



Lineamientos y Recomendaciones para **Documentar el Diccionario de Datos**

Dirección de
Normativas y Metodologías

Créditos

Dirección General:

Mildred Martínez, Directora General de la Oficina Nacional de Estadística (ONE)

Dirección de Normativas y Metodologías:

Crismairy Jiménez, Directora de Normativas y Metodologías

Coordinación Técnica:

Víctor Dietsch, Encargado del Departamento de Calidad de la Producción Estadística

Equipo técnico:

Anneurys Marmolejos, Analista del Departamento de Calidad de la Producción Estadística.

Apoyo técnico institucional:

Daygoro Díaz, Analista de la División de Gestión de Datos

Diomy Pereyra, Encargada del Departamento de Articulación del Sistema Estadístico Nacional

Eddy Tejeda, Encargado de la División de Encuesta de Actividad Económica

George Objío, Analista del Departamento de Calidad de la Producción Estadística.

Luis Alburquerque, Coordinador de Estadísticas de Comercio Exterior

Otto Rojas, Encargado Interino de la División de Estadísticas Sectoriales

Paola Rodríguez, Encargada del Departamento de Estadísticas Ambientales

Roberto Soriano, Encargado Interino de la División de Gestión de Datos

Xiomara Díaz, Técnica de Operaciones de Censos

Soporte Administrativo:

Juana Lebrón, Auxiliar administrativo de la Dirección de Normativas y Metodologías

Kassandra Sánchez, Auxiliar administrativo de la Dirección de Normativas y Metodologías

Cecilia Belliard, Secretaria ejecutiva Dirección General

Meri Guzmán, Secretaria ejecutiva Dirección General

Apoyo editorial:

Raysa Hernández, Encargada del Departamento de Comunicaciones

Carmen C. Cabanes, Encargada de la División de Diseño y Publicaciones

Iván Ottenwalder, Corrección de estilo

Alferny Eusebio, Diseño y Diagramación

Freepik, Foto de portada

Ficha técnica

Nombre de la publicación	Lineamientos y Recomendaciones para Documentar el Diccionario de Datos
Objetivo general del producto	Establecer un conjunto de pautas y recomendaciones para la documentación, gestión y mantenimiento del diccionario de datos, asegurando que sea una herramienta clara, accesible, estructurada y alineada con estándares internacionales, para el fortalecimiento de la calidad, coherencia e interoperabilidad de los datos producidos por las instituciones del Sistema Estadístico Nacional (SEN), facilitando su uso en la toma de decisiones e integración con otros sistemas.
Descripción general del producto	Este documento establece directrices para la correcta estructuración, gestión y mantenimiento de los datos en sistemas de información. Incluyen normas para la definición de variables, estandarización de formatos, aseguramiento de la calidad, interoperabilidad y seguridad de la información. También, abordan las etapas del proceso, desde la planificación y recolección de elementos hasta la construcción, difusión y actualización del DD.
Año de inicio del producto	2025
Tipo de producto	Lineamiento
Código de producto	LRDDD-CPE-001
Cobertura geográfica	Nacional
Versión del documento	1.0
Fecha de la publicación	20/08/2025
Medios utilizados para la difusión de las publicaciones	Digitales
Contactos	<p>Directora General de la Oficina Nacional de Estadística: Mildred Martínez Tel. 809-682-7777 ext. 2101 direccion@one.gob.do</p> <p>Directora de Normativas y Metodologías: Crismairy Jiménez Tel. 809-682-7777 ext. 3509/3718 dnm@one.gob.do</p> <p>Encargado del Departamento de Calidad de la Producción Estadística Víctor Dietsch Tel. 809-682-7777 ext. 3525 victor.dietsch@one.gob.do</p>
Unidad encargada	Departamento de Calidad de la Producción Estadística, Dirección de Normativas y Metodologías.

Contenido

Siglas y acrónimos	6
Introducción	7
Agradecimientos	8
Objetivos	9
Objetivo General	9
Objetivos Específicos	9
Alcance	9
Marco conceptual	10
Marco normativo	13
Marco metodológico	14
Principios contemplados en el CNBPE relacionados con la Documentación del Diccionario de Datos	15
Relación existente entre el Diccionario de Datos y el Proceso Estadístico	16
Generalidades	17
Lineamientos Generales para la Documentación de un Diccionario de Datos	18
Etapas del proceso	19
Etapa 1. Plan de diseño	19
1.1 Objetivo y alcance del DD	19
1.2 Responsabilidades y roles	20
1.3 Normas y estándares	21
1.4 Revisión y validación	23
Etapa 2. Definición técnica de variables y elementos	23
2.1 Inventario de variables	24
2.2 Diseño de estructura	26
2.3 Normalización y Optimización	27
2.4 Revisión y validación	27
Etapa 3. Construcción	27
3.1 Herramienta para modelo del Diccionario de Datos	28
3.2. Centralización de documentación y herramienta	29
3.3. Asignación de permisos	29

Contenido

Etapas 4. Difusión	30
4.1. Canales y medios de difusión	30
4.2. Medición	31
Etapas 5. Gestión de mantenimiento	32
5.1. Revisión y actualización	33
5.2. Capacitación y documentación	33
5.3. Sistematización	35
5.4. Versionamiento y control de cambios	36
Etapas 6. Evaluación y validación	36
6.1. Monitoreo y control de calidad	37
6.2. Evaluación de impacto	38
6.3. Mejora continua	38
Recomendaciones generales	39
Bibliografía	40
Anexos	41
Anexo 1. Definición de estándares de datos	41
Anexo 2. Ejemplo de cómo estructurar la documentación del DD para una Operación Estadística específica	42
Anexo 3. Ejemplo de un DD para una Base de Datos sobre Comercio Exterior	43
Ejemplo de Datos en la Tabla exportaciones	46

Siglas y acrónimos

ANDA	Archivo Nacional de Datos
API	Interfaz de Programación de Aplicaciones
CNBPE	Código Nacional de Buenas Prácticas para las Estadísticas Oficiales
END	Estrategia Nacional de Desarrollo
ENE	Escuela Nacional de Estadística
ODS	Objetivos de Desarrollo Sostenible
OE	Operación Estadística
ONE	Oficina Nacional de Estadística
PAD	Programa Acelerado de Datos
PEN	Plan Estadístico Nacional
PNPSP	Plan Nacional Plurianual del Sector Público
RRAA	Registros Administrativos
SEN	Sistema Estadístico Nacional
UOE	Unidad Organizacional de Estadística

Introducción

El Diccionario de Datos es una herramienta que describe de forma ordenada la información sobre los nombres, definiciones y características de cada campo o atributo de una base de datos y/o conjunto de datos; se especifica el significado, la organización, los tipos de datos y como se relacionan entre sí. Son útiles para proveer un lenguaje común entre el autor de los datos y los posibles usuarios.

En un contexto donde los datos se han convertido en un recurso fundamental para instituciones públicas, empresas privadas, investigadores y usuarios en general, su correcta interpretación es esencial. Por ello, este documento ofrece pautas y recomendaciones que permitirán la creación, documentación, mantenimiento y uso de los diccionarios.

La Oficina Nacional de Estadística, en su rol de coordinador del Sistema Estadístico Nacional (SEN), y responsable de la generación de estándares únicos, ha establecido, dentro de este documento, directrices y recomendaciones tomando como referencia normas internacionales, lineamientos, las estructuras de los diferentes Diccionarios de Datos y otras herramientas importantes, que se deben establecer, para diseñar, implementar y actualizar un diccionario de dato.

El proceso de diseño de un Diccionario de Datos (DD) se ha establecido en 6 etapas y 21 subetapas, las cuales se deben llevar a cabo para la gestión y actualización del DD, al aplicar este documento en el proceso de producción del DD se establece un hito que se alinea en su totalidad a los lineamientos establecidos para el Proceso Estadístico y buenas prácticas definidas en el Código Nacional de Buenas Prácticas para las Estadísticas Oficiales.

La estructura del documento ("Lineamientos y Recomendaciones para Documentar el Diccionario de Datos") está diseñada con una distribución que busca garantizar ser una herramienta clara y detallada, proporcionando una guía integral para la creación, gestión y mantenimiento de un DD, segregado en tres grandes componentes. A continuación, se presenta su estructura y contenido, destacando los principales elementos que lo componen.

1. Organización general del documento, este es el **primer componente** y se encuentra compuesto por la introducción y contexto, donde se explica la importancia del DD en la gestión de los datos y el papel que representa como herramienta clave para garantizar la coherencia, accesibilidad y calidad de los datos. Además, describe la alineación existente con estándares y normas internacionales.

Objetivos y alcance: en esta sección se define el propósito del documento, asegurando su aplicabilidad en las instituciones productoras y especifica los principios clave que rigen el DD, como claridad, estandarización y trazabilidad.

Por último, este primer componente tiene los marcos conceptual, normativo y metodológico: el marco conceptual incluye las definiciones claves que fueron utilizadas en el desarrollo del documento, el marco normativo detalla las leyes y regulaciones aplicables al contexto sobre el manejo de datos y metadatos y el marco metodológico explica el proceso llevado a cabo para el desarrollo de este documento.

2. Estructura metodológica del DD es el **segundo componente** del documento, el cual establece las vinculaciones existentes de este documento con el Código Nacional de Buenas Prácticas para las Estadísticas Oficiales, los Lineamientos para el Proceso Estadístico en el Sistema Estadístico Nacional, las diferentes etapas y subetapas que se han establecido para la estructuración del Diccionario de Datos y las Recomendaciones.
3. El **tercer y último componente** cuenta con las fuentes bibliográficas, control de cambios y anexos que pretenden evidenciar las investigaciones realizadas para el desarrollo de este documento, registrar los posibles cambios que se realice a través del tiempo y la colocación de una estructura de DD como ejemplo.

La estructura constituye ser un marco sólido para la documentación del Diccionario de Datos. Al replicar estos lineamientos y recomendaciones las instituciones productoras garantizan un DD robusto, accesible, seguro y en constante evolución, contribuyendo a la calidad y eficiencia en la gestión de información estadística.

Agradecimientos

Desde la Oficina Nacional de Estadística, extendemos nuestro más profundo agradecimiento a cada una de las instituciones que forman parte del Sistema Estadístico Nacional por su valiosa participación en la validación de los Lineamientos y Recomendaciones para Documentar el Diccionario de Datos.

Este ejercicio colectivo no solo permitió enriquecer el contenido técnico del documento, sino que reafirmó el compromiso que compartimos con la generación de estadísticas oficiales de calidad, transparente y útil para la toma de decisiones en todos los niveles.

Su disposición a colaborar, aportar experiencias, plantear desafíos y construir soluciones conjuntas es reflejo del espíritu de trabajo en red que nos mueve hacia una cultura estadística más sólida y coherente. Cada comentario, cada sugerencia y cada intercambio durante este proceso representa una huella de compromiso institucional con el fortalecimiento de nuestras capacidades estadísticas.

Gracias por ser parte activa de este paso decisivo. Los resultados que hoy plasmamos en este documento son evidencia de que avanzar en la armonización y estandarización de la documentación estadística es posible cuando se trabaja de manera articulada, con visión de país y voluntad de mejora continua.

Sigamos construyendo juntos un Sistema Estadístico Nacional más integrado, moderno y confiable.

Objetivos

Objetivo General

Establecer un conjunto de pautas y recomendaciones para la documentación, gestión y mantenimiento del DD, asegurando que sea una herramienta clara, accesible, estructurada y alineada con estándares internacionales. Este documento busca fortalecer la calidad, coherencia e interoperabilidad de los datos producidos por las instituciones del Sistema Estadístico Nacional (SEN), facilitando su uso en la toma de decisiones e integración con otros sistemas.

Objetivos Específicos

- Estandarizar la elaboración y documentación del DD, asegurando que cada elemento cuente con definiciones precisas, atributos bien estructurados y metodologías claras para su actualización.
- Definir las etapas y subetapas. clave en la construcción y mantenimiento del DD.
- Incorporar mecanismos de evaluación y mejora continua, permitiendo que el DD evolucione de manera estructurada y se mantenga alineado con las necesidades de la institución productora.
- Garantizar la accesibilidad y seguridad del DD, estableciendo controles y niveles de permisos para su consulta en las etapas de producción, edición y administración, protegiendo la integridad de la información.

