

Modelo del Proceso Estadístico y Geográfico (MPEG)

Sistema Nacional de Información Estadística y Geográfica (SNIEG)

Versión 1.0, 28 de abril de 2017

Resumen

Este documento describe el Modelo de Proceso Estadístico y Geográfico del SNIEG (MPEG). El MPEG está basado en el Modelo Genérico del Proceso Estadístico (GSBPM, por sus siglas en inglés) desarrollado por la Comisión Económica de las Naciones Unidas para Europa (UNECE). Este modelo describe las fases del proceso a través de las cuales se transforman los datos estadísticos y geográficos en información relevante para los usuarios. Asimismo, se describen los procesos transversales que coadyuvan a una adecuada ejecución del proceso principal como son la gestión de calidad, la gestión de datos y la gestión de usuarios. Finalmente se plantean varios esquemas de adopción.



Este trabajo está bajo la licencia “Creative Commons Atribución 3.0.”. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by/3.0/>. Si utiliza la totalidad o parte de este trabajo, por favor, atribúyalo a la Comisión Económica de las Naciones Unidas para Europa (UNECE), en nombre de la comunidad estadística internacional.

Contenido

I. Introducción	4
II. Antecedentes	5
III. Estructura y descomposición	5
IV. Fases y subprocesos	9
1. Fase Especificación de necesidades	10
2. Fase Diseño	12
3. Fase Construcción	14
4. Fase Recolección	16
5. Fase Procesamiento	18
6. Fase Análisis	20
7. Fase Difusión	22
8. Fase Evaluación	23
V. Procesos Transversales con componente estadístico o geográfico	24
1. Gestión de la calidad	24
2. Gestión de documentación	25
3. Gestión de datos	25
4. Gestión de metadatos	26
5. Gestión de parados (datos de proceso)	26
6. Gestión del conocimiento	26
7. Gestión del marco geoestadístico	27
8. Gestión del programa geoestadístico	27
9. Gestión de informantes	27
10. Gestión de usuarios	27
11. Gestión de tomadores de decisiones	27
VI. Procesos transversales sin componente estadístico o geográfico	27
1. Planeación estratégica	27
2. Administración del marco legal	28
3. Administración financiera	28
4. Administración de proyectos.	28
5. Administración de recursos humanos.	28
6. Armonización de Arquitectura de Aplicaciones (software)	28

7. Administración organizacional.	28
VII. Esquemas de adopción del Proceso Sustantivo	29
1. Adopción secuencial	29
2. Adopción por niveles	29
3. Adopción respecto al tipo de datos origen	30
4. Adopción por subconjuntos de fases	31
5. Adopción no secuencial por subprocesos	31
6. Adopción regional	32
VIII. Adopción de los procesos transversales	32

I. Introducción

Las oficinas nacionales de estadística y geografía han tenido una preocupación constante por mejorar la eficiencia de sus actividades y la calidad de sus productos de información. Históricamente, muchas de las mejoras giraban alrededor de la automatización de actividades teniendo a las tecnologías de la información como apoyo central para cumplir estos objetivos. Sin embargo, el esquema de la simple automatización tiene limitaciones. Por un lado, en varios proyectos se observa una herencia de sistemas legados que en el mejor de los casos permite una mejora marginal del producto de información. En este enfoque de proyecto aislado no se facilita la reutilización y por lo mismo se incurre en costos innecesarios generados por la multiplicidad de procesos y sistemas diferentes. En otras palabras, no se procuran mecanismos estandarizados que faciliten la generación de indicadores y permitan generar economías de escala.

El conjunto de países que colaboran en la Comisión Económica de las Naciones Unidas para Europa (UNECE), ha venido trabajando para compensar esta limitación, generando marcos de procesos estándares que faciliten la colaboración y la reutilización de procesos, datos y sistemas estadísticos. Uno de estos marcos es el Modelo Genérico del Proceso Estadístico (GSBPM, por sus siglas en inglés). El GSBPM plantea la posibilidad de usar un modelo estandarizado para la cadena de valor que sirva como marco conceptual de referencia y facilite la ejecución de actividades con una orientación a procesos. Asimismo, el trabajo de la UNECE ha permitido integrar en el mismo Modelo las mejores prácticas para la generación de estadísticas oficiales.

En este documento se describe el Modelo de Proceso Estadístico y Geográfico (MPEG). En su estructura conceptual está basado en el GSBPM. Sin embargo, esta versión trata de ofrecer un enfoque igualmente conceptual, pero con lenguaje más simple y orientado al productor de información estadística y geográfica del SNIEG. Asimismo, se limita a la descomposición del Modelo y a las estrategias de implantación.

El MPEG se puede descomponer el proceso sustantivo en tres niveles. El Nivel 0 conceptualiza el proceso de producción de información a alto nivel en un contexto entrada-proceso-salida. El Nivel 1 identifica la cadena de valor, o ciclo estadístico, en 8 fases para la producción de información, es decir, descompone el Nivel 0 en 8 componentes. El nivel 2, a su vez, descompone cada una de las fases del nivel 1 en subprocesos, en total se identifican 44 subprocesos en las 8 fases. Asimismo, el modelo propone una serie de procesos transversales que coadyuvan a la gestión del proceso sustantivo.

Además, el MPEG provee un marco conceptual con una terminología armonizada de tal forma que se facilite:

- Documentar las características de los procesos de manera consistente.
- Estandarizar datos y metadatos.
- Compartir métodos y componentes.
- Armonizar infraestructuras de aplicaciones de software.
- Proveer un marco de referencia para mejoras y evaluaciones del proceso de calidad.

En primera instancia, el MPEG podría parecer muy amplio. Sin embargo, su adopción puede hacerse de forma gradual, a través de iteraciones se puedan agregar elementos hasta llegar a su total adopción.

II. Antecedentes

Las sesiones de trabajo conjuntas entre UNECE/Eurostat/OCDE han preparado un marco común¹ del Ciclo Estadístico, también conocido como cadena de valor estadística, que descompone las diversas fases del proceso y proporciona los términos genéricos para describirlas.

El Generic Statistical Business Process Model (GSBPM), en el cual está basado el MPEG, fue un desarrollo basado en el modelo de proceso genérico de la Oficina Nacional de Estadísticas de Nueva Zelanda y complementado con aportaciones de otros organismos internacionales. Desde noviembre de 2013, este trabajo ha sido llevado a cabo por el Comité de Modernización de Estándares, bajo el Grupo de Alto Nivel para la Modernización de la Producción Estadística y Servicios (High Level Group – HLG)². La versión original del GSBPM se puede consultar en la página web de la UNECE (www.unece.org/stats/GSBPM).

En este documento se desarrolla una versión de Modelo Genérico que toma como referencia el trabajo desarrollado para el GSBPM. Se considera sustancialmente el contenido del Apartado IV, pero se modifican las secciones introductorias y de implantación con el propósito de tener una versión más cercana a las necesidades de los responsables de gestión y operación de los procesos estadísticos y geográficos del SNIEG. Por otro lado, se considera que el MPEG es un modelo de producción lo suficientemente general que se puede aprovechar para enmarcar la producción de información de diferentes tipos. En este caso, en el SNIEG se decidió adoptar este modelo para información estadística y geográfica.

III. Estructura y descomposición

El MPEG tiene una estructura que se puede descomponer de manera esquemática en tres niveles. El Nivel 0, es una abstracción en la que el modelo se considera una caja negra que integra los insumos, el proceso y la generación de productos estadísticos y geográficos relevantes. En la Figura 1, se ilustra esta abstracción, las fases y subprocesos se considerarían de manera implícita. La relevancia de este esquema es que propone un enfoque orientado a procesos en el que se pueden adoptar otros conceptos de procesos como son: roles, actores, controles, entregables e indicadores de gestión que estén especificados en la versión detallada del Modelo. Esto es, aunque no se maneje de manera explícita cada subproceso identificado en las fases, de manera explícita se puede cumplir con las especificaciones de responsabilidad, monitoreo y control.

¹ Ver: <http://www.unece.org/stats/cmfi>

² Ver: <http://www1.unece.org/stat/platform/display/hlgbas>



Figura 1

El Modelo General se descompone en niveles. La convención adoptada por el MPEG es descomponerlo en 2 niveles. En el primer nivel se identifican las fases. En el segundo nivel se descomponen las fases en subprocesos.

La descomposición del Nivel 0 al Nivel 1 implica la división del proceso global en 8 fases. Cada una de estas tiene la conceptualización general de un proceso con sus respectivos insumos, sus procesos de transformación y sus productos como entregables determinados. La Figura 2 ilustra la descomposición en 8 fases.



Figura 2

El siguiente nivel se forma al separar las 8 fases que componen el Nivel 1 en 44 subprocesos que forman el Nivel 2 (figura 3). El apartado IV describe a detalle el Nivel 2.

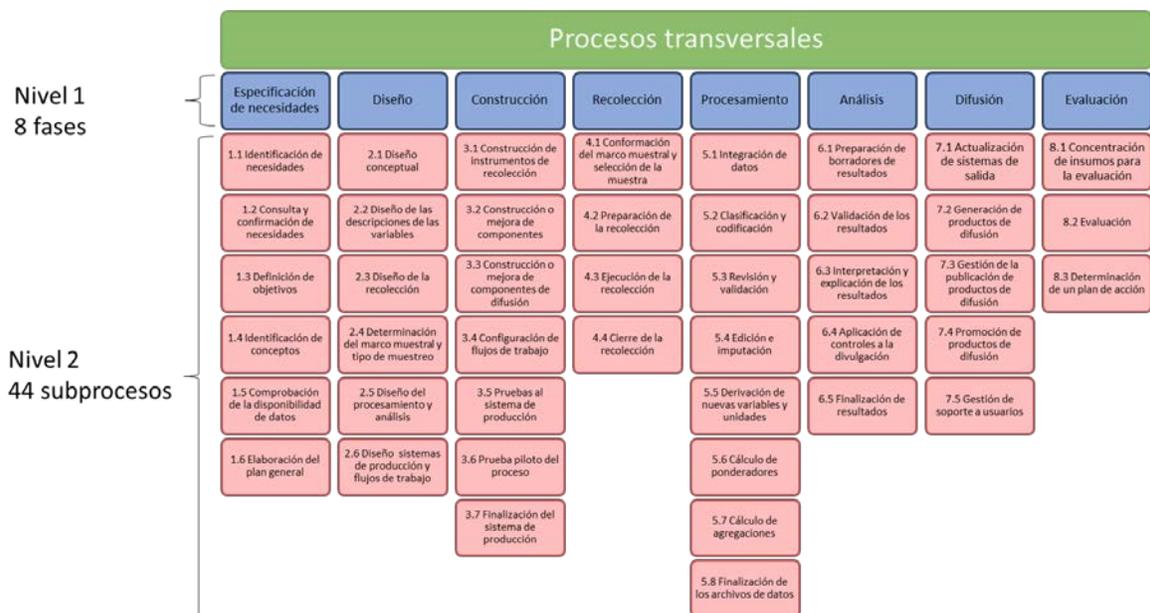


Figura 3

Adicional a los dos niveles de descomposición, en la misma Figura 3 se ilustran los procesos transversales complementarios a las funciones estadísticas y geográficas. Los procesos transversales se aplican a lo largo de las 8 fases y a través de los 44 subprocesos. Los procesos transversales se agrupan en dos categorías: los que apoyan funciones estadísticas o geográficas y los procesos generales de soporte que podrían ser aplicados a cualquier tipo de organización. Los procesos con un componente estadístico o geográfico incluyen:

