largo del proceso de producción estadística.

**HLG-MOS** – El Grupo de Alto Nivel para la Modernización de la Producción y Servicios Estadísticos (inglés: *The High-Level Group for the Modernization of Statistical Production and Services*).

**METIS** – El “nombre de marca” para el trabajo sobre Metadatos Estadísticos bajo la Conferencia de Estadísticos Europeos.

**OCDE** – La Organización para la Cooperación y el Desarrollo Económicos

**SDMX** – Intercambio de Datos y Metadatos Estadísticos (inglés: *Statistical Data and Metadata eXchange*): Un conjunto de normas técnicas y guías orientadas al contenido de la mano con arquitectura y herramientas de TI a utilizarse para hacer eficiente el intercambio y participación de datos y metadatos estadísticos.

**TIC** – Tecnologías de la información y la comunicación

**XML** – Lenguaje electrónico extensible a beneficio (inglés: *eXtensible Mark-up Language*): Un lenguaje que define un conjunto de reglas para codificar documentos en un formato que sea legible para un humano y una máquina.