Alcance

Este documento está dirigido a las áreas y equipos que gestionan y producen datos, dentro de las instituciones del Sistema Estadístico Nacional, estableciendo buenas prácticas para la elaboración y documentación del DD, abarcando todo el ciclo de vida de este, desde su planificación y diseño hasta su difusión, mantenimiento y mejora continua.

Marco conceptual

Accesibilidad: Facilidad con que la información estadística puede ser ubicada y obtenida por los usuarios. Contempla la forma en que esta se provee, los medios de difusión, así como la disponibilidad de los metadatos y los servicios de apoyo para su consulta.¹

Accesibilidad Web: consiste en lograr que las páginas web sean utilizables por el máximo número de personas, independientemente de sus conocimientos o sus capacidades personales, e independientemente de las características técnicas del equipo utilizado para acceder a la web. (DANE, Departamento Administrativo Nacional de Estadística, 2015).

Anonimización de Datos: proceso técnico que consiste en transformar los datos individuales de las unidades de observación, de tal modo que no sea posible identificar sujetos o características individuales de la fuente de información, preservando así las propiedades estadísticas en los resultados.²

Atributo: característica o variable que se utiliza para describir o explicar una variable dependiente o una variable explicada. Los atributos pueden ser utilizados para identificar patrones, tendencias o relaciones entre variables.³

Calidad: grado en que un conjunto de características inherentes de un objeto cumple con los requisitos. Una definición simple es “apta para el uso” o “apta para el propósito”. Son las necesidades de los usuarios las que definen la calidad. Diferentes usuarios pueden tener diferentes necesidades que deben equilibrarse entre sí. (UNNQAF, 2019).

Calidad de Datos: se refiere a la precisión, integridad, consistencia, confiabilidad y actualidad de los datos almacenados y utilizados en un sistema. Es esencial para la toma de decisiones basada en datos. (ISO 8000-8, 2020)

Calidad Estadística: es el grado de cumplimiento en que un proceso estadístico y sus productos cumplen con los requisitos y propiedades esenciales para satisfacer las necesidades de información de entidades y personas usuarias.

Campo: unidad mínima de almacenamiento de datos en un sistema, representando un atributo específico dentro de un conjunto de datos estructurados. (Kimball y Ross, 2013).

Datos Estadísticos: son el conjunto de datos numéricos resultantes de un proceso de aplicación de métodos y técnicas que ofrece la estadística como ciencia. (ONE, 2022).

Diccionario de Datos: es una recopilación organizada de metadatos que describe la estructura, características y relaciones de los datos dentro de un sistema de información o base de datos. Su propósito es estandarizar la documentación de los datos, mejorar su comprensión y facilitar su gestión. (ISO/IEC 11179-3, 2013).

Diccionario de Datos Activo: solo para bases de datos. Se generan y se actualizan automáticamente por el sistema gestor de la base de datos (SGBD). En la mayoría de los casos pueden configurarse para cumplir con las características que el usuario determine. (Agencia Digital de Innovación Pública, 2021)

Diccionario de Datos Pasivo: deben ser generados de forma manual, principalmente para los conjuntos de datos que no utilizan un sistema gestor de base de datos, como aquellos en XLSX o CSV. (Agencia Digital de Innovación Pública, 2021).

Dominio: el dominio de una variable o campo define el conjunto de valores permitidos para dicha variable. Puede incluir valores numéricos, cadena de texto o listas predefinidas de categorías. (ISO/IEC 11404, 2007).

Estadísticas Oficiales: son estadísticas que describen de manera representativa los fenómenos económicos, demográficos, sociales y ambientales de un país. (CEPAL, 2022).

Extracción Transformación y Carga (ETL): proceso de extraer datos de diferentes fuentes, transformarlos para adaptarlos a las necesidades empresariales y cargarlos en un sistema de almacenamiento centralizado, como un Data Warehouse. (inmon, W. H., 2005). Building the Data Warehouse.

Fuente Administrativa: dependencia u organismo público, privado o mixto responsable del registro administrativo y sus oficinas, donde se llevan a cabo los procesos de captación y mantenimiento de los datos del registro. (INEGI, 2012).

¹ Fuente de información: Numeral 3.1. Términos relativos a los atributos de calidad de la estadística. <https://www.dane.gov.co/files/sen/normatividad/NTC-Proceso-Estadistico-PE-1000-2020.pdf>

² Fuente de información: Numeral 3.5. Términos relativos al resultado. <https://www.dane.gov.co/files/sen/normatividad/NTC-Proceso-Estadistico-PE-1000-2020.pdf>

³ Fuente de información: Definición de Atributo en Estadística: Ejemplos, Autores y Concepto. Fuente de información: Numeral 3.4.9. Términos relativos al resultado. <https://www.dane.gov.co/files/sen/normatividad/NTC-Proceso-Estadistico-PE-1000-2020.pdf>

Fuente de Datos: ubicación específica del conjunto de datos o base de datos donde están disponibles los datos y los metadatos.⁴

Información Estadística: conjunto de resultados y la documentación que los soporta, los cuales se obtienen de las operaciones estadísticas y que describen o expresan características sobre un elemento, fenómeno u objeto de estudio. (DANE, 2019).

Grupos de interés (stakeholders): son personas y organizaciones que están interesadas e involucradas con una entidad y que interactúan con la misma.⁵

Llave primaria: es una columna especial o conjunto de columnas en una tabla de base de datos que identifica de manera única cada registro en la tabla. La clave puede ser una cadena corta o una combinación de múltiples columnas. Cuando se usa en una consulta SQL, la clave principal puede ayudar a recuperar los datos que está buscando más rápidamente. (García, 2023).

Llave foránea: conjunto de columnas en una tabla que se utiliza para relacionar esta tabla con otra tabla. Se utiliza para establecer la relación entre dos tablas que tienen una relación entre sí. (E.F. Codd).

Macrodatos (Big Data): “grandes conjuntos de datos” (McGuire et al., 2012). Loukides (2010) lo define como un dato para el cual “el tamaño de los datos en sí mismo se convierte en parte del problema”.⁶

Metadatos: información necesaria para el uso e interpretación de las estadísticas. Los metadatos describen la conceptualización, calidad, generación, cálculo y características de un conjunto de datos estadísticos. (DANE, 2018).

Microdatos: datos sobre las características de las unidades de estudio de una población (individuos, hogares, establecimientos, entre otros), que constituyen una unidad de información de una base de datos y que son recogidos por medio de una operación estadística. (DANE, 2014).

Normalización de Bases de Datos: es un proceso estructurado que organiza los datos en tablas para reducir redundancias y mejorar la integridad de la información. Se lleva a cabo mediante la aplicación de reglas llamadas formas normales (FN), que aseguran que cada dato esté correctamente ubicado y estructurado. (FreeCodeCamp, 2023).

Proceso Estadístico: conjunto sistemático de acciones que dan como resultado la producción de estadísticas oficiales. Este comprende las etapas Identificación de necesidades de la información; diseño, construcción, recolección, procesamiento, análisis, difusión, evaluación y archivo. (ONE, 2022).

Plan Estadístico Nacional: instrumento cuyo propósito es la sistematización y el ordenamiento de todas las operaciones estadísticas que deberán ejecutar las instituciones del sector público para generar las mediciones, con la rigurosidad, oportunidad y periodicidad necesarias. (PEN, 2021-2024).

Tipos de Datos: los tipos de datos definen qué tipo de valores un campo puede almacenar, como números, texto, valores verdaderos o falsos (booleanos), o fechas. No solo determinan el tipo de información que se puede guardar, sino también las acciones y cálculos que se pueden realizar con esos valores. Por ejemplo, los campos numéricos permiten realizar sumas y promedios, mientras que los campos de texto permiten realizar búsquedas o manipulaciones de cadenas. El tipo de dato es esencial para asegurar que los datos sean utilizados de manera adecuada y coherente dentro de la base de datos.⁷

Registro: conjunto de campos relacionados que conforman una entidad dentro de una base de datos. Cada registro representa una instancia única de la entidad que se está documentando. (Elmasri & Navathe, 2016).

Registro Administrativo: serie de datos sobre un hecho, evento, suceso o acción sujetos a regulación, vigilancia o control que recaba una institución del sector público o privado como parte de su función y en cumplimiento a las normativas jurídicas diferentes a la legislación estadística nacional. (ONE, 2022).

⁴ Fuente de información: Numeral 3.4.9. Términos relativos al resultado. <https://www.dane.gov.co/files/sen/normatividad/NTC-Proceso-Estadistico-PE-1000-2020.pdf>

⁵ Fuente de información: TÉCNICA No. 4 RELACIONAMIENTO DE LOS GRUPOS DE INTERÉS - Rendición de Cuentas - Función Pública

⁶ Fuente de información: OECD (2013), “Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by “Big Data””, OECD Digital Economy Papers, No. 222, OECD Publishing, Paris, <https://doi.org/10.1787/5k47zw3fcp43-en>.

⁷ Tomado de: https://www.conectasoftware.com/magazine/glosario/diccionario-de-datos/?utm_source=2024

Restricciones: las restricciones son las reglas que los datos deben seguir para ser aceptados en el sistema. Estas reglas aseguran que la información ingresada sea válida y consistente. Pueden incluir límites sobre los valores permitidos (como un rango de edades entre 0 y 100), exigir que ciertos campos sean obligatorios (por ejemplo, campos que no pueden dejarse vacíos), o verificar que los datos se ajusten a formatos específicos, como fechas en el formato DD/MM/AAAA o textos con una longitud máxima de caracteres. Las restricciones son clave para mantener la calidad y fiabilidad de los datos.⁸

Relaciones entre los Datos: las relaciones entre los datos son las conexiones que se establecen entre diferentes elementos de información en una base de datos. En los sistemas de bases de datos, estas relaciones se establecen mediante el uso de claves primarias y claves foráneas. Una clave primaria es un identificador único para cada registro dentro de una tabla, lo que garantiza que cada dato sea distinto y fácilmente accesible. Por otro lado, una clave foránea es un campo en una tabla que apunta a la clave primaria de otra tabla, creando un vínculo entre ellas. Esto permite que los datos se mantengan consistentes y que se puedan realizar consultas de manera eficiente sin que la información se repita innecesariamente entre tablas. Al usar claves primarias y foráneas, se asegura que los datos estén correctamente relacionados, lo que favorece la integridad y coherencia de la base de datos.⁹

Variable: característica o propiedad de un objeto de estudio que puede asumir diferentes valores dentro de un conjunto determinado. En estadística, una variable puede ser categórica o numérica y es fundamental para la recopilación y análisis de datos. (DANE, 2018).

Versionamiento: proceso esencial para gestionar los cambios dentro de un sistema. Facilita la documentación de cada modificación, desde la incorporación de nuevos campos hasta la actualización de descripciones, asegurando que todas las personas usuarias trabajen con la versión más reciente y evitando cualquier tipo de inconsistencia en los datos. Este proceso asegura que los cambios sean controlados, claros y fácilmente rastreables, lo que permite mantener la integridad del sistema y la precisión de la información.¹⁰

⁸ Fuente de información: https://www.conectasoftware.com/magazine/glosario/diccionario-de-datos/?utm_source=2024

⁹ Fuente de información: <https://definicionwiki.com/definicion-de-relaciones-de-base-de-datos-ejemplos-segun-autor-que-es-concepto-significado/>, s.f.

¹⁰ Fuente de información: <https://our-academy.org/posts/versionamiento-de-base-de-datos>, s.f.)

Marco normativo

Ley No. 5096 sobre Estadísticas y Censos Nacionales, promulgada en 1959, establece en su **Artículo 1.** - Se establece un organismo técnico con el nombre de Dirección General de Estadística a cuyo cargo estará la recolección, revisión, elaboración y publicación de las estadísticas nacionales en relación con las actividades económicas, agrícolas, comerciales, industriales, financieras, sociales, condiciones de la población, censos nacionales, así como la coordinación de los servicios estadísticos de la República.

En enero de 2012, el Poder Ejecutivo promulgó la **Ley 1-12 sobre la Estrategia Nacional de Desarrollo 2030, Artículo 46.** Generación de información. La Oficina Nacional de Estadística es la instancia encargada de coordinar la generación de las informaciones nacionales necesarias para realizar la labor de monitoreo y evaluación de la Estrategia Nacional de Desarrollo 2030. Para tal fin, se deberá fortalecer el Sistema Estadístico Nacional y establecer estándares nacionales únicos para la generación de información confiable, oportuna y de uso colectivo.