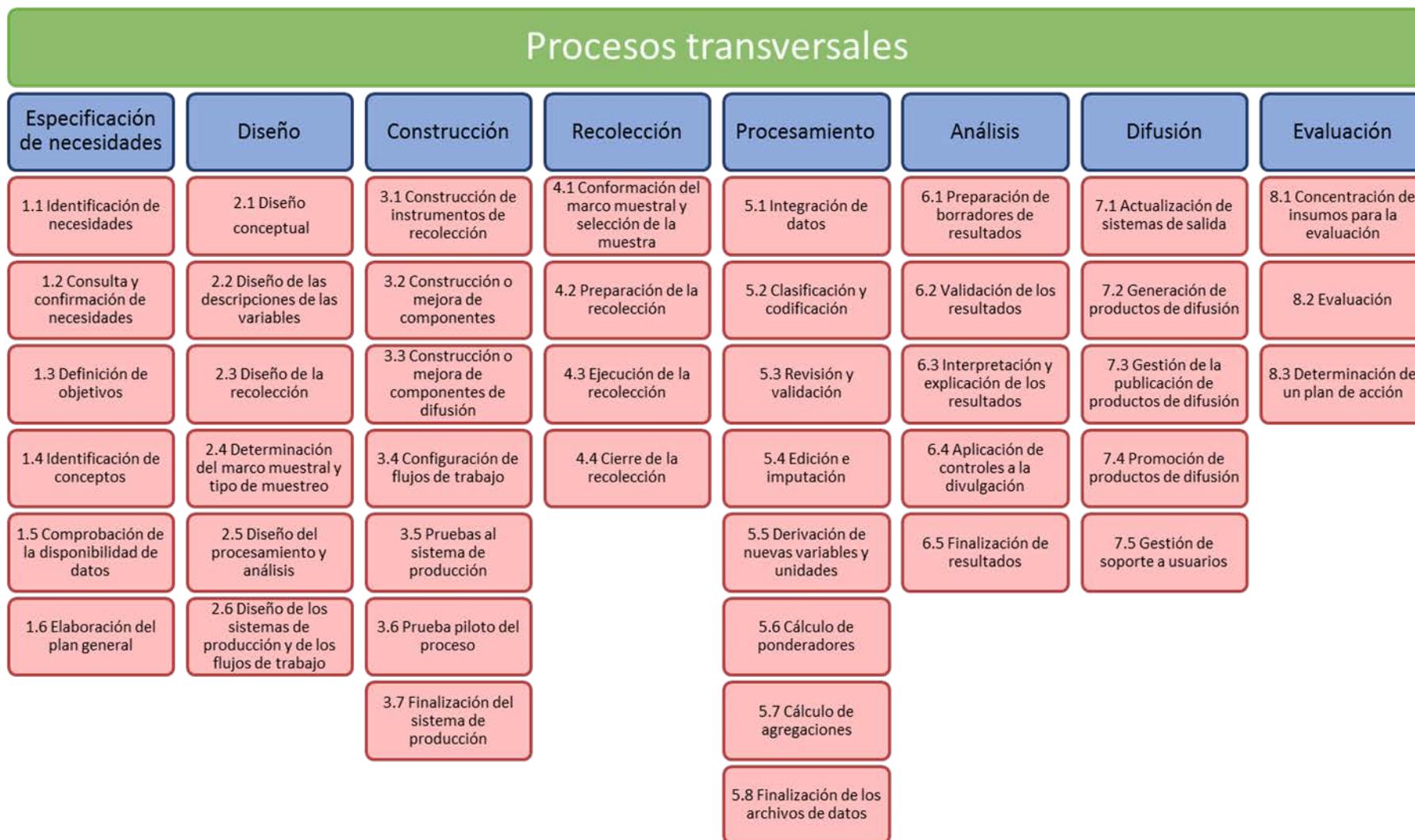
- Gestión de la calidad.
- Gestión de datos.
- Gestión de metadatos.
- Gestión del conocimiento.
- Gestión del marco geoestadístico.
- Gestión del programa geoestadístico.
- Gestión de informantes.
- Gestión de usuarios.
- Gestión de tomadores de decisiones.

Los procesos generales de soporte incluyen:

- Planeación estratégica.
- Planeación operativa.
- Administración de recursos humanos.
- Administración financiera.
- Administración de proyectos.
- Administración del marco legal.
- Administración del marco organizacional.

En los apartados V y VI se describen a detalle estos procesos.

IV. Fases y subprocesos



En esta sección se describe cada fase (Nivel 1) y los subprocesos que componen cada fase (Nivel 2).

1. Fase Especificación de necesidades



Esta fase se inicia cuando se identifica una nueva necesidad o la retroalimentación del proceso provoca una revisión. Incluye las actividades asociadas con la participación de los usuarios para identificar sus necesidades detalladas, proponiendo opciones de solución de alto nivel y la preparación de casos de negocio para satisfacer estas necesidades.

En esta fase, la organización:

- identifica la necesidad de la información estadística y/o geográfica;
- confirma con detalle las necesidades de las partes interesadas;
- establece los objetivos de alto nivel para los resultados estadísticos y/o geográficos;
- identifica los conceptos y variables relevantes para los que se requiere información;
- comprueba el grado en que las fuentes de datos actuales pueden satisfacer estas necesidades;
- prepara el caso de negocio para obtener la aprobación para producir la información estadística y/o geográfica.

Esta fase se divide en seis subprocesos que son generalmente secuenciales, pero pueden ocurrir en paralelo y pueden ser iterativos. Los subprocesos son:

1.1 Identificación de necesidades

Incluye la investigación inicial e identificación de qué información se necesita. Puede ser desencadenada por una nueva solicitud de información, un cambio de medio ambiente o una reducción presupuestal. También se puede derivar de trabajos de los Comités Técnicos Especializados o por requerimientos internacionales como la Agenda 2030 de Desarrollo Sostenible. También incluye la consideración de prácticas de otras organizaciones estadísticas y geográficas (nacionales e internacionales) y en particular los métodos usados por éstas. Puede involucrar la consideración de las necesidades específicas de diferentes comunidades de usuarios, como los discapacitados o diferentes grupos étnicos. Finalmente, pueden ser acciones derivadas de los resultados de las evaluaciones de la ejecución previa del proceso.

1.2 Consulta y confirmación de necesidades

Este subproceso se centra en consultas con las partes interesadas, así como en confirmar a detalle las necesidades a satisfacer por la información estadística y geográfica. Las consultas se pueden llevar a cabo de manera directa con usuarios y expertos, o a través de grupos de enfoque. Una buena comprensión de las necesidades del usuario es necesaria para que la organización comprenda no sólo lo que se espera entregar, sino también cuándo, cómo y, tal vez lo más importante, por qué. Para segunda y sucesivas iteraciones de esta fase, el enfoque principal será determinar si las necesidades identificadas

previamente han cambiado, las razones del cambio y las implicaciones en la comparabilidad con los resultados de ejercicios previos. Esta comprensión detallada de las necesidades del usuario es parte primordial de este subproceso.

1.3 Definición de objetivos

Este subproceso identifica los resultados estadísticos y geográficos que se requieren para satisfacer las necesidades de usuario identificadas en el subproceso 1.2 (consultar y confirmar las necesidades). Incluye la adecuación de los resultados propuestos y sus medidas de calidad de acuerdo con la retroalimentación de los usuarios. Los marcos legales (por ejemplo, relativas a la confidencialidad) y los recursos disponibles probablemente actúen como restricciones al momento de establecer objetivos de salida.

1.4 Identificación de conceptos

Este subproceso aclara los conceptos dentro del proceso estadístico que, desde el punto de vista del usuario final, precisan una medición. En esta etapa los conceptos identificados deben, en la medida de lo posible, alinearse con normas y estándares estadísticos nacionales e internacionales. Por ejemplo, aquellos definidos en la Normatividad del SNIEG. La alineación y la elección o definición de los conceptos estadísticos y variables a utilizar, se lleva a cabo en el subproceso 2.2.

1.5 Comprobación de la disponibilidad de datos

Este subproceso comprueba si fuentes de datos actuales podrían satisfacer los requerimientos de usuario y las condiciones bajo las cuales estarían disponibles, incluyendo las restricciones sobre su uso. Una evaluación de posibles alternativas normalmente incluye investigación de datos administrativos u otras fuentes potenciales de datos no estadísticos, para determinar si sería adecuado su uso para fines estadísticos. Una vez evaluadas las fuentes existentes, se prepara una estrategia para llenar los huecos en los requerimientos de datos. Este subproceso también incluye una evaluación más general del marco jurídico en el que descansaría la recolección y uso de los datos y, por lo tanto, puede identificar propuestas de cambio a la legislación existente o la generación de un nuevo marco legal.

1.6 Elaboración del plan general

Este subproceso documenta los resultados de los otros subprocesos pertenecientes a esta fase, en la forma de un caso de negocio, con el fin de obtener la aprobación para implementar un proceso estadístico o geográfico, nuevo o modificado. El caso de negocio deberá conformarse a los requerimientos del cuerpo de aprobación, pero típicamente incluye elementos tales como:

- Una descripción del estado actual en el que se encuentra el proceso estadístico o geográfico (si ya existe), con información sobre cómo se producen las estadísticas actuales y la información geográfica, destacando cualquier ineficiencia o problema a abordarse;
- La propuesta de solución posible detalla cómo se desarrollará el proceso estadístico y geográfico para producir la información nueva o revisada;
- Una evaluación de los costos y beneficios, así como cualquier restricción externa, como puede ser compromisos con la oportunidad, precisión o comparabilidad de la información.