Marco metodológico

Para el desarrollo de los Lineamientos para la Documentación del Diccionario de Datos, se llevaron a cabo las siguientes tareas:

Tarea 1. Revisión de las consideraciones establecidas en el Código Nacional de Buenas Prácticas para las Estadísticas Oficiales y los Lineamientos para el Proceso Estadístico en el Sistema Estadístico Nacional. Esta primera tarea se llevó a cabo para determinar la necesidad de establecer directrices y recomendaciones para llevar a cabo el diseño y estructuración de un diccionario de datos. Además, la revisión bibliográfica de documentos e instrumentos relacionados a guías y diseños para este.

Tarea 2. Diseño y construcción de un cuestionario para captura de la información relevante relacionada a la construcción de un DD, este cuestionario se aplicó a las diferentes áreas productoras de la Oficina Nacional de Estadística (Dirección de Censos y Encuestas, Dirección de Estadísticas Económicas, Dirección de Estadísticas Demográficas, Sociales y Ambientales, Departamento de Procesamiento de Datos, entre otras...).

Tarea 3. Revisión de la estructura de los Lineamientos y Recomendaciones para Documentar el Diccionario de Datos, en esta tarea se estableció la homologación de la estructura con otros lineamientos ya establecidos y difundidos, ofreciendo mejoras en el desarrollo, enfocando su diseño para ser incluido en el Sistema de Gestión Documental (SGD) de la Oficina Nacional de Estadística.

Tarea 4. Socialización y validación con instituciones que conforman el Sistema Estadístico Nacional que producen y difunden estadísticas oficiales, con la finalidad de recibir retroalimentación para su estructura y dar a conocer la importancia de contar y aplicar estos lineamientos en sus procesos de producción y difusión de sus metadatos relacionados a los diccionarios de datos.

Principios contemplados en el CNBPE relacionados con la Documentación del Diccionario de Datos



Principios relacionados del **Nivel A. Gestión del Entorno Institucional**, este nivel es de vital importancia en el reforzamiento de la credibilidad y eficiencia de las estadísticas oficiales. Los principios establecidos en este nivel que tienen vinculación directa con los Lineamientos y Recomendaciones para Documentar el Diccionario de Datos, son los siguientes:

Principio 1. Asegurar la imparcialidad y objetividad, CNBPE

- 1.1 Las fuentes de datos, conceptos, clasificaciones y metodologías se deben elegir de forma objetiva, considerando las mejores prácticas internacionales.

Principio 4. Asegurar la calidad, CNBPE

- 4.1 Debe de existir en la ONE y el SEN una política de calidad que siga los lineamientos de los Principios Fundamentales de las Estadísticas Oficiales.
- 4.2 La ONE debe promover la evaluación y mejora continua de los procesos y productos estadísticos.



En relación con el **Nivel B. Gestión del Proceso Estadístico**, la vinculación que este nivel tiene con los Lineamientos y Recomendaciones para Documentar el Diccionario de Datos se da a través de los siguientes principios:

Principio 6. Procedimientos estadísticos apropiados, CNBPE

- 6.1 Los metadatos, la documentación de los métodos y de los diferentes procesos estadísticos se deben gestionar en todos los procesos y compartirse de manera apropiada.



La vinculación existente entre los Lineamientos y Recomendaciones para Documentar el Diccionario de Datos y el **Nivel C. Gestión de Resultados Estadísticos** se establece a través de los siguientes principios:

Principio 13. Asegurar la accesibilidad y la claridad, CNBPE

- 13.1. Las estadísticas y sus metadatos deben presentarse de manera que se facilite la interpretación adecuada y las comparaciones significativas.

Principio 14. Asegurar la coherencia y la comparabilidad, CNBPE

- 14.1. Se debe utilizar estándares internacionales, regionales y nacionales con respecto a definiciones, unidades, variables y clasificaciones.

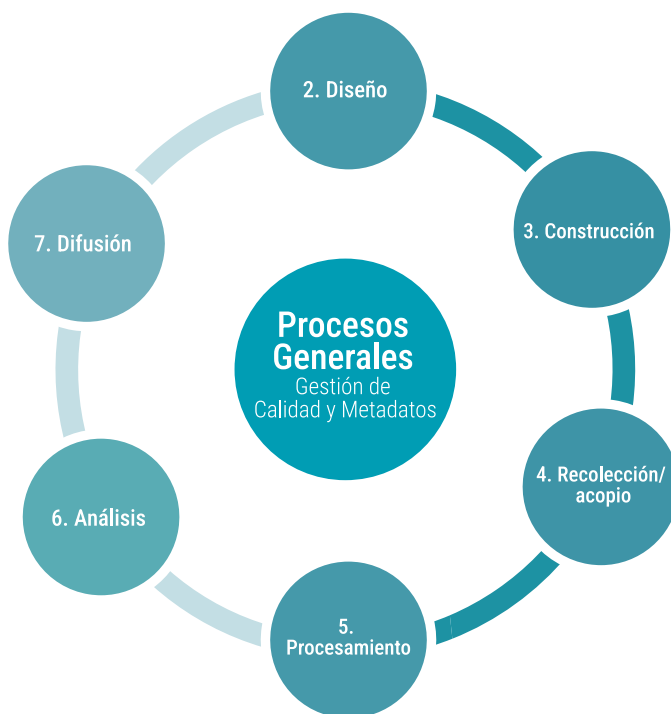
Principio 15. Gestión de los metadatos, CNBPE

- 15.1. Se debe encontrar definido y documentado correctamente el sistema de gestión de metadatos de la institución productora de estadística.
- 15.2. Los metadatos se deben documentar, archivar y difundir de acuerdo con las normas internacionalmente aceptadas.
- 15.3. Deben existir programas de capacitación y desarrollo del personal relacionados con la gestión de metadatos y con los sistemas de información y documentación.

Relación existente entre el Diccionario de Datos y el Proceso Estadístico

El proceso de documentación del DD se asocia a diferentes fases del proceso estadístico, según lo establecido en los Lineamientos para el Proceso Estadístico en el *Sistema Estadístico Nacional*¹¹. Aunque, el DD se diseña principalmente en las fases **2. Diseño** y **3. Construcción**, su uso se extiende a lo largo de todo el proceso estadístico, debido a que el DD es un documento vivo que debe ir actualizándose continuamente para garantizar la calidad y coherencia de la base de datos relacionada a este.

Ilustración 1. Vinculación del DD en el proceso estadístico:



En la Fase 2. Diseño, se relaciona durante el diseño de una Operación Estadística (OE), donde se definen los datos a recolectar / acopiar y su estructura, en esta fase, también se establecen las variables, clasificaciones, tipos de datos y relaciones, que luego pasan a ser documentadas en el DD.

En la Fase 3. Durante la construcción, se implementan las estructuras de datos en los sistemas informáticos, bases de datos o plataformas tecnológicas definidas. El Diccionario de Datos sirve aquí como guía para garantizar que las estructuras digitales respeten las especificaciones diseñadas: nombres de variables, tipos, longitudes, codificaciones, formatos y relaciones entre elementos.

Adicionalmente, en esta fase se prepara el soporte necesario para la **Fase 4. Recolección/Acopio**, asegurando que los instrumentos, canales y sistemas utilizados para captar la información estén alineados con las definiciones documentadas en el DD. Esto incluye la validación de campos, dominios aceptados, reglas de captura y controles de calidad automatizados que facilitan el cumplimiento de los estándares establecidos. En este sentido, la fase de construcción no solo habilita el almacenamiento y procesamiento, sino que también establece las condiciones técnicas que garantizan que la recolección se ejecute conforme al Diccionario de Datos.

¹¹ Los Lineamientos para el Proceso Estadístico en el Sistema Estadístico Nacional, es un documento descriptivo donde se establece el modelo de producción de estadísticas, describiendo las diferentes fases y subfases que existen dentro del proceso de producción de estas. [Lineamientos para el Proceso Estadístico en el Sistema Estadístico Nacional - Oficina Nacional de Estadística \(ONE\)](#)

En la Fase 5. Procesamiento, durante esta fase, el Diccionario de Datos se utiliza como referencia esencial para interpretar correctamente los datos recolectados, aplicar transformaciones y realizar validaciones conforme a las reglas y formatos establecidos. El DD proporciona la definición precisa de cada variable, su tipo, dominio válido, unidades de medida y codificación, lo que permite asegurar la calidad y coherencia de los datos en cada paso del procesamiento. Además, el Diccionario de Datos, se sustenta, en los procedimientos analíticos, que se desarrollan durante esta misma fase, facilitando así la trazabilidad y la comprensión de los datos transformados. Las tareas de análisis exploratorio, generación de indicadores y validación de resultados, que utilizan la información técnica contenida en el DD para interpretar, adecuadamente, las salidas del procesamiento y verificar que correspondan con las definiciones conceptuales y operativas previstas. En este sentido la Fase 6 el análisis se considera una extensión natural del procesamiento y está profundamente vinculado al uso correcto del Diccionario de Datos.

En la Fase 7. Difusión, finalmente, se asocia en el uso de las generaciones de los metadatos y tiene como propósito asegurar que los metadatos documentados acompañen y respalden técnicamente las publicaciones estadísticas. En este sentido, los metadatos deben estar alineados con los datos contenidos en la base de datos final, garantizando la coherencia entre lo que se publica y su correspondiente documentación técnica. El DD permite la interpretación correcta de los resultados dados y la garantía de su trazabilidad.

Es por esto por lo que, documentar el diccionario de datos es de vital importancia en la captura, procesamiento y difusión de resultados estadísticos ya que en el proceso se pueden identificar e implementar las mejoras necesarias, permitiendo resultados de calidad, comprensibles y utilizables por entidades y personas usuarias.

Generalidades

El **Diccionario de Datos** constituye una herramienta fundamental para la gestión de la información en cualquier sistema estadístico o informático. Se trata de una recopilación organizada de metadatos que permite describir, de forma estructurada, las características, relaciones y atributos de los datos contenidos en una base de datos o conjunto de información. Su propósito principal es **estandarizar la documentación**, facilitando así la comprensión, interoperabilidad y uso eficiente de los datos por parte de los distintos actores involucrados. Esta práctica se sustenta en marcos internacionales como la norma **ISO/IEC 11179-3:2013**, la cual promueve el uso de catálogos de elementos de datos como parte esencial de una gestión de datos orientada a la calidad.

Dentro de esta lógica, se distinguen dos tipos de diccionarios de datos según el entorno tecnológico en el que operan: **activos y pasivos**.

El **Diccionario de Datos Activo** se aplica en contextos donde se utiliza un **Sistema Gestor de Bases de Datos (SGBD)**. En estos entornos, el diccionario se genera y actualiza automáticamente a medida que se crean o modifican las estructuras de datos. Además, puede configurarse para reflejar atributos personalizados definidos por el usuario, permitiendo una gestión más dinámica y eficiente. Este tipo de diccionario es habitual en plataformas robustas que manejan grandes volúmenes de información estructurada de manera relacional. Así lo destaca la Agencia Digital de Innovación Pública (2021), al señalar que estos diccionarios activos permiten mantener actualizada la documentación de manera sincronizada con los cambios en la base de datos.

En contraste, el **Diccionario de Datos Pasivo** se desarrolla de manera **manual** y está pensado para contextos en los que los datos se almacenan fuera de un sistema gestor, como en archivos planos del tipo **XLSX, CSV o TXT**. En estos casos, es necesario que los responsables del conjunto de datos documenten explícitamente la estructura y características de las variables, garantizando así su interpretación y reutilización adecuada. Si bien requiere de un esfuerzo adicional, este tipo de diccionario es clave en muchos entornos institucionales donde aún se utilizan formatos no automatizados para el almacenamiento de información.

Ambos enfoques, aunque diferentes en su implementación, comparten un mismo fin: asegurar la calidad, transparencia y trazabilidad de los datos a través de una documentación estandarizada, confiable y accesible.

A través del DD es posible conocer, entre otros aspectos: el significado de cada campo o variable, el tipo de datos que contiene, los valores que puede tomar y el nivel de sensibilidad de la información.

Contar con un Diccionario de Datos adecuadamente documentado permite reducir ambigüedades, prevenir errores de interpretación y promover la interoperabilidad entre sistemas. Además, facilita el trabajo colaborativo entre equipos técnicos, mejora la trazabilidad de los datos, y fortalece los principios de transparencia y rendición de cuentas en la producción estadística. Para el DD, la calidad de las fuentes de información constituye un elemento determinante para la solidez técnica del Diccionario de Datos (DD) y, en consecuencia, para la calidad de todo el proceso estadístico. Evaluar dichas fuentes garantiza que, los elementos de datos documentados, reflejen con precisión la realidad que se pretende medir, evitando inconsistencias que puedan comprometer la interpretación, comparabilidad y utilidad de los resultados estadísticos.

En este sentido, la evaluación de calidad de las fuentes especialmente bases de datos, registros administrativos o sistemas de información permite identificar debilidades estructurales, problemas de cobertura, errores de codificación o ausencia de metadatos esenciales, lo cual es clave para asegurar la coherencia entre los datos utilizados y la documentación generada en el DD.

Además, una fuente de información validada contribuye a la trazabilidad de los datos, la interoperabilidad entre sistemas y la posibilidad de reutilización futura, elementos esenciales para una gestión moderna de la información estadística. Por tanto, incorporar este análisis como parte del ciclo de documentación no solo mejora la calidad técnica del DD, sino que también fortalece la transparencia y la rendición de cuentas en la producción estadística.¹²

Lineamientos Generales para la Documentación de un Diccionario de Datos

La etapa Inicial que compone este documento establece los principios basados en buenas prácticas y normas fundamentales para la elaboración y documentación del diccionario de datos, garantizando uniformidad, claridad y estructuración de la información dentro de este. La replicación de estos lineamientos asegura que los datos sean comprensibles, accesibles y, también, estandarizados para su correcta gestión e interoperabilidad.

Para la definición de los lineamientos, se ha contemplado el conjunto de etapas y subetapas claves para lograr una documentación clara y efectiva de los datos:

Ilustración 2. Vista General de actividades para la Documentación del Diccionario de Datos

Actividades Generales					
Gestión del Diccionario de Datos / Metadatos					
1. Plan de diseño	2. Definición técnica de variables y elementos	3. Construcción	4. Difusión	5. Gestión de mantenimiento	6. Evaluación y validación
1.1. Objetivo y alcance del DD	2.1. Inventario de variables	3.1. Herramientas para modelo de datos	4.1. Canales y medios de difusión	5.1. Revisión y actualización	6.1. Monitoreo y control de calidad
1.2. Responsabilidades y roles	2.2. Diseño de estructura	3.2. Centralización de documentación y herramienta	4.2. Medición	5.2. Capacitación y documentación	6.2. Evaluación de impacto
1.3. Normas y estándares	2.3. Normalización y Optimización	3.3. Asignación de permisos		5.3. Sistematización	6.3. Mejora continua
1.4. Revisión y validación	2.4. Revisión y validación			5.4. Versionamiento y control de cambios	

¹² Fuente de referencia: Guía práctica para la elaboración de diccionarios de datos – Agencia Digital de Innovación Pública, febrero 2021.

Etapas del proceso

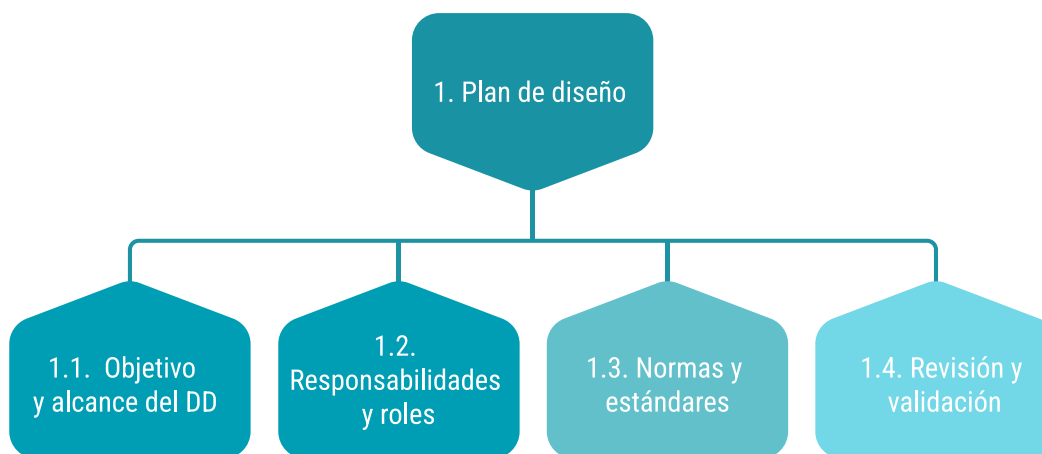
Las etapas y subetapas identificadas en este documento componen el proceso de documentación del DD, este conjunto de etapas y subetapas se han identificado en el proceso de levantamiento de información en las áreas que producen y difunden Metadatos e igualmente de distintos materiales que fueron utilizados como referencias para su diseño. Las cinco etapas del proceso de documentación se encuentran divididas en 18 subetapas que corresponden a las actividades a ejecutar durante el proceso de desarrollo del Diccionario de Datos de una Operación Estadística.

Estos lineamientos establecen directrices y recomendaciones basándose en las mejores prácticas establecidas para la documentación de los diccionarios de datos.

Etapa 1. Plan de diseño

La etapa de la planificación de diseño es de vital importancia para garantizar que el DD sea estructurado, funcional y útil para todas las personas involucradas en el diseño, entidades y personas usuarias. Esta etapa define todo lo relacionado al diccionario de datos, ejemplo; definición de objetivo, alcance, normas y estándares, diseño, construcción, recopilación, almacenamiento, gestión de la información, difusión, gestión del mantenimiento, codificación, gestión de la evaluación, entre otros aspectos claves.

Ilustración 3. Etapa y subetapas establecidas en el Plan de diseño.



1.1 Objetivo y alcance del DD

Esta subetapa da lugar al establecimiento del propósito inicial, alcance y aplicabilidad que tendrá el DD con el fin de identificar y garantizar que su desarrollo responda a los requerimientos de la Operación Estadística (OE) o del sistema de información al que será vinculado.

El objetivo define qué se espera lograr con su implementación, asegurando que sea un recurso funcional para la correcta interpretación, gestión y reutilización de datos, lo que permitirá una mejor comprensión de la información para fortalecer la toma de decisiones, facilitar el análisis y promover su uso eficiente. Las instituciones productoras deben tomar en cuenta:

- a. ¿Para qué se crea el DD?
- b. ¿Por qué se crea el DD?
- c. ¿Quiénes serán los grupos de interés (stakeholders) del DD?
Nota: La identificación de los grupos de interés dependerá del carácter de los datos documentados en el DD, es decir si son público o privado.
- d. ¿Cómo se beneficiará el proceso, la institución u Operación Estadística con ese DD?

Nota: En esta pregunta se debe considerar a los actores internos que apoyarían en la construcción del DD.

Al establecer el alcance del DD se define los límites de este, es decir, qué se incluirá y qué quedará omitido. Al momento de definirlo, las instituciones productoras deben tomar en cuenta lo siguiente:

- Establecer la cobertura¹³ del DD donde se describen y asociación de todos los componentes asociados a la Base de Datos que incluyan sistemas o módulos involucrados, las fuentes de datos a considerar y los elementos documentados (ejemplo: tablas, bases de datos, atributos, relaciones, otros...).
- Establecer el nivel de detalle que tendrá el DD, es decir: Definiciones básicas o avanzadas, relación entre datos y formatos de datos (ejemplo: incluir identificación clara de cada elemento de datos, su tipo, formato, valores, solo lectura de datos o también reglas de transformación, llaves foráneas, integridad referencial, tipo de datos, longitud, otros...).
- Definir exclusiones (fuera de alcance), se debe establecer la exclusión de datos temporales o no relevantes, igualmente, para todos los sistemas o módulos no documentados en el diccionario.

1.2 Responsabilidades y roles

En la documentación del DD, la subetapa que hace referencia y que debe establecer las responsabilidades y roles es clave para garantizar la correcta gestión y trazabilidad del DD, su buen mantenimiento y uso, asegurando la garantía de la calidad de este.

Esta subetapa debe definir con claridad las responsabilidades de la documentación del DD, incluyendo las tareas de creación, actualización, validación, difusión y consulta de los elementos que lo componen. La asignación de roles asegura que la información documentada sea confiable y esté alineada con los estándares estadísticos nacionales e internacionales. Las instituciones productoras deben tomar en cuenta lo siguiente:

Tabla 1. Responsabilidades y roles en la gestión del DD.

Rol	Descripción	Permisos
Responsable de la gobernanza de los datos	Supervisar la aplicación de normativas y estándares nacionales e internacionales (ejemplo: clasificaciones, ISO 11179, IDD, SDMX, otros...), definir las políticas para la gestión del DD y asegurar la coherencia y alineación del DD con otros sistemas de información estadística.	Supervisar y definir las políticas para la gestión del DD
Administrador	Ejecutar las tareas de mantenimiento y actualización del DD, coordina con los equipos técnicos y de análisis para validar definiciones y estructuras y gestiona las solicitudes que se presenten para modificar el DD (ejemplo: inclusión o exclusión de elementos).	Acceso total (crear, modificar y eliminar)
Especialista en documentación de datos y Metadatos	Define y gestiona la estructura de metadatos asegurando su correcta documentación y coherencia.	Acceso de edición
Analista Calidad de Datos	Evaluar la integridad, consistencia y precisión de los datos dentro del DD.	Acceso para evaluar
Usuario de Consulta	Acceso para la búsqueda y uso de información disponible dentro del DD.	Solo lectura
Auditor/Revisor	Revisar la calidad, cumplimiento de necesidades y generación de auditorías.	Lectura con permiso para comentarios

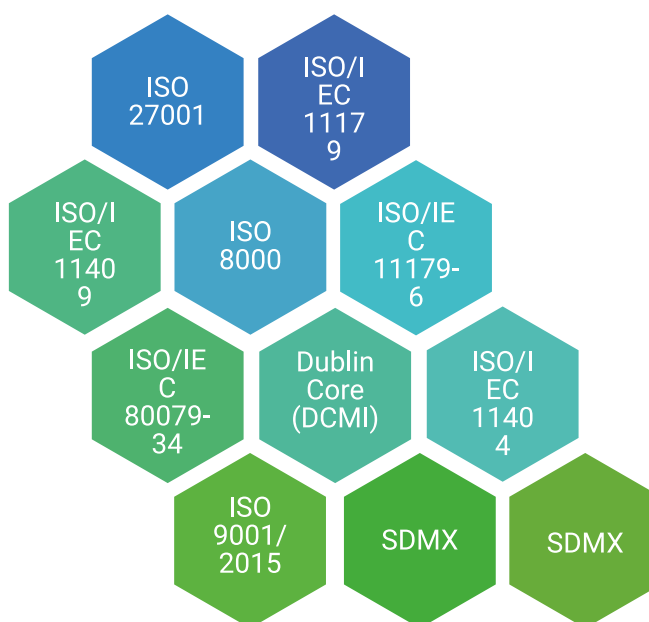
¹³ Ejemplo de Cobertura de un DD: "Este DD abarca la documentación de los elementos de datos utilizados en la base de datos del registro civil de la Junta Central Electoral (JCE), incluyendo nombres de tablas, columnas, tipos de datos, relaciones y reglas establecidas por la entidad. No incluye información de sistemas externos ni detalles sobre procesos ETL."

Adicional a los roles establecidos en la tabla anterior, se debe establecer un equipo de soporte y capacitación: estos deben proporcionar formación a las entidades y personas usuarias sobre el uso y mantenimiento del DD y brindar soporte a los que consultan la información dentro del DD (en caso de requerirlo).

1.3 Normas y estándares

Al momento de planificar el diseño del Diccionario de Datos (DD), es fundamental definir qué normas y estándares serán aplicados para asegurar la consistencia, la calidad técnica y la alineación del DD con otras herramientas de documentación utilizadas en el sistema estadístico. La adopción de estándares nacionales e internacionales (como la norma ISO 11179 o el estándar SDMX) permite fortalecer la estructura y utilidad del diccionario, facilitando su integración como instrumento de apoyo a procesos de análisis, difusión y gestión de metadatos. Las instituciones productoras deben tomar en cuenta el uso de estándares tales como:

Ilustración 4. Referencias de normas internacionales



La planificación y gestión del Diccionario de Datos (DD) puede fortalecerse significativamente mediante la adopción de estándares internacionales que aseguren su integridad, utilidad y alineación con buenas prácticas.