2. Fase Diseño



En esta fase se describen las actividades de diseño conceptual, metodológico y de sistemas; así como, cualquier trabajo práctico de investigación asociado que se necesite para definir los productos estadísticos y geográficos, los instrumentos de recolección y los procesos operativos. Incluye todos los elementos de diseño necesarios para definir o refinar los productos estadísticos y geográficos o los servicios identificados en el caso de negocios. Esta fase especifica todas las fuentes de información y metadatos correspondientes, dejándolos listos para su uso posterior en el proceso, así como en los procedimientos de aseguramiento de la calidad. Para los productos producidos de forma regular, esta fase se produce normalmente durante la primera iteración y siempre que las acciones de mejora sean identificadas en la fase de “Evaluación” de una iteración anterior.

Las actividades de diseño hacen uso sustancial de estándares 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. Adicionalmente, los productos de los procesos de diseño pueden formar la base para los futuros estándares a nivel organizacional, nacional o internacional.

Esta fase se divide en seis subprocesos que generalmente son secuenciales, pero también pueden darse en paralelo o ser iterativos.

2.1 Diseño conceptual

Este subproceso contiene el diseño detallado de los resultados de información estadística y geográfica que se obtendrán. Estas tareas incluyen el trabajo del diseño de apoyo a la preparación de los sistemas y herramientas usados en la fase de “Difusión”, los métodos de reserva de datos, así como los procesos encargados de acceder a la información confidencial resultante. Los productos deben ser diseñados, en la medida de lo posible, siguiendo estándares existentes por lo que la inclusión de datos en este proceso puede incluir metadatos de colecciones previas o similares, estándares internacionales e información acerca de prácticas en otras organizaciones estadísticas identificados en el subproceso 1.1 (Identificación de necesidades).

2.2 Diseño de las descripciones de las variables

Este subproceso define las variables que deben ser recolectadas directamente a través del instrumento de recolección. Lo mismo ocurre con cualquier otra variable que derivará de ellas en el subproceso 5.5 (Derivar nuevas variables y unidades) y de cualquier clasificación estadística que pueda ser utilizada. Se espera que se sigan estándares nacionales e internacionales siempre que sea posible. Puede que este subproceso necesite correr paralelo al subproceso 2.3 (Diseño de la recolección) ya que la definición de las variables a recolectar, así como la elección del instrumento de recolección pueden ser

interdependientes. La preparación de las descripciones de los metadatos derivados de las variables y sus clasificaciones son una condición necesaria para las fases subsecuentes.

2.3 Diseño de la recolección

Este subproceso determina el(los) método(s) de recolección e instrumento(s) más apropiado(s). Las actividades que de hecho se llevan a cabo en este subproceso van a variar de acuerdo con el tipo de instrumentos de recolección requeridos, los cuales pueden incluir el diseño de entrevistas asistidas por computadora, cuestionarios en papel, interfaces de información administrativa, métodos de trabajo de campo, diseño de croquis y técnicas de integración de datos. Este subproceso incluye el diseño de instrumentos de recolección, preguntas y formatos de respuesta, en conjunto con las variables y clasificaciones estadísticas diseñadas en el subproceso 2.2 (Diseño de las descripciones de las variables). También incluye el diseño de cualquier acuerdo formal relacionado con el suministro de datos, tales como memorandos de acuerdo, así como la confirmación de las bases legales para la recolección de los datos. Este proceso es habilitado por herramientas como las bibliotecas de preguntas (que facilitan la reutilización de preguntas y atributos relacionados con las mismas), herramientas de cuestionario (que permiten la recolección fácil y rápida de preguntas en formatos propios de la evaluación cognitiva) y formatos de acuerdo (que ayudan a estandarizar los términos y condiciones). Este subproceso también incluye el diseño de sistemas administrativos proveídos por procesos específicos.

2.4 Determinación del marco muestral y tipo de muestreo

Este subproceso sólo aplica a procesos que involucran la recolección de datos basados en muestreo tales como estudios estadísticos. Esta actividad identifica y especifica la población objetivo, define el marco muestral (y, cuando es necesario, el registro del cual deriva) y determina los criterios de muestreo, y la metodología más adecuados (que podrían incluir la enumeración completa). Fuentes comunes del marco muestral son registros administrativos y estadísticos, censos e información de otros estudios o encuestas. Este subproceso describe cómo pueden combinarse estas fuentes si así se requiere. Debe realizarse un análisis para verificar que el marco cubra la población objetivo. Además, debe realizarse un plan de muestreo, aunque la muestra como tal se crea en el subproceso 4.1 (Conformación del marco muestral y selección de la muestra) utilizando la metodología que haya sido especificada en este subproceso.

2.5 Diseño del procesamiento y análisis

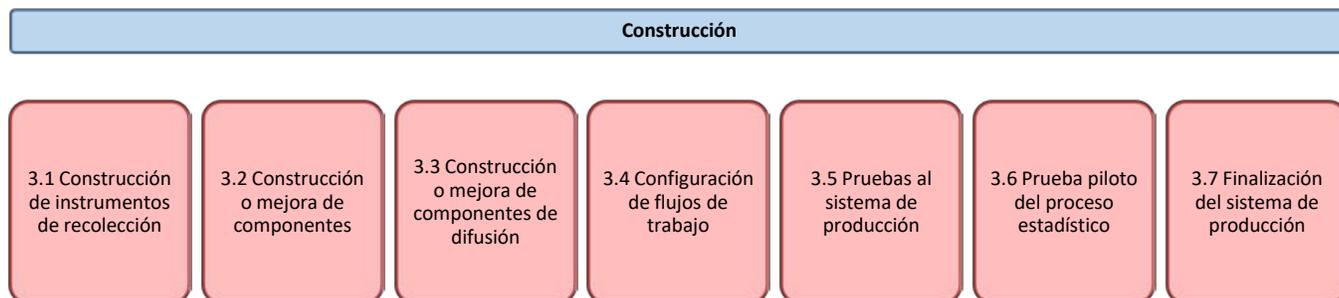
Este subproceso diseña la metodología de procesamiento estadístico y geográfico que deberá ser aplicada en las fases de “Procesamiento” y “Análisis.” Ésta puede incluir especificaciones de rutina para la codificación, edición, imputación, estimación, integración, validación y finalización de los conjuntos de datos.

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

Este subproceso determina el flujo de trabajo desde la recolección de los datos hasta la difusión, tomando en cuenta una visión general de los procesos requeridos dentro de la totalidad del proceso de producción y se asegura de que se ajusten de forma eficiente en su conjunto sin brechas ni redundancias. Se requieren diversos sistemas y bases de datos a lo largo del proceso. Un principio general es reutilizar procesos y tecnología a lo largo de varios procesos estadísticos y geográficos. De esta forma, las soluciones de producción existentes (por ejemplo: servicios, sistemas y bases de datos) deben ser examinadas, primero para determinar si se ajustan al objetivo de un proceso específico. Si se da el caso de que existen brechas, se deben diseñar nuevas soluciones. Este subproceso también considera la forma

como el personal interactuará con los sistemas y quién será responsable de qué y dónde tendrá que hacerlo.

3. Fase Construcción



Esta fase construye y prueba la solución de producción al punto en el que está listo para ser usado en un ambiente “en vivo.” Los resultados de la fase de “Diseño” dirigen la selección de procesos reutilizables, instrumentos, información y servicios que son ensamblados y configurados en esta fase para crear el ambiente operativo completo para correr el proceso. Por excepción, se construyen nuevos servicios generados en respuesta a brechas en el catálogo existente de fuentes de servicio que se encuentran dentro y fuera de la organización. Estos nuevos servicios son construidos para ser ampliamente reutilizados dentro de la arquitectura de producción estadística y geográfica.

Para productos, generalmente de software, producidos de forma regular en esta fase se generan usualmente en la primera iteración. Las iteraciones subsecuentes normalmente siguen adecuaciones por razones de cambios en metodología o por innovaciones o restricciones tecnológicas.

Esta fase se divide en siete subprocesos, los cuales son generalmente secuenciales, de izquierda a derecha en la figura de arriba, pero que también pueden ocurrir en paralelo y pueden ser iterativos.

3.1 Construcción de instrumentos de recolección

Este subproceso describe las actividades para construir los instrumentos de recolección que se usarán durante la fase de “Recolección”. El instrumento de recolección es generado o construido con base en las especificaciones de diseño creadas durante la fase de “Diseño.” La recolección puede utilizar uno o más modos para recibir los datos. Por ejemplo: entrevistas personales o telefónicas, cuestionarios en papel, cuestionarios electrónicos, estructuras SDMX, formatos de colecta de información de campo, técnicas de fotointerpretación y fotoidentificación. Los instrumentos de recolección pueden ser también rutinas de extracción usados para recolectar datos de conjuntos de datos estadísticos, geográficos o administrativos existentes. Este subproceso también incluye preparar y probar los contenidos y funcionamiento del instrumento (por ejemplo: probar las preguntas de un cuestionario). Es recomendable considerar la conexión directa de instrumentos de recolección al sistema de metadatos para que de esta forma los metadatos puedan ser capturados fácilmente en la fase de recolección. La conexión entre los metadatos y datos en el punto de captura puede reducir la cantidad de trabajo en fases posteriores. Capturar la métrica de la recolección de datos (paradatos) es también una consideración importante en este subproceso.

3.2 Construcción o mejora de componentes

Este subproceso describe las actividades para construir y mejorar componentes de software, así como servicios de software necesarios para las fases de “Procesamiento” y “Análisis”, de acuerdo a lo que

haya sido diseñado en la fase de “Diseño.” Los servicios de los componentes pueden incluir funciones y características del tablero, servicios de información, funciones de transformación, marcos de referencia del flujo de trabajo, así como servicios de gestión de proveedores y metadatos.

3.3 Construcción o mejora de componentes de difusión

Este subproceso describe las actividades necesarias para construir o mejorar componentes y servicios de software que sean necesarios para la difusión de los productos generados. La construcción o mejora debe seguir el diseño especificado en la fase de Diseño. Los tipos de componentes y servicios de software van desde aquellos que son usados para producir publicaciones tradicionales en papel, hasta aquéllos que brindan servicios web, productos de datos abiertos o acceso a microdatos.