- **ISO/IEC 11179** proporciona el marco central para la gobernanza y gestión de metadatos, mientras que su parte **11179-6** detalla los requisitos para el registro estructurado de estos metadatos. Para asegurar la claridad semántica,
- **ISO/IEC 11404** define tipos de datos genéricos, apoyando una correcta clasificación dentro del DD.
- **ISO 8000**, enfocada en la calidad de los datos, complementa al DD al ofrecer lineamientos para la verificación de precisión, completitud y consistencia de los elementos documentados.
- **ISO 27001**, centrada en la seguridad de la información, es clave para proteger la confidencialidad, integridad y disponibilidad de los datos registrados en el DD.
- **ISO/IEC 80079-34**, aunque desarrollada para ambientes industriales, ofrece principios relevantes para la gestión de documentación técnica compleja.

- La aplicación de un enfoque transversal de calidad bajo **ISO 9001:2015** permite que el DD se integre a un sistema de gestión institucional, garantizando su mantenimiento continuo.
- En cuanto a la interoperabilidad y el intercambio de información, el estándar **SDMX** permite que los diccionarios sean compatibles con plataformas estadísticas nacionales e internacionales. Finalmente, **Dublin Core** (DCMI) es útil en contextos donde se desea una estandarización mínima para catalogar contenidos informacionales, aportando valor a los metadatos descriptivos del DD.

En relación con el uso y definición de estándares de datos, estos se deben de establecer de manera interna y externa para la correcta estructura y gestión del DD. Las instituciones productoras deben tomar en cuenta el uso de estándares tales como:

Ilustración 5. Definición de estándares de datos



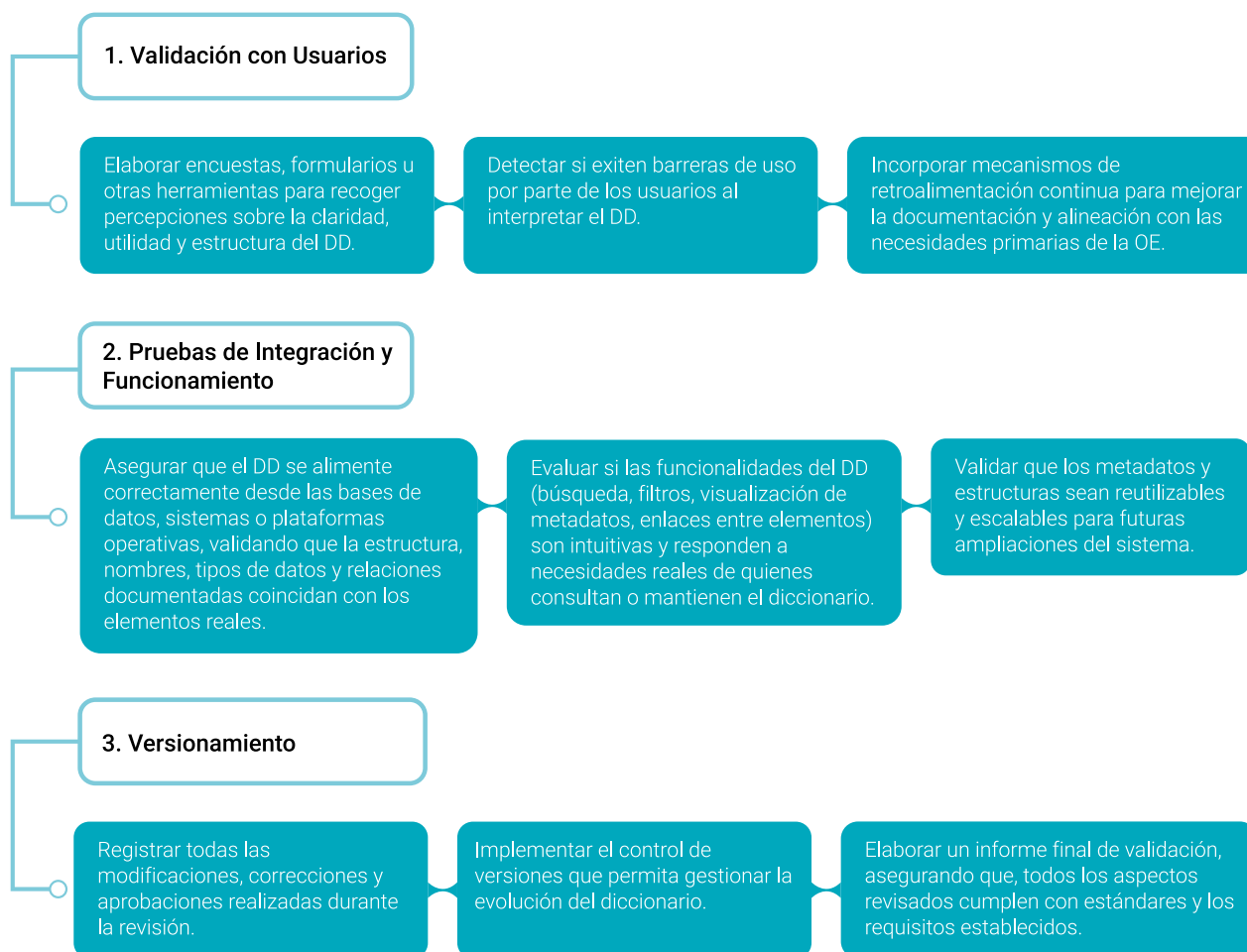
Nota: ver Anexo 1. Detalles de componentes establecidos en Ilustración 5.

Asegurar que el DD sea robusto, confiable y que se encuentre alineado con las mejores prácticas nacionales e internacionales conlleva que, en esta subetapa, se establezcan las normas y los estándares adecuados.

1.4 Revisión y validación

La última subetapa establecida dentro de la etapa del Plan de Diseño es la revisión y validación, en esta subetapa se debe tomar en cuenta seis aspectos claves, los cuales permiten garantizar la calidad, coherencia y utilidad del diccionario. Las instituciones productoras deben tomar en cuenta:

Ilustración 6. Aspectos claves para la garantía de la calidad, coherencia y utilidad del DD.

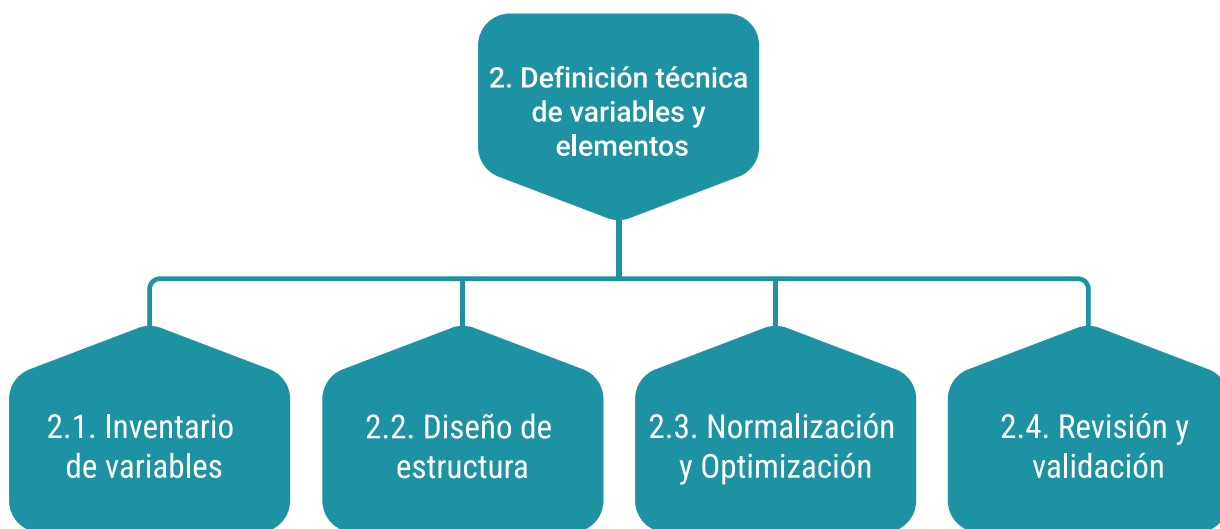


Revisar y validar el diseño del DD es esencial para garantizar que este sea preciso, confiable y alineado con las necesidades establecidas en la Operación Estadística en cuestión.

Etapa 2. Definición técnica de variables y elementos

La etapa de definición técnica de variables y elementos busca identificar, organizar y documentar todos los elementos de datos que formarán parte del DD. Para ello, se recopila información técnica clave sobre cada variable, como su nombre, tipo de dato, longitud, definiciones operativas y relaciones con otros elementos. Asimismo, se documentan los metadatos esenciales asociados como: fuente de origen, periodicidad, nivel de desagregación y unidad de medida, asegurando la completitud, coherencia e integridad de los elementos definidos. Esta etapa también promueve el involucramiento de actores clave, con el fin de garantizar que las definiciones estén alineadas con los requerimientos del proceso estadístico y los estándares establecidos.

Ilustración 7. Etapa y subetapas establecidas para la definición técnica de variables y elementos.



2.1 Inventario de variables

La subetapa que compone el inventario de variables juega un papel esencial en la identificación, documentación y estandarización de los elementos de datos que componen el sistema de información. Esta subetapa permite establecer una base estructurada y organizada de todas las variables utilizadas en los procesos estadísticos, administrativos y operacionales, asegurando su coherencia, interoperabilidad y alineación con los estándares de calidad de datos.

Crear un catálogo detallado que facilite la gestión, análisis y reutilización de los datos en las instituciones productoras es uno de los objetivos principales de contar con un inventario de variables. Por lo cual, para lograrlo se requiere una identificación rigurosa de cada variable, definiendo sus atributos esenciales como nombre, tipo de dato, formato, unidad de medida, dominio de valores y relaciones con otras variables. La adecuada documentación de estas características permite garantizar la claridad en la interpretación de los datos, reduciendo redundancias y evitando inconsistencias en su uso.

Es por esto que, tener un inventario de variables sienta las bases esenciales para un DD bien estructurado. Para desarrollar un inventario de variables robusto, las instituciones productoras deben tomar en cuenta los siguientes siete requisitos claves:

Ilustración 8. Requisitos claves para estructurar un inventario de variables robusto.



- 1. Identificar y clasificar las variables:** Es el proceso mediante el cual se reconocen todas las variables utilizadas en los procesos estadísticos y se agrupan por tipo, función, nivel de desagregación, unidad de análisis o frecuencia.
- 2. Establecer definiciones claras y estandarizadas:** Descripción precisa y no ambigua de cada variable, alineada con marcos conceptuales institucionales, nacionales o internacionales que faciliten su interpretación y uso coherente. Esta definición debe incluir el nombre técnico, tipo de dato, formato, longitud, unidades de medida, valores válidos, restricciones y frecuencia de actualización.
- 3. Documentación de atributos esenciales:** Registro técnico y estructurado de los metadatos clave asociados a cada variable, tales como nombre estándar, tipo de dato, formato, longitud, unidad de medida, dominios de valores permitidos, reglas de validación, restricciones de uso y frecuencia de actualización, con el fin de garantizar su correcta interpretación, uso y mantenimiento.
- 4. Relacionamiento con otras variables:** Establecer vínculos entre variables que se derivan, se componen o están asociadas, lo que mejora la trazabilidad y comprensión del sistema de datos.
- 5. Evaluar la calidad de las variables:** Verificar la completitud, coherencia y precisión de cada variable y su documentación, asegurando que cumpla estándares mínimos de calidad.
- 6. Asegurar la interoperabilidad y comparabilidad:** Atributo que garantiza que las variables puedan ser utilizadas en conjunto con otras fuentes de datos o sistemas, asegurando su integración técnica y su coherencia conceptual en contextos comparativos.
- 7. Definir mecanismos de seguridad y acceso:** Establecer niveles de permiso para acceder, editar o eliminar información sobre las variables, de acuerdo con su sensibilidad y uso institucional.

A través de la aplicación de directrices y mejores prácticas en el proceso de estructuración de las variables, se sientan las bases para la construcción de un DD bien estructurado, robusto, accesible y alineado a estándares internacionales para la gestión y calidad de datos.

2.2 Diseño de estructura

La subetapa que define el diseño de la estructura del DD es neurálgica dentro de su desarrollo, pues, esta define su organización, representación y disposición de los elementos de datos dentro del sistema. Contemplar el diseño de una estructura bien elaborada permite una documentación clara, accesible y estandarizada, brindando mayor facilidad en la gestión, interoperabilidad y uso eficiente de los datos a capturar según la Operación Estadística en cuestión.

Contemplar el diseño de una estructura coherente y alineada con estándares internacionales mitiga errores, mejora la calidad del dato y permite la integración con otros sistemas de información, agregando a esto, facilidad en la comprensión de los datos por parte de entidades y personas usuarias.

Ilustración 9. Beneficios esperados de un diseño estructurado del DD:



Para la realización de una estructura robusta de un DD, las instituciones productoras deberán tomar en cuenta lo siguiente:

- Definir y establecer una clasificación jerárquica de los datos
- Establecer un orden lógico por categoría, tales como datos administrativos, estadísticos, operativos, financiero, otro...
- Establecer y asignar códigos únicos para la identificación de cada elemento.
- Adicionar a cada elemento del DD la siguiente información mínima estructurada: nombre de variable, definición (variable o campo), tipo de dato, dominio de valor (valores permitidos o rango de datos aceptables), formato, fuente, periodicidad, reglas de validación y relación con otras variables.
- Asegurar la implementación de identificadores únicos para las Bases de Datos.
- Utilizar marcos de referencias (estándares internacionales) para la estructura del DD.
- Diseñar estructura de fácil acceso, búsqueda y recuperación de información.
- Asegurar una estructura que permita la interoperabilidad con otros sistemas.
- Establecer controles de calidad y seguridad (alineados a normas nacionales e internacionales).
- Proporcionar un diseño que garantice una documentación comprensible para las entidades y personas usuarias (técnicos y no técnicos).

2.3 Normalización y Optimización

A través de esta subetapa, se establecen reglas y procedimientos que permiten mitigar el riesgo de redundancias, mejorar la precisión de los datos y optimizar su almacenamiento y gestión. Normalizar y optimizar el DD permite contribuir a la consistencia, interoperabilidad y facilidad de uso a la información. Las instituciones productoras deben tomar en cuenta:

- a. Aplicar estándares de normalización que permitan la adopción de nomenclaturas uniformes para nombres de variables, evitando ambigüedades y redundancias.
- b. Utilizar catálogos y códigos estandarizados.
- c. Detectar y consolidar duplicidad de variables.
- d. Identificar y documentar los datos obsoletos o en desuso, señalando su estado dentro del Diccionario de Datos para evitar su incorporación activa en procesos estadísticos y mantener la vigencia del inventario.
- e. Aplicar mecanismos de validación para la corrección de errores en el DD.
- f. Implementar versionamiento de datos que permita el seguimiento de cambios y actualizaciones.
- g. Asegurar la compatibilidad de los datos con diferentes plataformas y sistemas.

2.4 Revisión y validación

La subetapa de Revisión y Validación es la última etapa establecida dentro de la etapa de Definición técnica de variables y elementos. Esta subetapa es esencial en la construcción del DD, ya que permite asegurar la precisión, coherencia y calidad de la información recopilada. Mediante este proceso, se analizan los elementos de datos para identificar posibles errores, inconsistencias o duplicidades, garantizando que el DD cumpla con los estándares establecidos dentro de su diseño.

Esta subetapa garantiza mitigar los riesgos de errores en el procesamiento de datos, mejorar la interoperabilidad y asegurar una toma de decisiones basada en información producida con la calidad esperada. Las instituciones productoras deben tomar en cuenta:

- a. Evaluar la precisión y consistencia de las variables, documentando en el DD los tipos de datos y formatos requeridos para cada una, asegurando que sean coherentes y adecuados para su uso en los sistemas asociados.
- b. Identificar y eliminar las inconsistencias en el DD.
- c. Aplicar controles de calidad para las variables y elementos y, verificar, la exactitud, completitud e integridad de estos.

Etapa 3. Construcción

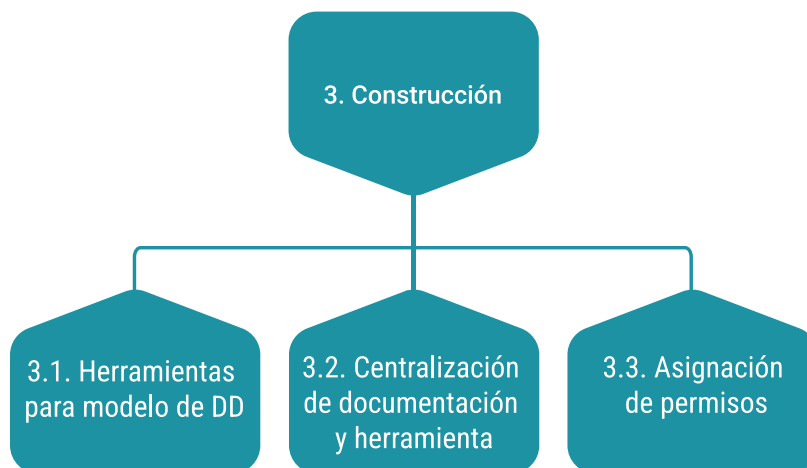
La etapa que contempla la construcción del DD es el proceso donde se materializa toda la información recopilada, normalizada y validada en una estructura organizacional y funcional. En esta etapa, se consolidan las definiciones, atributos y relaciones de cada variable que se ha identificado para asegurar que el DD sea preciso, accesible y que este se encuentra alineado con los estándares de calidad e interoperabilidad definidos.

Asimismo, durante esta etapa deben documentarse las relaciones clave entre variables, entidades o estructuras de datos, como las llaves primarias y foráneas, cuando corresponda. Esto permite representar adecuadamente la organización lógica de los datos, garantizando la integridad referencial, la unicidad y la trazabilidad de la información registrada en el Diccionario de Datos.

Un diseño eficiente del DD genera una construcción robusta del mismo, lo que facilita su consulta, mantenimiento y actualización, también permite que las entidades y personas usuarias accedan y utilicen las bases de datos de manera eficiente y debidamente estructuradas.

La siguiente ilustración presenta las diferentes subetapas que se han establecido para la construcción de un DD de calidad, robusto, confiable, accesible para entidades y personas usuarias y que permita la gestión y análisis de datos.

Ilustración 10. Etapa y subetapas establecidas para la construcción del DD.



3.1 Herramienta para modelo del Diccionario de Datos

Esta subetapa se inserta dentro de la fase de construcción del Diccionario de Datos (DD), y tiene como propósito central seleccionar, configurar y utilizar herramientas tecnológicas que permitan estructurar, documentar y gestionar el modelo del DD de forma eficiente y estandarizada.

El modelo del DD representa la estructura lógica y técnica de todos los elementos de datos que forman parte de una operación estadística, incluyendo variables, dominios, clasificaciones, relaciones, formatos y restricciones. Por tanto, la herramienta elegida debe responder no solo a necesidades técnicas, sino también a las exigencias documentales y normativas del entorno estadístico institucional.

Estas herramientas no solo permiten visualizar modelos relacionales, sino también facilitar la interoperabilidad, automatizar procesos de validación de metadatos y asegurar la trazabilidad entre el diseño, la implementación en bases de datos, y la documentación de salida.

Esto incluye nombre, tipo, longitud, dominio, unidad de medida, relaciones con otras variables, notas técnicas y referencias, garantizando así su integración con el contenido del DD.

Las instituciones productoras deben tomar en cuenta:

Se debe priorizar que la herramienta soporte estructuras alineadas con ISO/IEC 11179 (registro de metadatos), SDMX (intercambio de datos estadísticos) o IDD (normas de documentación del Banco Mundial).

Esto implica capacidad para exportar modelos a formatos interoperables (XML, SQL, XMI), sincronizarse con bases reales y permitir la reutilización de modelos para distintos dominios temáticos u operaciones estadísticas.

3.2. Centralización de documentación y herramienta

Para evitar problemas sobre la dispersión en diferentes documentos, bases de datos y sistemas se crea la subetapa de Centralización de documentación y herramientas en la etapa de construcción del DD, esta subetapa busca unificar y estructurar toda la documentación en un solo lugar, permitiendo que los datos sean accesibles, se encuentren organizados y sean fáciles de mantener.

Que las instituciones productoras centralicen sus DD no solo les permite mejorar la eficiencia en la gestión de datos, sino que también garantiza su seguridad, integridad y estandarización, para lograrlo, es clave que las instituciones productoras del SEN elijan herramientas adecuadas que permitan automatizar procesos, optimizar consultas y garantizar la actualización del DD de manera ágil y confiable, las siguientes pautas deben ser implementadas por las instituciones productoras:

- a. Centralizar en un repositorio único y estructurado que permita integrar la documentación del DD junto a la base de datos en una plataforma centralizada (base de datos, sistemas de gestión documental, otros...).
- b. Implementar acceso controlado y gestión de permisos, asegurando que la información sea consultable y modificable por los responsables del DD.
- c. Elegir la herramienta adecuada (software) que permita la creación, consulta y actualización automática del DD (Power BI Data Catalog, Talend Data Catalog, otros...).
- d. Garantizar que la herramienta utilizada brinde la garantía del versionamiento, trazabilidad de cambios y auditorías.
- e. Asegurar que el formato siga normas y estándares internacionales como el SDMX, DDI, otros...

3.3. Asignación de permisos

Tener los datos de la institución productora en un solo lugar es eficiente para el proceso, pero que cualquier persona pueda editarlos, eliminarlos o acceder a información sensible sin restricciones, representa un alto riesgo para la seguridad de la información y la generación de errores que conllevan a la distorsión del DD.

Para reducir los riesgos, la subetapa de Asignación de Permisos es clave en el desarrollo del DD. Su principal objetivo es determinar quién tiene la capacidad de visualizar, modificar o gestionar la información, asegurando así que solo las personas apropiadas tengan acceso a ciertos datos. Esta subetapa es clave para asegurar una gestión eficiente del sistema y proteger la información, lo que a su vez garantiza la integridad, la trazabilidad y el cumplimiento de las normativas de seguridad. A continuación, se presentan las pautas que deben seguir las instituciones productoras en el establecimiento y asignación de los permisos:

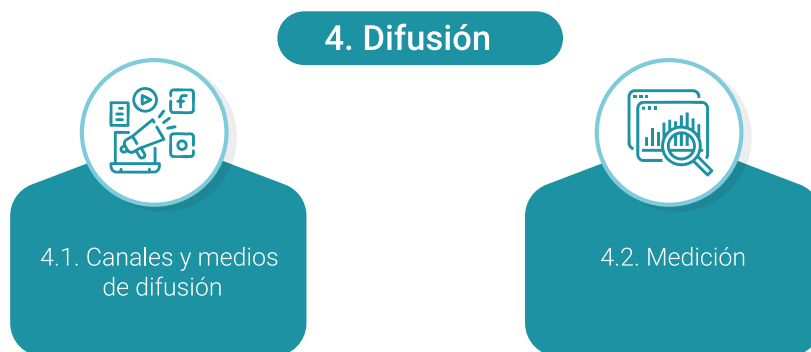
- a. Definir a cada rol los niveles de acceso o categorías de personas usuarias según su función (administradores, editores, lectores auditores, otros...), para cada categoría la asignación de permisos va desde el control total del DD, modificación o simplemente lectura.
- b. Implementar métodos y sistemas de autenticación para mejorar el control de acceso.
- c. Asegurarse de que cada usuario solo pueda acceder a la información necesaria de acuerdo a su rol.
- d. Definir fechas de expiración para los permisos de acceso de los colaboradores en cada área.
- e. Llevar un registro de accesos, modificaciones, tiempos y la recuperación de versiones anteriores.
- f. Crear políticas y normas que garanticen la seguridad, accesibilidad y protección de datos.
- g. Automatizar la gestión de permisos e introducir flujos de aprobación para los cambios de acceso, asegurando y validando las credenciales antes de otorgarlos.

Etapa 4. Difusión¹⁴

Esta etapa tiene como finalidad asegurar que el DD esté disponible, sea conocido, accesible, comprendido y reutilizable por los distintos grupos de interés, promoviendo su consulta y aprovechamiento para fines técnicos, metodológicos, analíticos y de interoperabilidad. Esta etapa es clave en la garantía de que el DD no solo esté disponible, sino que también sea comprendido y aprovechado por quienes trabajan con la información en la institución productora.