3.4 Configuración de flujos de trabajo

Este subproceso configura el flujo de trabajo, sistemas y transformaciones utilizadas dentro del proceso estadístico, desde la recolección de los datos hasta la difusión de los mismos. Esto asegura que el flujo de trabajo especificado en el subproceso 2.6 (Diseño de los sistemas de producción y de los flujos de trabajo) funcione en la práctica.

3.5 Pruebas al sistema de producción

Este subproceso se ocupa de probar los servicios ensamblados y configurados y los flujos de trabajo relacionados a ellos. Incluye la prueba técnica y aprobación de nuevos programas y rutinas, así como brindar confirmación de que las rutinas de otros Modelos Generales de Procesos sean adecuadas para utilizarlos en esta ocasión. Mientras que parte de la actividad relacionada con la prueba de componentes individuales y servicios podría lógicamente unirse con el subproceso 3.2 (Construir o mejorar componentes del proceso), este subproceso también incluye probar las interacciones entre los servicios ensamblados y configurados y asegurarse de que la solución de la producción funcione como un conjunto de procesos, información y servicios coherentes.

3.6 Prueba piloto del proceso

Este subproceso describe las actividades relacionadas con gestionar una prueba de campo o piloto del Proceso. Típicamente, éste incluye una recolección de datos a menor escala para probar los instrumentos de recolección seguido del procesamiento y análisis de los datos recolectados para asegurarse de que el MPEG funcione de la manera esperada. Después de la prueba piloto, puede que sea necesario regresar a pasos anteriores y hacer ajustes a los instrumentos, sistemas o componentes. Para procesos importantes, un censo poblacional, por ejemplo, puede haber varias iteraciones hasta que el proceso funcione satisfactoriamente.

3.7 Finalización del sistema de producción

Este subproceso incluye las actividades para poner los componentes y servicios de software - ensamblados y configurados - en producción. Lo anterior permitirá el uso de los componentes y servicios de software por parte de las áreas de negocio. Estas actividades incluyen:

- Producir documentación acerca de los componentes y servicios de software, incluyendo documentación técnica y manuales del usuario.
- Capacitación a los usuarios sobre cómo operar la funcionalidad del software.

- Poner los componentes del proceso en el ambiente de producción y asegurarse de que funcionen de la forma esperada en dicho ambiente.

4. Fase Recolección



Esta fase recopila la información necesaria (datos y metadatos) utilizando diferentes métodos de recolección (incluyendo extracciones de registros y bases de datos estadísticos, administrativos y no estadísticos) y los carga al ambiente adecuado para mayor procesamiento. Mientras que puede incluir validación de formatos de conjuntos de datos, no incluye transformaciones de los datos en sí mismos, ya que esto se lleva a cabo en la fase de “Procesamiento.” Con respecto a la información geoespacial, la recolección de datos puede implicar la adquisición de información de otras instituciones (imágenes de satélite, archivos vectoriales, geodatabases), así como la creación de la información geoespacial por medio del trabajo de campo, encuestas y técnicas de teledetección. Para productos estadísticos y geográficos generados de forma regular, esta fase se lleva a cabo en cada iteración.

La fase de “Recolección” está conformada por cuatro subprocesos 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

Este subproceso conforma el marco y selecciona la muestra para la iteración de la recolección, como se especifica en el subproceso 2.4 (marco y diseño muestrales). También incluye la coordinación de muestras entre instancias del mismo Modelo Genérico de Proceso Estadístico (por ejemplo: para gestionar superposiciones o rotación) y entre procesos distintos usando un marco y registro común (por ejemplo: para gestionar superposiciones o expandir la carga de respuesta). La certeza de la calidad y la aprobación del marco y la muestra seleccionada también se llevan a cabo en este subproceso, aunque el mantenimiento de los registros subyacentes, de los cuales se obtienen los marcos de varios procesos estadísticos, se maneja como un proceso genérico distinto. El aspecto del muestreo de este subproceso no es usualmente relevante en procesos basados completamente en fuentes preexistentes (por ejemplo, fuentes administrativas) ya que estos procesos generalmente crean marcos de los datos existentes y continúan posteriormente en un enfoque de censo.

4.2 Preparación de la recolección

Este subproceso se asegura de que las personas, procesos y tecnología estén listos para la recolección de los datos y metadatos en todas las modalidades que se han diseñado. Se lleva a cabo en un período de tiempo ya que incluye las actividades de estrategia, planeación y entrenamiento en preparación para las instancias específicas del MPEG. Cuando el proceso se repite de forma regular, algunas (o todas) de

estas actividades pueden no ser explícitamente requeridas para cada iteración. Para procesos únicos y nuevos, estas actividades pueden prolongarse.

Este subproceso incluye:

- preparar el esquema de recolección;
- capacitar al personal que hará la recolección;
- asegurarse de que los recursos para recolectar estén disponibles (por ejemplo, inmuebles, computadoras portátiles o tabletas);
- acordar las condiciones de trabajo con colaboradores o intermediarios que recolecten información (por ejemplo: oficinas regionales, subcontratistas);
- configurar sistemas de recolección para solicitar y recibir los datos;
- asegurarse de la seguridad de los datos que son recolectados;
- preparar los instrumentos de recolección (por ejemplo: impresión de los cuestionarios, pre llenarlos con datos existentes, cargar de cuestionarios y datos en las computadoras de los entrevistadores).

Para fuentes que no provengan de encuestas, este subproceso incluirá -asegurándose de que los procesos, sistemas y procedimientos de seguridad estén en su lugar- recibir o extraer la información necesaria de la fuente.

4.3 Ejecución de la recolección

En este subproceso es donde se implementa la recolección de información con los diferentes instrumentos diseñados para su recolección o recopilación. Esto puede incluir microdatos crudos o agregados generados en la fuente, así como cualquier metadato asociado. Incluye el contacto inicial con el proveedor, así como cualquier acción de seguimiento o recordatorios subsecuentes. Puede incluir datos agregados de forma manual en el punto de contacto o gestión del campo de trabajo dependiendo de la fuente y modo de recolección. Debe registrarse cuándo y cómo se contactó a los proveedores y si estos contestaron.

Este subproceso también incluye el manejo de las unidades de estado y otros proveedores que estén involucrados en la recolección en curso al asegurarse de que la relación entre la organización y el proveedor se mantenga positiva. Del mismo modo, se debe registrar y responder a los comentarios, quejas o dudas. Para fuentes administrativas y no estadísticas, este proceso es breve: el proveedor es contactado para que envíe la información o la manda cuando se había acordado. Cuando la recolección alcanza sus objetivos, se cierra y se realiza el reporte de la recolección. La validación básica de la estructura e integridad de la información recibida puede llevarse a cabo en este subproceso. Por ejemplo, revisar que los archivos estén en el formato adecuado y que contengan los campos esperados. Toda la validación del contenido ocurre en la fase de “Procesamiento.”

4.4 Cierre de la recolección

Este subproceso incluye el cargar los datos recolectados (datos y metadatos) a un ambiente electrónico adecuado para su posterior procesamiento. Puede incluir el tomar los datos de forma manual o de forma automática. Por ejemplo: utilizar personal administrativo o instrumentos de reconocimiento óptico de caracteres de cuestionarios en papel, convertir los formatos de documentos recibidos de otras organizaciones o transferencia de información existente en equipos de cómputo móvil. Puede incluir también el análisis de cifras de cierre, o los datos del proceso (paradatos) asociados con la recolección

para asegurarse de que las actividades de recolección hayan cumplido con los requisitos establecidos. En los casos en los que existe un instrumento de recolección física como un cuestionario en papel, que no requiere procesamiento posterior, este subproceso gestiona el almacenamiento del material.

5. Fase Procesamiento



Esta fase describe la limpieza de los datos y su preparación para ser analizados. Está conformado por subprocesos que revisan, limpian y transforman los insumos para que puedan ser analizados y difundidos como resultados estadísticos o geográficos. Puede repetirse varias veces si es necesario. Para los resultados que se producen de forma regular, esta fase ocurre en cada iteración. Los subprocesos en esta fase pueden ser aplicados a datos de fuentes estadísticas y no estadísticas (con la posible excepción del subproceso 5.6 [Calcular ponderaciones] que usualmente es específico de los datos de la encuesta).

Las fases del “Procesamiento” y “Análisis” pueden ser iterativas y paralelas. El análisis puede revelar mayor entendimiento de los datos, lo que puede llevar a notar que se requiere procesamiento adicional de los datos. Las actividades que conforman las fases de “Procesamiento” y “Análisis” pueden empezar antes de que la fase de “Recolección” termine. Esto permite la compilación de resultados provisionales cuando el tiempo es una preocupación relevante para el usuario y aumenta el tiempo disponible para el análisis.

Esta fase se subdivide en ocho subprocesos que pueden ser secuenciales, de izquierda a derecha, pero que pueden ocurrir en paralelo y pueden ser iterativos. Estos subprocesos son:

5.1 Integración de datos

Este subproceso integra los datos de una o más fuentes. Es en donde los resultados de subprocesos de la fase de “Recolección” se combinan. Los datos que entran pueden provenir de una mezcla de fuentes internas o externas; así como de una variedad de modalidades de recolección, incluyendo extractos de datos administrativos. El resultado es una serie de datos vinculados. La integración de los datos puede incluir:

- la combinación de datos de múltiples fuentes como parte de la creación de estadísticos integrados como cuentas nacionales.
- emparejar o registrar vínculos habituales o de rutina con el objetivo de vincular micro y macrodatos de diversas fuentes.
- priorizar, cuando dos o más fuentes contienen datos de la misma variable, con valores potencialmente distintos.

La integración de los datos puede llevarse a cabo en cualquier punto en esta fase: antes o después de los subprocesos 5.2 a 5.7. Pueden existir diferentes ocasiones en las que se integre la información. Después

de la integración (dependiendo de los requisitos de protección de los datos) los datos pueden convertirse en anónimos, quitando identificadores como nombres y domicilios, para promover la confidencialidad.

5.2 Clasificación y codificación

Este subproceso clasifica y codifica la aportación de datos. Por ejemplo, rutinas de codificación automáticas o administrativas pueden asignar códigos numéricos a respuestas en forma de texto por medio de un esquema de clasificación predeterminada. También se puede incluir la revisión y codificación de preguntas abiertas.