Esta etapa tiene como objetivo principal dar a conocer el DD de manera clara y accesible, asegurando que su contenido llegue a las entidades y personas usuarias de la información.

Ilustración 11. Etapa y subetapas establecidas para la Difusión del DD.



4.1. Canales y medios de difusión¹⁵

La subetapa de Canales y medios de difusión establecida en la etapa de difusión del DD busca equilibrar la facilidad de consulta con la protección de la información, garantizando que el diccionario de datos sea accesible para quienes lo necesiten, sin comprometer esto su integridad, el acceso debe ser seguro. La base de datos deberá estar acompañada siempre del DD.

Para garantizar la trazabilidad, interpretación y uso correcto del DD, se recomienda que toda versión difundida esté claramente vinculada a la base de datos (BD) o a la operación estadística (OE) que documenta. Esta vinculación debe estar explícita en el encabezado o metadatos del DD, permitiendo identificar su fuente de origen. Asimismo, no debe difundirse ningún DD que no cuente con esta referencia claramente definida.

De tal manera, las instituciones productoras deben establecer los mecanismos y medios adecuados para garantizar una difusión efectiva del DD. Estos canales deben ajustarse a las capacidades tecnológicas institucionales, el tipo de usuarios y el propósito del DD:

Canales institucionales:¹⁶

- Portal Web Institucional: publicación en la sección de metadatos, biblioteca técnica o apartado de transparencia.
- Catálogo Nacional de Metadatos: si existe un catálogo centralizado a nivel del SEN.
- Repositorios documentales internos: accesibles mediante autenticación para usuarios técnicos internos.
- Plataformas de datos abiertos: si el DD está vinculado a microdatos anonimizados o bases abiertas.
- Publicaciones especializadas: boletines, informes técnicos, fichas metodológicas.
- Correo institucional y circulares internas: para difusión dirigida.

¹⁴ PRINCIPIOS DE LA DIFUSIÓN DEL DICCIONARIO DE DATOS

- El DD es un **producto técnico independiente**, pero **complementario** a la base de datos o conjunto de microdatos.
- Su difusión debe **respetar los principios de acceso, seguridad, trazabilidad y confidencialidad** de la información.
- La difusión del DD debe **garantizar su integridad, comprensión y disponibilidad** a todos los perfiles de usuarios autorizados.

¹⁵ Se recomienda la aplicación de los Lineamientos para la Difusión de Resultados Estadísticos en el Sistema Estadístico Nacional.

¹⁶ Se recomiendan los siguientes formatos: PDF, HTML, Documento en XLSX o CVS (estructura tabular para reutilización técnica) y cualquier otro que se ajuste a las necesidades de las entidades y personas usuarias.

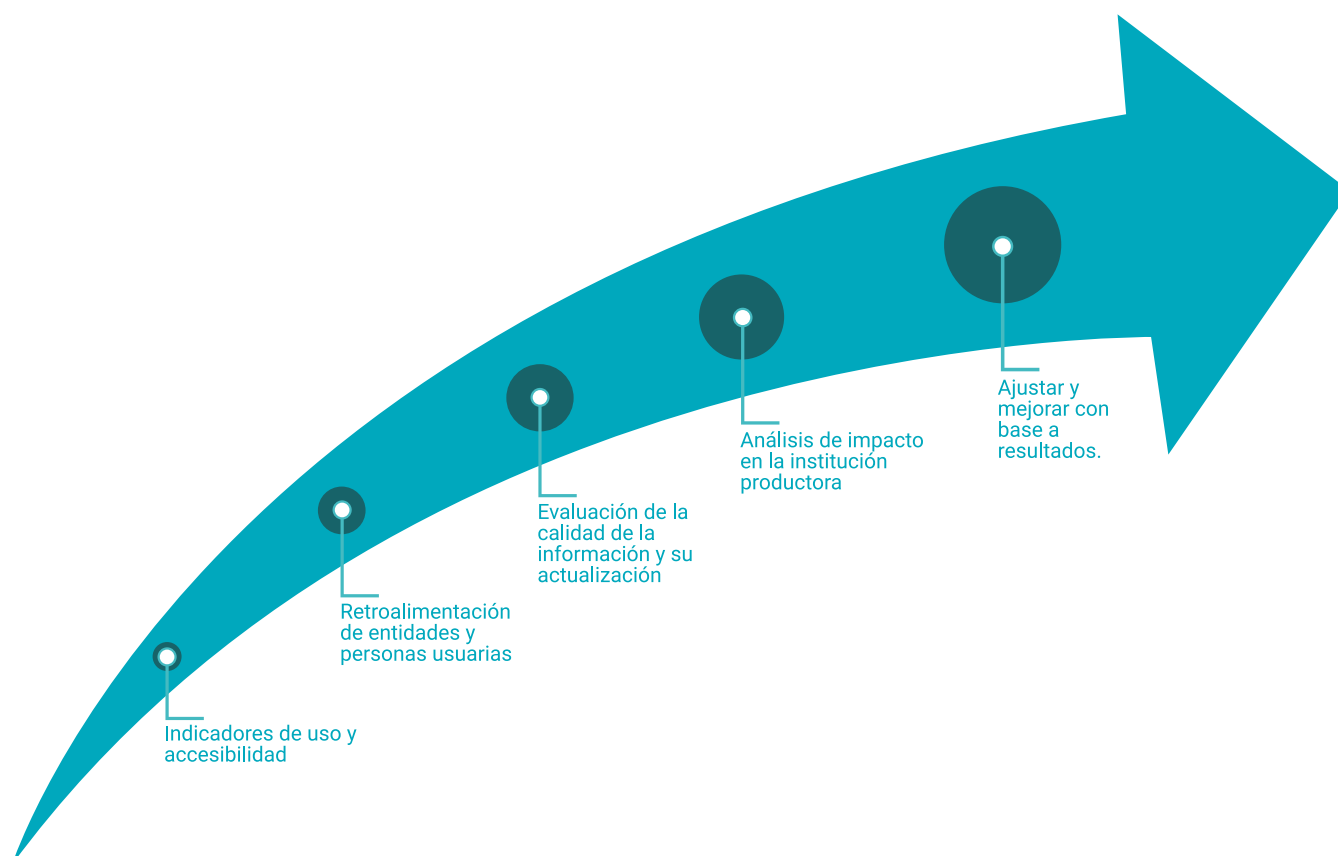
4.2 Medición

El DD además de difundirlo, compartirlo o darlo a conocer, debe asegurarse de que realmente está siendo utilizado de manera efectiva por las entidades y personas usuarias. Esta subetapa permite evaluar qué tan accesible, comprensible y útil ha sido el DD para las entidades y personas usuarias, identificando con las retroalimentaciones oportunidades de mejora, para garantizar que este cumpla con su propósito. Es necesario establecer métricas y mecanismos de seguridad que permitan entender como interactúan las entidades y personas usuarias con el DD. Esto incluye analizar la frecuencia de uso, los accesos, la retroalimentación de entidades y personas usuarias y la calidad de la información consultada. Con estas evaluaciones, se pueden hacer las mejoras optimizando que se asegure, que el DD, agrega valor a la institución productora.

La subetapa 4.1. Medición se vincula directamente con la fase de Evaluación del Proceso Estadístico, ya que permite valorar la efectividad del Diccionario de Datos (DD) como herramienta operativa y documental dentro del ciclo estadístico. Esta medición no solo se limita a verificar su existencia o difusión, sino que se enfoca en evaluar su uso real por parte de las entidades y personas usuarias. A través del análisis de métricas como la frecuencia de acceso, los patrones de consulta, la retroalimentación recibida y la utilidad percibida, se identifican oportunidades de mejora que retroalimentan tanto la calidad del DD como la eficiencia del proceso estadístico en su conjunto. De este modo, se garantiza que el DD no sea un producto estático, sino un recurso vivo que agrega valor, fortalece la transparencia, y asegura la trazabilidad y comprensión de los datos producidos.

En la siguiente ilustración se muestran los cinco criterios claves que se han establecido para la medición de la difusión del DD.

Ilustración 12. Criterios establecidos para la medición de la difusión del DD.



Al establecer los cinco criterios antes mencionados, se espera que el DD sea realmente útil y que esté en su máxima optimización según las experiencias de las entidades y personas usuarias. Basado en los cinco criterios ya establecidos, las instituciones productoras deben tomar en cuenta:

- a. Establecer indicadores para el uso y la accesibilidad para medir el número de accesos y consultas que se realizan al DD en un periodo determinado.
- b. Identificar qué secciones o elementos son más consultados para entender su relevancia y evaluar el tiempo de permanencia en la plataforma por parte de las entidades y personas usuarias.
- c. Capturar retroalimentación de entidades y personas usuarias con la aplicación de encuestas de satisfacción que midan la experiencia de estos.
- d. Habilitar un canal que permita capturar recomendaciones para mejora o reporte de problemas por parte de entidades y personas usuarias.
- e. Evaluar la calidad de la información y su actualización verificando si los datos documentados están actualizados y alineados con las necesidades de la Operación Estadística en cuestión.
- f. Determinar si los términos y definiciones siguen siendo claros y relevantes para las entidades y personas usuarias.
- g. Analizar el impacto en la institución productora evaluando si el uso y acceso del DD ha facilitado la toma de decisiones.
- h. Ajustar y mejorar con base a los resultados obtenidos a través de la retroalimentación de entidades y personas usuarias.
- i. Implementar mejoras en la interfaz, estructura o contenido del DD según los resultados obtenidos de la retroalimentación.
- j. Reforzar la capacitación o difusión si se detecta que las entidades y personas usuarias desconocen algunas funcionalidades del DD.

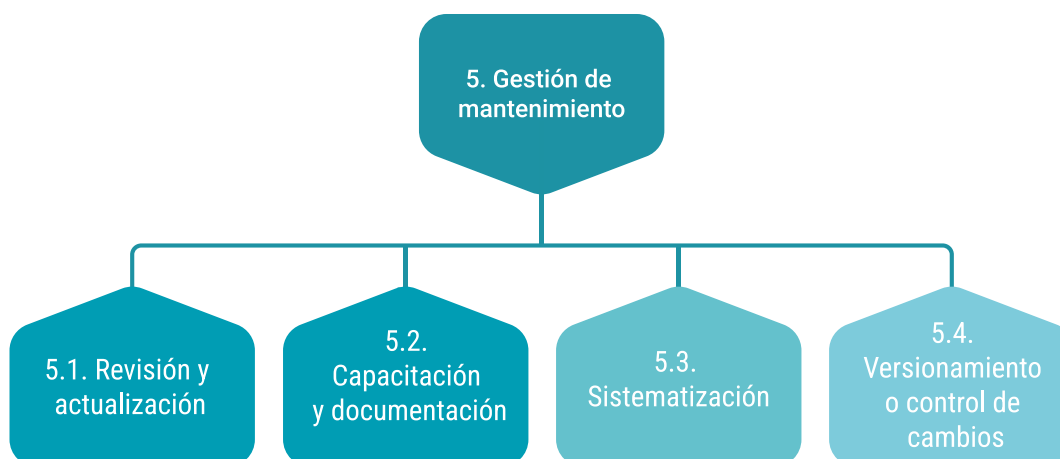
Etapa 5. Gestión de mantenimiento

Elaborar un DD es solo el comienzo. Para que realmente el DD sea de utilidad, debe mantenerse actualizado, relevante y alineado con las necesidades definidas. Es por esto por lo que, la etapa que se ha orientado a este proceso es la etapa para la gestión de mantenimiento, la cual se encarga de preservar la calidad y vigencia del diccionario, asegurando que refleje con precisión la información que respalda los procesos y decisiones.

Definir estrategias para mantener la constancia en el mantenimiento del DD, ya que los datos evolucionan incorporándose nuevas variables, cambian de definiciones, surgen nuevas relaciones y otros son obsoletos. Sin el establecimiento de una buena estrategia para el mantenimiento, el DD puede volverse confuso, desactualizado o incluso contradictorio.

La siguiente ilustración presenta la etapa y subetapas que se han establecido para la gestión del mantenimiento del DD.

Ilustración 13. Etapa y subetapas establecidas para la Gestión del Mantenimiento del DD.



5.1. Revisión y actualización

A medidas que la institución incrementa la producción, también lo hacen sus datos, estructuras y necesidades. La subetapa establecida para la revisión y actualización garantiza que el DD se mantenga alineado con los procesos actuales de la institución, refleje cambios en la información y siga siendo una herramienta confiable y útil para quienes lo consulten (entidades y personas usuarias).