5.3 Revisión y validación

Este subproceso examina los datos para tratar de identificar problemas potenciales, errores y discrepancias como valores atípicos, respuestas faltantes o errores en la codificación. También puede llamarse validación de los datos entrantes. Puede ser ejecutada de forma iterativa, validando los datos por medio de reglas de edición preestablecidas, usualmente en un orden fijo. Puede marcar datos para inspección automática o manual. La revisión y validación puede aplicarse para datos de cualquier tipo de fuente, antes y después de la integración. Mientras que la validación se considera parte de la fase de “Procesamiento”, en la práctica, algunos elementos de validación pueden ocurrir a la par de procesos de recolección, particularmente en los casos en los que la recolección de información es llevada a cabo a través de la web. Aunque este proceso se enfoca en la detección de errores o errores potenciales, cualquier actividad de corrección que altere los datos se debe llevar a cabo en el subproceso 5.4.

5.4 Edición e imputación

Cuando los datos son considerados incorrectos, faltantes o poco confiables, se pueden insertar nuevos valores en este subproceso. Los términos de edición e imputación cubren una gran variedad de métodos para realizar estos cambios y frecuentemente usan un enfoque basado en reglas. Los pasos específicos por lo general incluyen:

- determinar si se agregan o se cambian los datos;
- la selección del método que se va a utilizar;
- añadir o cambiar valores de los datos;
- escribir los nuevos valores en el nuevo conjunto de datos y marcarlos como cambiados;
- la producción de metadatos en el proceso de edición e imputación.

5.5 Derivación de nuevas variables y unidades

Este subproceso deriva los datos para variables y unidades que no son provistas de forma explícita en la recolección, pero que se necesitan para obtener los resultados requeridos. Deriva las nuevas variables al aplicar fórmulas aritméticas a una o más de las variables que ya están presentes en el conjunto de datos o aplica diferentes supuestos del modelo. Esta actividad puede que requiera ser iterativa ya que algunas variables derivadas pueden estar basadas en otras variables derivadas. Es por esto que es importante asegurarse de que las variables sean derivadas en el orden correcto. Unidades nuevas pueden ser derivadas al agregar o separar datos para unidades de recolección o por medio de otros medios de estimación. Un ejemplo puede ser derivar hogares cuando la unidad de recolección son personas o derivar negocios cuando la unidad de recolección de datos son unidades legales.

5.6 Cálculo de ponderadores

Este subproceso crea ponderaciones para unidades de registros de datos de acuerdo con la metodología creada en el subproceso 2.5 Diseño del procesamiento y análisis. Cuando se trata de una encuesta con muestreo, las ponderaciones se pueden utilizar para hacer que los resultados sean representativos de la población objetivo o para ajustar la falta de respuesta en enumeraciones totales. En otras situaciones, las variables pueden necesitar ser ponderadas para propósitos de normalización.

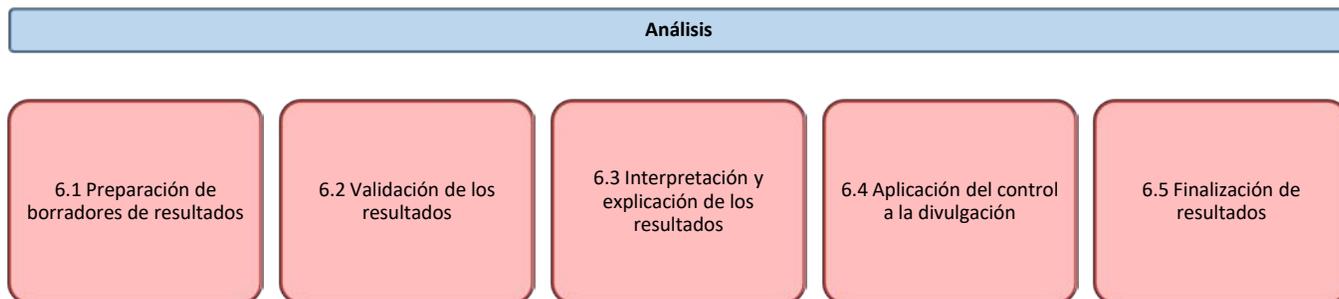
5.7 Cálculo de agregaciones

Este subproceso crea datos agregados y totales de la población a partir de microdatos y agregados de menor nivel. Incluye la suma de los datos para registros que compartan ciertas características, determinar medidas de tendencia central y de dispersión, y aplicar ponderaciones del subproceso 5.6 para derivar totales apropiados. Cuando utilizamos encuestas con muestreo, los errores de muestreo pueden ser calculados en este subproceso y ser asociados con los agregados más relevantes.

5.8 Finalización de los archivos de datos

Este subproceso une los resultados de los otros subprocesos en esta fase y genera una carpeta de datos (usualmente de macrodatos) que se utiliza como la base para la fase de “Análisis.” A veces, ésta puede ser una carpeta intermedia y no la final, particularmente en procesos en los que hay fuertes presiones de tiempo y se hayan solicitado resultados preliminares y finales.

6. Fase Análisis



En esta fase se obtienen los resultados estadísticos, se examinan a detalle y se preparan para la difusión. Incluye preparar el contenido estadístico (incluyendo comentarios, notas técnicas, entre otros) y hay que asegurarse de que los resultados sean “aptos para su propósito” antes de la difusión ante los clientes. Esta fase también incluye los subprocesos y actividades que permiten a los analistas estadísticos entender los productos estadísticos obtenidos. Para resultados estadísticos generados con regularidad, esta fase ocurre en cada iteración. Para el caso del análisis espacial se puede entender como investigar o examinar los lugares, atributos y relaciones de las características de los objetos espaciales a través de superposición, las distancias, la ubicación, la selección espacial, intersección, agregación y otras técnicas de análisis. En ocasiones el análisis espacial crea nueva información a partir de la de origen. La fase de “Análisis” y sus subprocesos son genéricos para todos los productos estadísticos sin importar de qué fuentes se obtuvieron los datos.

La fase de “Análisis” se divide en cinco subprocesos que generalmente son secuenciales, de izquierda a derecha, pero que pueden también ocurrir en paralelo y pueden ser iterativos. Los subprocesos son:

6.1 Preparación de borradores de resultados

En este subproceso los datos son transformados en resultados estadísticos. Incluye la producción de medidas adicionales tales como índices, tendencias o series ajustadas estacionalmente, así como el registro de características de calidad.

6.2 Validación de los resultados

Este subproceso es en el que se valida la calidad de los resultados obtenidos de acuerdo con el marco de referencia de calidad, así como las expectativas establecidas. En la parte geográfica corresponde a la validación de acuerdo al modelo y diccionario de datos. Este subproceso incluye las actividades involucradas con la obtención de inteligencia y con el efecto acumulativo de construir un cuerpo de conocimiento sobre un dominio estadístico específico. Este conocimiento se aplica a la colección actual, en el ambiente actual, para identificar cualquier divergencia con las expectativas y hacer un análisis informado. Las actividades de validación pueden incluir:

- revisar que la cobertura de la población y las tasas de respuesta son las requeridas;
- comparar las estadísticas con ciclos anteriores (si aplica);
- revisar que los metadatos y parámetros (metadatos de proceso) están presentes y van de acuerdo con las expectativas;
- confrontar las estadísticas con otros datos relevantes (tanto internos como externos);
- investigar inconsistencias en los estadísticos obtenidos;
- tratamiento de errores;
- análisis de la no respuesta y su sesgo;
- realizar ediciones masivas;
- validar las estadísticas en comparación con las expectativas y la información existente en este dominio.

6.3 Interpretación y explicación de los resultados

Este subproceso es en el que los analistas obtienen un entendimiento profundo de los resultados, al interpretar y explicar los datos obtenidos en este ciclo y valorar qué tanto los mismos reflejan sus expectativas; mediante la revisión de los resultados desde todas las perspectivas, utilizando diferentes herramientas y medios y haciendo análisis profundos.

6.4 Aplicación de controles a la divulgación

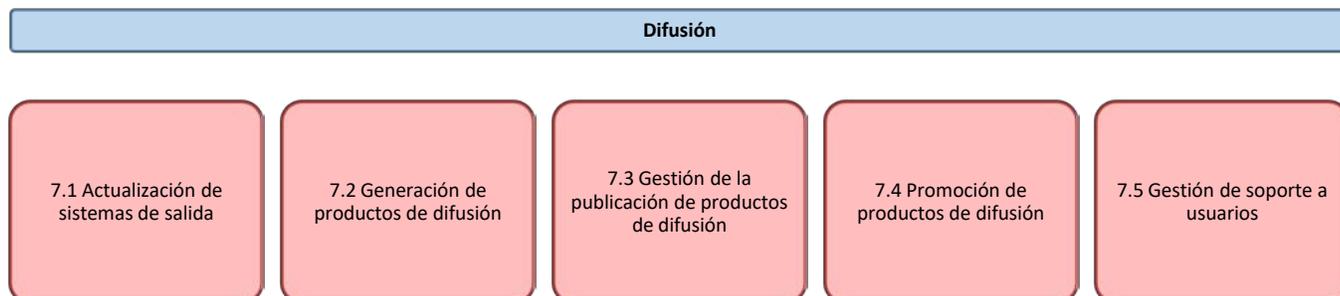
Este subproceso se refiere al resguardo de la información. Asegura que los datos (y metadatos) difundidos no violen las reglas de confidencialidad. Esto puede incluir difusión primaria y secundaria, así como la aplicación de eliminación de datos o técnicas de perturbación. El grado y el método de control de difusión pueden variar entre diferentes tipos de resultados. Por ejemplo, el enfoque que se utiliza con conjuntos de microdatos para fines de investigación va a ser diferente del correspondiente para la publicación en mapas o tablas.

6.5 Finalización de resultados

Este subproceso asegura que las estadísticas y la información que se asocie a ellas sean aptas para el objetivo que se persigue y que alcancen el nivel de calidad requerido y que de esta forma estén listas para ser usadas. Esto incluye:

- completar revisiones de consistencia;
- determinar el nivel de publicación y aplicar advertencias;
- recopilar información de soporte incluyendo la interpretación, comentarios, notas técnicas, informes, medidas de incertidumbre y cualquier otro metadato necesario;
- producir los documentos internos de soporte;
- discutir una versión preliminar con expertos internos o externos en la materia;
- aprobar el contenido estadístico y geográfico para su publicación.

7. Fase Difusión



Esta fase maneja la publicación de los productos a los usuarios. Incluye todas las actividades asociadas con ensamblar y publicar un rango de productos dinámicos y estáticos por medio de distintos canales. Estas actividades ayudan a los clientes a acceder y usar los datos publicados por la organización. Para algunos casos se pueden ofrecer sólo servicios para la visualización de datos, mientras que otros pueden ser a través del desarrollado de plataformas más completas que integran los servicios de mapas con capacidades para el análisis espacial en línea y los servicios de descarga.

Para datos producidos regularmente, esta fase ocurre en cada iteración y la conforman cinco subprocesos que generalmente son secuenciales, de izquierda a derecha, pero que también pueden ocurrir en paralelo y pueden ser iterativos. Estos subprocesos son:

7.1 Actualización de sistemas de salida

Este subproceso gestiona la actualización de sistemas en los que los datos y metadatos se almacenan cuando están listos para propósitos de difusión. Incluye:

- Dar formato a los datos y metadatos listos para ser colocados en bases de datos de salida;
- cargar datos y metadatos en bases de datos de salida;
- asegurarse de que los datos estén vinculados con metadatos relevantes.

Formatear, cargar y vincular metadatos debe llevarse a cabo preferentemente en fases anteriores pero este subproceso incluye una revisión final con el propósito de que todos los metadatos necesarios están en su lugar y listos para ser difundidos.

7.2 Generación de productos de difusión

Este subproceso produce los productos pertinentes, tal y como se diseñaron previamente en el subproceso 2.1, para cubrir las necesidades del usuario (pueden incluir publicaciones impresas, comunicados de prensa y sitios web). Los productos pueden adquirir muchas formas, incluidos gráficos

interactivos, mapas, cubos de datos, archivos vectoriales, imágenes satelitales, tablas, conjuntos de microdatos de uso público y documentos descargables. Pasos típicos en este proceso incluyen:

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

7.3 Gestión de la publicación de productos de difusión

Este subproceso tiene como propósito asegurar que los elementos a ser publicados están en su lugar e incluye gestionar el momento en el que serán publicados. Incluye también informar a grupos específicos como la prensa o ministros; así como, los arreglos para cualquier embargo pre-lanzamiento. Incluye la provisión de productos a los subscriptores y manejar el acceso a datos confidenciales a grupos de usuarios autorizados (como los investigadores). En ocasiones, una organización debe retirar un producto (por ejemplo, si se descubre un error). Esto también se incluye en este subproceso.

7.4 Promoción de productos de difusión

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

7.5 Gestión de soporte a usuarios

Este subproceso se asegura de que las dudas y solicitudes de los usuarios de servicios, como el acceso a microdatos, mapas y geoservicios sean registrados y que se proporcionen respuestas puntuales dentro de un rango de tiempo establecido. Estas consultas y solicitudes deben revisarse periódicamente para proporcionar un insumo al proceso transversal de gestión de la calidad; ya que, éstas pueden indicar necesidades nuevas o cambios en las necesidades de los usuarios.

8. Fase Evaluación



Esta fase maneja la evaluación de una instancia específica del Modelo Genérico del Proceso Estadístico y Geográfico. Al contrario del control de calidad más general que se traslapa a lo largo del modelo, este lógicamente toma lugar al final de la instancia del proceso, pero depende de información obtenida en diferentes fases. Implica evaluar el éxito de una instancia específica del proceso, tomando información cuantitativa y cualitativa e identificando y priorizando mejoras potenciales.

Para resultados producidos de forma regular, la evaluación debe, por lo menos en teoría, ocurrir para cada iteración, determinando si deben ocurrir futuras iteraciones y, si esto ocurre, si se deben implementar mejoras. Sin embargo, en algunos casos, particularmente para procesos generales regulares y bien establecidos, la evaluación puede no ser llevada a cabo en cada iteración. En estos casos, esta fase puede ser vista como la determinante para decidir si la iteración debe iniciar en la fase de Especificación de necesidades o en la fase de “Recolección”.

Esta fase está compuesta por tres subprocesos que por lo general son secuenciales, pero que pueden traslaparse en cierta medida en la práctica. Estos subprocesos son:

8.1 Concentración de insumos para la evaluación

El material de evaluación puede ser elaborado en cualquier otra fase o subproceso. Puede tener diversos formatos, incluyendo retroalimentación de los usuarios, metadatos de procesos (paradatos), métricas de sistemas y sugerencias del personal. Reportes de progreso comparados con un plan de acción con respecto a una iteración previa pueden también formar un insumo para las evaluaciones de iteraciones futuras. Este subproceso reúne todos estos insumos y los pone a disposición de la persona o equipo que está realizando la evaluación.

8.2 Evaluación

Este subproceso analiza los insumos de evaluación y los sintetiza en un reporte de evaluación. El reporte que resulte de esto, debe resaltar cualquier problema de calidad que sea específico de este Modelo Genérico de Proceso Estadístico y Geográfico y debe hacer recomendaciones para realizar los cambios apropiados. Estas recomendaciones pueden cubrir cambios en cualquier fase o subproceso para futuras iteraciones del proceso o puede sugerir que el proceso no se vuelva a repetir.

8.3 Determinación de un plan de acción

Este subproceso reúne el poder de toma de decisión para formar y acordar un plan de acción basado en el reporte de evaluación. También debe incluir la consideración de un mecanismo para monitorear el impacto de dichas acciones que, al mismo tiempo, proveen insumos para evaluaciones de iteraciones futuras del proceso.

V. Procesos Transversales con componente estadístico o geográfico

En el apartado III Estructura, se describió que además del proceso sustantivo el Modelo define una serie de procesos transversales que complementan la labor del proceso estadístico y geográfico. Estos procesos transversales están divididos en dos: 1) los procesos con componente estadístico o geográfico y, 2) Los procesos generales. En esta sección describiremos el primer caso.

1. Gestión de la calidad

En el estándar ISO 9000-2005 la calidad se define como: “El grado en el que un conjunto de características inherentes cubren los requisitos.”. El MPEG provee un marco de referencia estandarizado que facilita obtener datos operativos e indicadores estandarizados que se usen precisamente para evaluar si se cumplen o no los requisitos. En este marco se determina la calidad del producto bajo la primicia de que sólo con un proceso de calidad se puede obtener un producto de calidad. En este contexto, para mejorar la calidad del producto la gestión de calidad debe estar presente a lo largo del proceso definido en el MPEG.

El esquema de Gestión de Calidad diseñado por el INEGI plantea cuatro actividades fundamentales:

- **Gobierno.** Establecido a través de la Norma para el Aseguramiento de la Calidad de la Información Estadística y Geográfica del INEGI y de las acciones que lleva a cabo el Comité del Aseguramiento de la Calidad (CAC). Dentro de las actividades que define y promueve el CAC están:
 - Establecer y mantener el marco de calidad;
 - Establecer criterios de calidad globales;
 - Establecer objetivos de calidad y monitorear su cumplimiento;
 - Buscar y analizar la retroalimentación de los usuarios;
 - Revisar las operaciones y documentar las lecciones aprendidas;
 - Examinar los procesos de metadatos y los indicadores de calidad;
 - Promover evaluaciones internas o externas de los procesos.
- **Monitoreo.** Un papel fundamental dentro de la gestión de calidad es ocupado por un conjunto de acciones de control que deben ser implementadas dentro de los subprocesos para prevenir y monitorear errores. Esta actividad se debe incorporar en la actividad diaria y de ser posible manifestarse en un pizarrón de seguimiento.
- **Evaluación:** La multiplicidad de marcos de referencia de calidad aumenta la importancia de establecer evaluaciones comparativas y revisiones entre pares. Aunque esta perspectiva es difícilmente aplicable para cada iteración del ciclo del proceso. La evaluación debe llevarse a cabo de forma sistemática de acuerdo con un calendario predeterminado que permita revisar el estado y los resultados principales del proceso dentro de un período específico de tiempo.
- **Mejora:** El monitoreo y las evaluaciones resultan en retroalimentación que debe ser usada para mejorar el proceso, fase o subproceso en el que se hayan detectado fallas o acciones susceptibles de mejora. En esta etapa se plantea el diseño de la mejora, su desarrollo, implantación y futuro esquema de seguimiento. En términos del proyecto de mejora, se diseña el estado objetivo del proceso, se desarrolla, se implanta y se evalúa el impacto. En su caso se disparan nuevas acciones que crean un círculo de calidad.

La gestión de calidad también involucra la evaluación de otros procesos transversales como pueden ser la Gestión de recursos humanos, la Gestión de datos y la Gestión del marco geo estadístico.

2. Gestión de documentación

El MPEG proporciona una estructura para organizar y guardar documentos dentro de una organización, y promover la estandarización e identificación de prácticas adecuadas. Esta documentación auspicia que los procesos sean repetibles, principalmente a través de la formalización de la obligatoriedad de la documentación, establecida en políticas, normas y procedimientos de la institución.

La Oficina Nacional de Estadísticas de Nueva Zelanda ha desarrollado una base de datos para almacenar el modelado de procesos y productos y permitir que estos se vinculen al GSBPM. En este mismo contexto, el MPEG se puede utilizar como una taxonomía para construir y almacenar conocimiento del modelado de procesos, para promover y aumentar el entendimiento del mismo y facilitar la mejora.

3. Gestión de datos

En esta gestión se incluyen consideraciones independientes del proceso, como la seguridad de los datos en general, custodia y propiedad, calidad de los datos, reglas de archivo, preservación, conservación y

eliminación. De igual forma la gestión debe completar el manejo de datos maestros, datos de referencia, copias doradas (Golden copies), catálogos nacionales e internacionales.

Existe una serie de modelos y normas relacionados con el GSBPM que resultan de iniciativas de modernización. El más significativo de ellos es el Modelo Genérico de Información Estadística (GSIM) el cual permite descripciones genéricas para la identificación, gestión y uso de datos y metadatos a lo largo del proceso de producción estadística. La aplicación conjunta del GSIM y el MPEG puede facilitar la construcción de sistemas eficientes de metadatos y ayudar a armonizar las infraestructuras de aplicaciones de cómputo. Por ejemplo, completando las definiciones conceptuales del GSIM, documentadas en DDI e implementadas en SDMX.

Para el caso de los datos geográficos se debe realizar a través de la Norma Técnica para la Elaboración de Metadatos Geográficos (DOF 24 de diciembre de 2010), la cual está basada en el estándar ISO 19115.

4. Gestión de metadatos

El marco común que provee el MPEG puede ayudar a integrar el trabajo de metadatos al proveer un marco de referencia común y terminología común para describir el proceso general. La gestión adecuada de los metadatos es esencial para la operación del Modelo. Esto incluye consideraciones que son independientes del proceso, como la custodia de metadatos y su propiedad; su calidad, y las reglas de preservación, conservación y eliminación de archivo.

Podemos identificar dos grupos de metadatos: el primero asociado a la descripción y caracterización del proceso y el segundo asociado a la descripción de la información. Los metadatos del proceso definen el que, como y cuando de cada una de las fases. Por ejemplo, en esta clasificación se incluye información de los informantes, las características de la recolección, las reglas de transformación, las reglas de validación, los mecanismos de agregación, la relación de usuarios finales.

Los metadatos asociados a la información normalmente se definen como parte del diseño y caracterización de los datos de entrada y los productos de información de salida. Aquí se incluyen las características de los datos de insumo, los productos de información y la presentación de los mismos. El principal reto es asegurarse que los metadatos sean capturados, se almacenen, se actualicen y sean accesibles junto con los datos a los que hacen referencia.

Dependiendo de la implantación los dos conjuntos de metadatos podrían estar en un solo depósito de información. Este depósito podría implementarse en una base de datos de propósito específico. El Modelo de los metadatos se conoce como metamodelo.

5. Gestión de parados (datos de proceso)

Los parados son los rastros generados por un proceso. Ejemplo de parados son: la hora de inicio de la ejecución, el tiempo del proceso, el ejecutante de los procesos y el número y tipo de errores. Los parados son generados en cada fase e incluso en cada subprocesso. Es conveniente contar con un Sistema de Gestión de parados ya que puede ser la base para producir indicadores de gestión que pueden retroalimentar los procesos de evaluación y mejora.

6. Gestión del conocimiento

La gestión del conocimiento incluye el almacenamiento y gestión de todos los datos de documentación del proceso de datos, las evaluaciones y las propuestas de mejora. Es la administración de la “memoria

institucional” que permitirá a cualquier persona examinar la historia de los eventos, los problemas y sus soluciones.

7. Gestión del marco geoestadístico

Esta gestión incluye el desarrollo de estándares, metodologías, conceptos y clasificaciones que se aplican a través de múltiples procesos. Incluye también la gestión de las bases de datos y de las actualizaciones cartográficas que se utilizan como marco de referencia para relacionar la información estadística con el espacio geográfico correspondiente, la documentación de sus procesos de selección de muestra y las metodologías para generar los factores de expansión.

8. Gestión del programa geoestadístico

Esta gestión incluye un monitoreo y revisión sistemáticos, de los requerimientos emergentes de información y el cambio de las fuentes de datos a través de todos los dominios estadísticos y geográficos. Puede dar lugar a la definición de nuevos procesos estadísticos o al rediseño de los ya existentes.

9. Gestión de informantes

Esto incluye la gestión de la carga entre procesos, el perfilado de proveedores y la gestión de la información de contacto (y por lo tanto tiene vínculos estrechos con procesos que mantienen registros).

10. Gestión de usuarios

Esto incluye las actividades generales de mercadotecnia, la promoción de la cultura estadística y hacer frente a la retroalimentación por parte de clientes no especificados.

11. Gestión de tomadores de decisiones

Esto incluye las actividades de los usuarios, generalmente del sector público, que utilizan la información que se genere.

VI. Procesos transversales sin componente estadístico o geográfico

Los procesos transversales generales de soporte incluyen:

1. Planeación estratégica

El MPEG es un marco conceptual que puede ser usado con referencia para la modernización de procesos estadísticos. Un plan estratégico de una organización genera objetivos y líneas de acción (el ¿Qué?) pero no necesariamente hace explícito los elementos necesarios para llevarlas a cabo. En este contexto una organización productora de información estadística y geográfica puede basarse en el MPEG como guía para la ejecución de estrategias y líneas de acción (el ¿Cómo?).

Un aspecto importante del MPEG, es que se puede tomar como base de una modernización del proceso estadístico y geográfico para tomarse de referencia y evaluar el avance que las áreas productoras puedan tener con respecto a la adopción del estándar.

2. Administración del marco legal

Para una adecuada gestión de procesos, no es suficiente que se establezcan las definiciones conceptuales. En toda institución se deben incorporar la adopción de los procesos como parte integral de la normativa que puede incluir: principios, políticas, normas, procedimientos y lineamientos. La gradualidad dependerá de la capacidad de adopción de cada institución y de sus capacidades en la implementación del enfoque a procesos.

3. Administración financiera

El MPEG provee un marco, por etapas a través de las cuales se puede descomponer un proceso de producción estadística o geográfica y así dar seguimiento a los recursos presupuestales y a los costos de producción. La evaluación de costos ayuda a impactar actividades de modernización que mejoren la eficiencia de las partes que se identifiquen como más caras.

Relacionado con los costos, el MPEG puede también identificar componentes que no están siendo ejecutados de forma eficiente, que se duplican innecesariamente o que requieran ser reemplazados. De forma similar, puede identificar brechas para componentes que deban ser desarrollados.

4. Administración de proyectos.

La adopción del MPEG es un proyecto en sí mismo. Esta adopción puede asociarse a la implantación de un proceso o proyecto estadístico y geográfico que se lleva a cabo por primera vez o a la adaptación de proyectos que ya estén completamente implantados. Para la administración de proyectos es conveniente adoptar estándares internacionales como los proporcionados por el Project Management Institute (PMI).

5. Administración de recursos humanos.

El MPEG puede ser usado para desarrollar un marco de referencia que evalúe el conocimiento y aptitudes que existen dentro de una organización, así como identificar las brechas que deben ser cubiertas para mejorar la eficiencia operativa.

6. Armonización de Arquitectura de Aplicaciones (software)

El MPEG puede ser visto como un modelo operacional que apoye las definiciones de las aplicaciones de cómputo que procesan los datos. El MPEG Identifica los componentes clave del proceso estadístico general, promueve terminología estándar y promueve formas de trabajo estandarizadas a lo largo de procesos estadísticos y geográficos.

El MPEG define los componentes de procesos estadísticos y geográficos de una forma que promueve el compartir herramientas de software entre procesos globales. En este contexto también facilita el compartir software entre diferentes organizaciones que aplican el modelo. Un ejemplo de este uso lo llevo a cabo el Consejo Asesor de Participación (Sharing Advisory Board).

7. Administración organizacional.

El MPEG, en su definición básica, propone las acciones para la producción de información que se deben llevar a cabo en una organización. Como se comentaba en el apartado de Estructura, el MPEG identifica procesos sustantivos y procesos de soporte. De manera implícita, el MPEG asume que las organizaciones tienen una orientación a procesos. Sin embargo, existen muchas organizaciones con

estructura funcional que, podrían adoptar el Modelo como referencia primaria de una estructura orientada a procesos.

VII. Esquemas de adopción del Proceso Sustantivo

El MPEG es un modelo de referencia. Se pretende que pueda ser utilizado por organizaciones en diferentes grados. Una organización puede elegir implementarlo directamente o utilizarlo como base para desarrollar una versión específica de la organización que se adapte a su contexto organizacional. Sin embargo, al ser una recomendación de mejores prácticas de la industria, la recomendación es no adaptar el modelo sino adaptar las actividades de la organización a las actividades propuestas en las fases y los subprocesos del MPEG.

No obstante, la adopción del modelo no implica una estructura rígida y absoluta. En algunos casos puede ser apropiado utilizar solo un subconjunto de la funcionalidad, en otros se podría suponer una adopción por fases, sin necesidad de adoptar a detalle la funcionalidad de cada subproceso. En contraste, en algunos casos se podría necesitar un detalle mayor al del modelo de tal forma que fuera necesario desarrollar un Nivel 3 o un Nivel 4.

A continuación, se desarrollan algunas perspectivas sobre las cuales se podría adoptar el MPEG. Cabe comentar que la elección de alguna de ellas en una primera etapa, de ninguna manera restringe la posibilidad de usar otra en una segunda o tercera etapa hasta llegar al nivel de adopción deseado.

1. Adopción secuencial

La forma más simple de adoptar el MPEG es simplemente seguir la secuencia de las acciones que recomienda, tanto a nivel fase como a nivel subproceso. En este caso las actividades iniciarían con la fase de Análisis de Requerimientos y llevarían a cabo cada una de las acciones de los 6 subprocesos. Posteriormente se haría lo mismo con la fase de diseño y así hasta llegar a la fase de evaluación. En este caso vale la pena reflexionar sobre cada subproceso y, antes de omitirlo, considerar que es una recomendación de mejor práctica.

Un caso diferente sería cuando el proceso no sea necesario debido a las características de la información que se está recolectando. Por ejemplo, si se trata de registros administrativos, el subproceso:

2.4 Determinación del marco muestral y tipo de muestreo.

no es necesario.

2. Adopción por niveles

En el apartado III Estructura de este documento se describen los diferentes niveles de descomposición que están definidos en el MPEG. La adopción del Modelo se puede llevar a cabo de manera incremental por niveles.

Nivel 0 - Adopción global.

En este caso se utiliza la conceptualización de un proceso general con entradas, transformación de información y salida. La ventaja de este primer paso es que se pueden empezar a generar algunos indicadores de calidad, o productos de cierre, dentro del proceso global. Esta adopción aunque pareciera simple, se puede tomar como base para futuras adopciones.

Nivel 1 - Adopción por fases

En este esquema cada fase toma la función general de un proceso en donde se identifican entradas, procedimientos a realizar y un cierre formal de cada fase con entregables e indicadores específicos.

Nivel 2 - Adopción por subprocesos

La adopción a Nivel 2 implica la realización de cada una de las actividades identificadas en los 44 subprocesos. Para cada subproceso se debe tener una especificación formal de sus entradas, las actividades a realizar y las salidas que marquen la terminación adecuada del mismo.

3. Adopción respecto al tipo de datos origen

El MPEG esta conceptualizado para que pueda ser aplicado en todas las actividades realizadas por los productores de información estadística y geográfica oficiales. El modelo está diseñado para ser independiente de la fuente de datos, por lo que puede utilizarse para la descripción y evaluación de la calidad en procesos basados en encuestas, censos, registros administrativos, agregación de estadísticas y otras fuentes no estadísticas o mixtas.

Mientras que el típico proceso incluye la recolección y tratamiento de datos para producir resultados, el MPEG también se aplica a los casos en que los datos existentes se revisan, son transformados o recalculados, ya sea como resultado de contar con fuentes de datos mejoradas o por un cambio en la metodología. En estos casos, los datos insumo están dados por información anteriormente publicada, que es procesada y analizada para producir nueva información. En tales casos, algunas fases (particularmente las primeras) se omiten. Del mismo modo, el MPEG se puede aplicar a procesos tales como la compilación de las cuentas nacionales y los procesos típicos de organizaciones internacionales de estadística.

Además de ser aplicable en los procesos que dan lugar a información estadística, el MPEG también se puede aplicar en el desarrollo y mantenimiento de registros estadísticos, en los que los insumos son similares a aquellos de la producción de información estadística (aunque por lo general con un mayor énfasis en los datos administrativos) y los productos son típicamente marcos u otras extracciones de datos, que luego son utilizadas como insumos para otros procesos³. Una posible versión del MPEG en el contexto de registros administrativos se ilustra en la Figura 4.

³ Ver: <http://www1.unece.org/stat/platform/download/attachments/57835551/BR+meeting+paper2.doc>

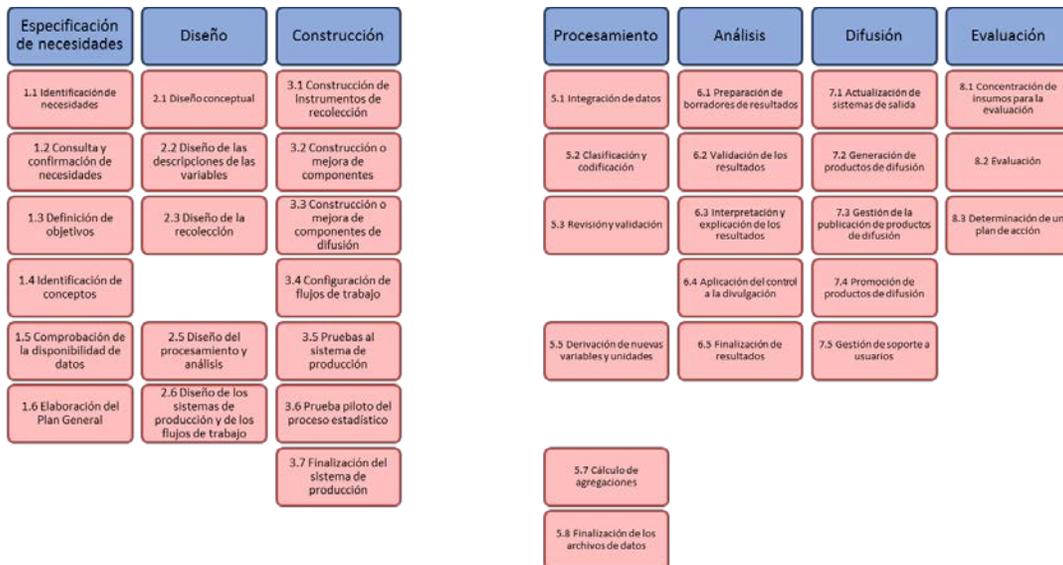


Figura 4

4. Adopción por subconjuntos de fases

Las instituciones con estructura organizacional con enfoque funcional (sin enfoque a procesos) tendrán necesariamente que utilizar el MPEG en el subconjunto que identifique la función específica. Por ejemplo, si la estructura organizacional es un área de TI que se concentra en la construcción de aplicaciones, esta unidad debería adoptar la Fase 3. Construcción con sus 6 subprocesos para definir sus acciones primordiales. En forma similar, si existe un área dedicada a la difusión de información producida ésta deberá tomar como base la Fase 6. Difusión con sus 5 subprocesos.

5. Adopción no secuencial por subprocesos

El MPEG puede ser aplicado 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 y las interdependencias entre ellos.

Aunque la presentación del MPEG sigue una secuencia lógica de pasos usados en la mayoría de los procesos, los elementos del modelo pueden presentarse en diferente orden, de acuerdo a las diferentes circunstancias. Además, algunos subprocesos podrán ser ejecutados en varias ocasiones, formando ciclos iterativos, particularmente dentro de las fases de “Procesamiento” y “Análisis”.

El MPEG debe ser visto como una matriz a través de la cual hay muchos caminos posibles. De esta manera, pretende ser lo suficientemente genérico para poder ser aplicable y para promover una vista estandarizada del proceso, sin llegar a ser demasiado restrictivo. Una ilustración del aprovechamiento de esta flexibilidad se describe en el Figura 5.

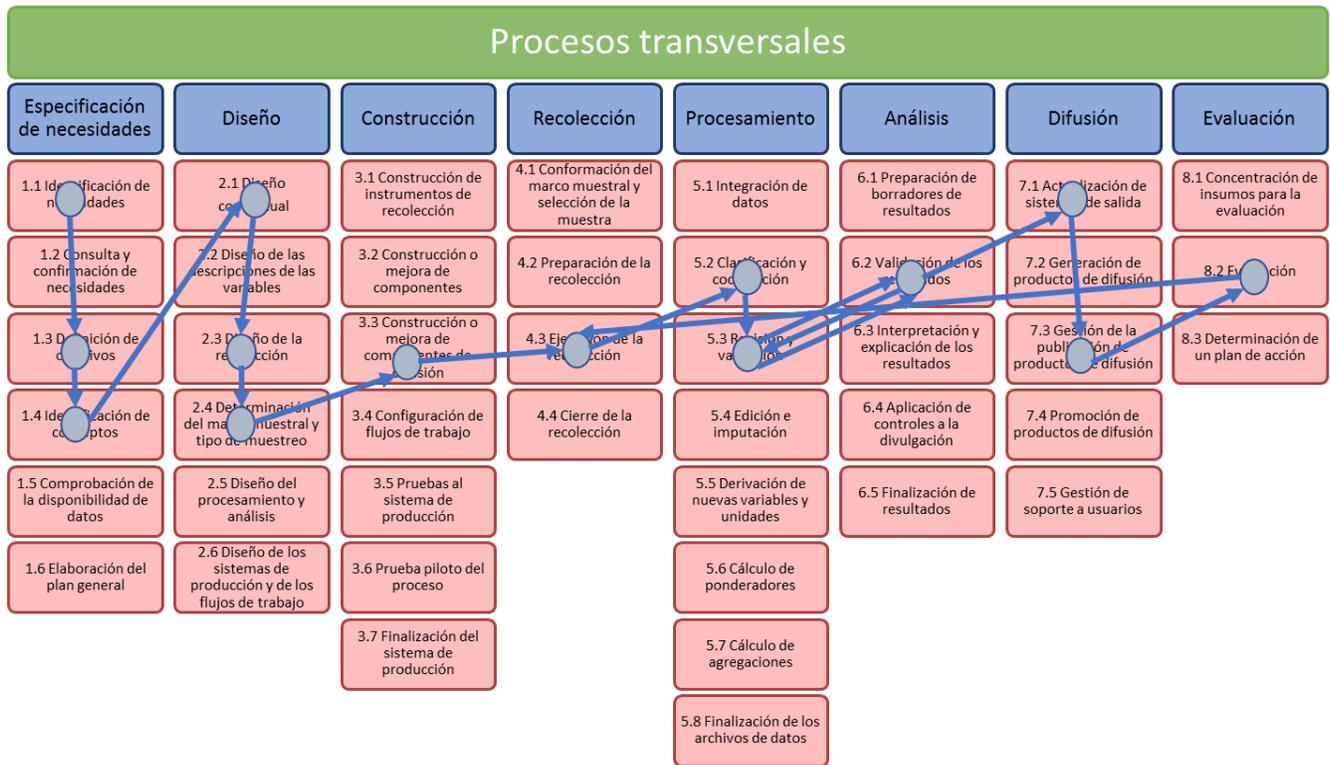


Figura 5

6. Adopción regional

Las instituciones con una cobertura geográfica amplia tendrán que utilizar el MPEG de manera distribuida y sincronizada. Algunas de las fases se podrán llevar a cabo en la oficina central y otras se llevarán a cabo en las oficinas regionales, muy probablemente de manera sincronizada. En este contexto el MPEG resulta de gran valor ya que sirve de base para estandarizar y armonizar las acciones. El típico ejemplo de una ejecución combinada es llevar a cabo las primeras tres fases en la oficina central, la fase de recolección y parte del proceso de manera regional, y las últimas fases de manera centralizada.

VIII. Adopción de los procesos transversales

En forma análoga a la adopción del proceso principal, la adopción de los procesos transversales se puede llevar a cabo de forma gradual e iterativa. En cada iteración se pueden sumar uno o más procesos transversales hasta llegar a adoptar el total de los mismos.

Anexo- Lista de acrónimos

Nota – esta lista sólo cubre algunas de las abreviaciones clave que se utilizan en este documento. Para un glosario completo de los términos relacionados con el proceso genérico estadístico, vea la documentación GSIM.

CMF- Marco Común de Metadatos (Common Metadata Framework): Un conjunto de recursos relacionados con el uso de metadatos por organizaciones estadísticas. Incluye información sobre estándares y prácticas adecuadas. Véase: <http://www.unece.org/stats/cmf/>.

CSPA - Arquitectura Común de Producción Estadística (Common Statistical Production Architecture): Una arquitectura de industria que une el MPEG y el GSIM, incluye nuevos marcos de referencia para Servicios Estadísticos para crear una descripción de alto nivel base del sistema de producción de estadística que se alinea con la iniciativa de modernización.

DDI – Iniciativa de Documentación de Datos (Data Documentation Initiative): Un estándar internacional para describir datos para las ciencias sociales, conductuales y económicas.

GLBPM – Modelo Genérico Longitudinal/General de Procesos (Generic Longitudinal /Business Process Model): Un modelo basado en el MPEG, desarrollado por la comunidad de investigación de encuestas sociales.

MPEG – Modelo Genérico de proceso estadístico (Generic Statistical Business Process Model): Una herramienta flexible que describe y define el conjunto de procesos generales necesarios para producir estadística oficial.

GSIM – Modelo Genérico de Información Estadística (Generic Statistical Information Model): Un marco de referencia de información objetiva que permite las descripciones genéricas de definiciones, manejo y el uso de datos y metadatos a lo largo del proceso de producción estadístico.

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

METIS – El “nombre de marca” para el trabajo sobre Metadatos Estadísticos bajo la Conferencia de Estadísticos Europeos (Conference of European Statisticians).

OECD - Organización para la Cooperación y el Desarrollo Económicos, OCDE (Organisation for Economic Cooperation and Development).

SDMX – Intercambio de Datos y Metadatos Estadísticos (Statistical Data and Metadata eXchange): Un conjunto de estándares técnicos 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.

UNECE – Comisión Económica de las Naciones Unidas para Europa (United Nations Economic Commission for Europa).

XML – Lenguaje de señal extensible (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.