Llevar a cabo el proceso de revisión y actualización del DD mitiga los riesgos de que este se vuelva obsoleto, e incluso, confuso e inexacto. Esta subetapa establece procesos donde se debe verificar, validar y ajustar el DD, asegurando que cada dato registrado siga cumpliendo con los estándares de calidad y utilidad, evitando que se vuelva una herramienta estática y se convierta en una herramienta viva, precisa y en constante mejora continua. Las instituciones productoras deben tomar en cuenta:

- Definir, establecer e implementar procesos de revisión periódica (mensual, trimestral, semestral, otros...).
- Comprobar que los datos documentados sigan las normas y estándares establecidos.
- Actualizar la información basada en cambios establecidos y a nuevas necesidades incorporadas.
- Evaluar si las relaciones entre los datos continúan siendo coherente y no representa redundancia o conflictos.

5.2. Capacitación y documentación

Vincular dentro de la capacitación del proceso estadístico y, así luego, especificar la capacitación de la fase.

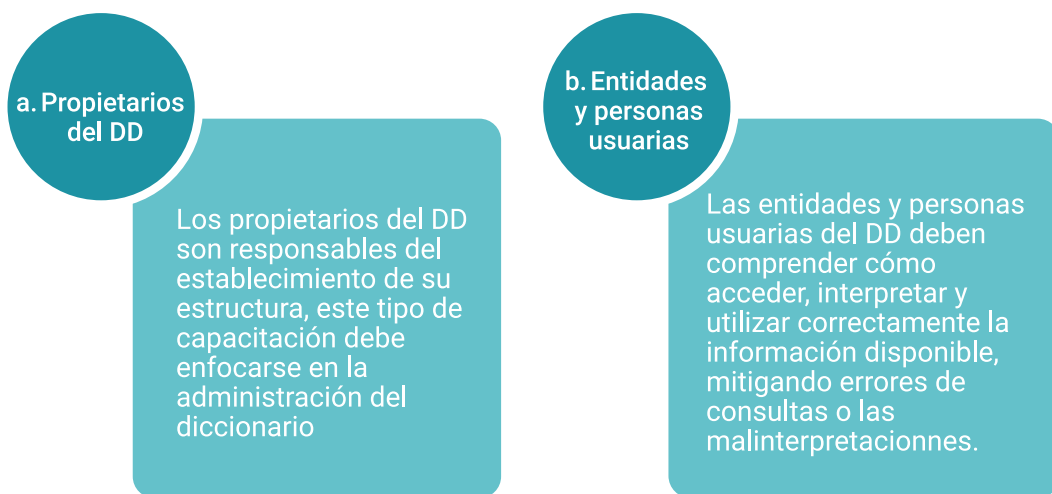
En este apartado, se busca resaltar la importancia de capacitar, específicamente sobre el Diccionario de Datos, su estructura, propósito y uso correcto. Aunque está relacionado con el proceso estadístico, se considera pertinente tratarlo como una actividad formativa diferenciada, ya que fortalece directamente la calidad de la documentación y su mantenimiento.

La etapa de capacitación y documentación busca garantizar que todas las personas involucradas en el proceso estadístico comprendan el propósito, la estructura y el funcionamiento del DD, asegurando su correcta aplicación dentro de las operaciones, y contribuyendo así a su mantenimiento y uso eficiente.

Elaborar planes de capacitación para los equipos involucrados, entidades y personas usuarias y documentar buenas prácticas no solo facilita el acceso a la información, sino que también mitiga errores y mantiene la coherencia en la gestión de los datos.

La siguiente ilustración presenta los tipos de capacitación que se han identificado y se recomiendan dentro de esta subetapa.

Ilustración 14. Tipos de capacitación.



Algunos temas que son claves al momento de diseñar y ejecutar las capacitaciones son:

Ilustración 15. Temas de capacitación relacionado a propietarios.



Ilustración 16. Temas de capacitación para entidades y personas usuarias del DD.



5.3. Sistematización

Mantener un DD actualizado y bien organizado puede no ser una tarea fácil si la información se gestiona de forma manual o dispersa. Esta subetapa busca optimizar este proceso mediante la implementación y uso de herramientas y metodologías que permitan automatizar, centralizar y estructurar la gestión del DD de forma eficiente y segura.

Cuando el DD se ha sistematizado, se mitiga el riesgo de errores, se facilita su mantenimiento y se mejora la accesibilidad para las entidades y personas usuarias. Esto permite que la información fluya de una manera ordenada, evitando duplicaciones, inconsistencias o datos desactualizados. La clave es transformar el DD en una herramienta dinámica, accesible y confiable, integrándolo con otros sistemas para que su consulta y actualización sean procesos rápidos, controlados y sin embotellamientos. Las instituciones productoras deben tomar en cuenta:

- Implementar una plataforma para la gestión centralizada del DD y sus bases de datos donde se utilice una herramienta tecnológica para su almacenamiento y administración (ej. Bases de datos estructurados, software de gestión para la documentación o plataformas especializadas).
- Eficientizar los procesos de actualización y validación automatizando los mismos desarrollando mecanismos de sincronización automática con bases de datos operativas para evitar duplicidades y errores.
- Implementar y alertar con notificaciones, para avisar cuando un dato se encuentre desactualizado o, que el mismo, se encuentre pendiente de revisión.
- Estandarizar la estructura y los Metadatos del DD al definir un formato único para la documentación de variables, asegurando que cada elemento cuente con información completa y clara.

- e. Garantizar la interoperabilidad con otros sistemas de gestión de datos.
- f. Integrar el DD con otras plataformas que permitan la exportación en formatos abiertos y estructurados (JSON, XML, CSV) para facilitar su uso en diferentes entornos.
- g. Implementar un sistema de registros que documente quién hizo cambios, para qué se realizaron los cambios, qué y cuándo se modificó.

5.4. Versionamiento y control de cambios

El DD no es ni debería ser un documento estático, es decir, este evoluciona constantemente a medida que la producción de las bases de datos de la institución crece, se ajustan definiciones o se incorporan nuevas variables. Es por lo que, sin un adecuado control de cambios, se compromete la pérdida de información valiosa, generan inconsistencias o dificultan la trazabilidad de las modificaciones y el uso de las bases de datos.

La subetapa de Versionamiento y control de cambios asegura que cada ajuste realizado en el DD quede registrado, documentado y pueda ser consultado en cualquier momento. Un buen sistema de Versionamiento facilita la recuperación de versiones anteriores y previene errores que puedan comprometer la calidad del DD, es por esto por lo que, las instituciones productoras deben tomar en cuenta:

- a. Definir una política de control de versiones y establecer protocolos para la realización de cambios.
- b. Implementar un registro estructurado de cambios, donde se registre cada ajuste con detalles (ej. Fecha del cambio, descripción de la modificación, persona responsable, motivo de cambio, versión).
- c. Establecer proceso de revisión y aprobación de cambios que describa el perfil de usuario y el flujo de validación.
- d. Comunicar y documentar los cambios realizados a las entidades y personas usuarias de manera transparente.
- e. Garantizar la recuperación y seguridad de la información realizando copias de seguridad periódicamente.

Etapas 6. Evaluación y validación

La última etapa definida dentro del proceso de diseño, producción y difusión del DD es la etapa de Evaluación y validación. La etapa de Evaluación y validación tiene como objetivo principal, verificar la calidad, consistencia y aplicabilidad del DD, asegurando que este cumpla con los estándares definidos y que refleje con fidelidad la estructura de los datos. En esta etapa se debe revisar si el DD está bien documentado, si las definiciones son claras y si la información se mantiene actualizada.

Con la estructuración de procesos de validación, se logra que el DD sea una herramienta efectiva, accesible y que la misma se encuentre alineada con los procesos institucionales establecidos en la producción de las estadísticas.

La siguiente ilustración presenta la etapa y subetapas que se han establecido para la Evaluación y validación del DD.

Ilustración 17. Etapa y subetapas establecidas para la Evaluación y Validación del DD.



6.1. Monitoreo y control de calidad

Un DD robusto (bien estructurado y documentado) pierde su valor si no se asegura su calidad a lo largo del tiempo. Es por esto por lo que, el monitoreo y control de calidad tiene como el objetivo principal supervisar constantemente la precisión, integridad y consistencia de la estructura e información contenida en el DD, garantizando que siga siendo una herramienta confiable y útil para las instituciones productoras estadísticas.

La subetapa 6.1. Monitoreo y control de calidad se integra de forma transversal en el proceso estadístico, con una conexión directa con la fase de Evaluación y los mecanismos de Gestión de la Calidad. Su propósito es asegurar que el Diccionario de Datos (DD) mantenga, a lo largo del tiempo, su precisión, integridad y consistencia estructural y documental. Un DD bien construido puede deteriorarse si no se supervisa continuamente su actualización, alineación con los estándares y coherencia con las bases de datos y operaciones estadísticas activas. Mediante este monitoreo, se establecen controles periódicos que permiten detectar errores, omisiones, duplicidades o desactualizaciones, reforzando la confiabilidad del DD como instrumento clave del proceso estadístico y garantizando su utilidad para las instituciones productoras en la toma de decisiones y la trazabilidad de los datos.

Llevar a cabo este proceso permite la detección temprana de errores, identificar inconsistencias y aplicar cualquier tipo de mejora que requiera el DD antes de que afecten la toma de decisiones o la interoperabilidad de los datos. Con la finalidad de que las instituciones productoras monitoreen y mantengan el control de la calidad del DD, estas deben tomar en cuenta:

- a. Establecer indicadores de calidad para el DD.
- b. Implementar un plan de monitoreo periódico conforme a la OE que permita detectar errores, duplicidades, información obsoleta, otros...
- c. Aplicar controles de calidad en la gestión de los datos validando su formato, definiciones, otros...
- d. Garantizar que los cambios realizados al DD sean revisados y validados antes de su implementación.
- e. Asignar un equipo responsable del monitoreo y mantenimiento del DD que asegure el control de la calidad y documente de forma transparente y trazable cada modificación realizada al mismo.
- f. Documentar internamente las actualizaciones y mejoras implementadas en el DD como parte del monitoreo continuo, asegurando su trazabilidad y disponibilidad para los procesos de verificación y control de calidad.

- g. Elaborar plan de auditoría interna que permita detectar posibles inconsistencias o datos desactualizados.
- h. Establecer un ciclo de revisiones periódica que permita el aseguramiento de la información y la mejora continua del DD.

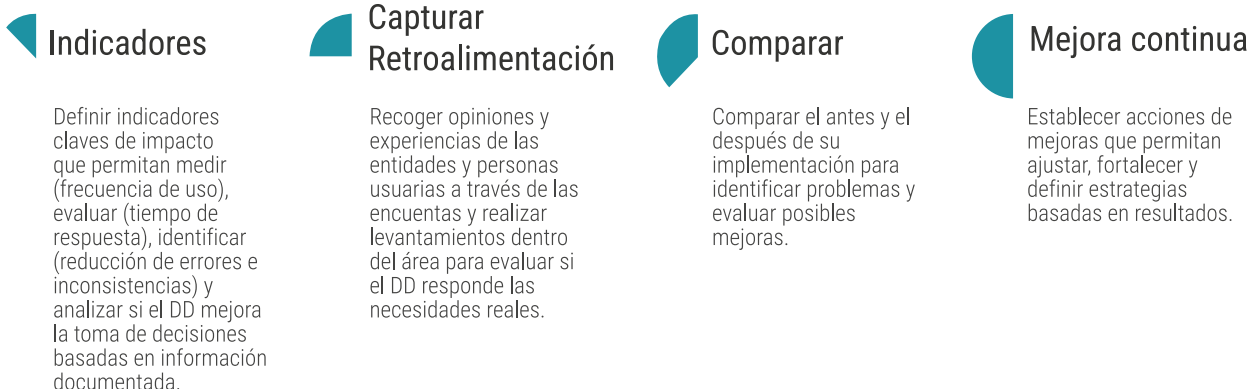
6.2. Evaluación de impacto

La subetapa sobre la Evaluación de impacto tiene como objetivo medir cómo el DD ha mejorado la gestión de datos, optimizado el trabajo de las entidades y personas usuarias y fortalecido la calidad de la información. Esta subetapa permite entender si el DD realmente está en cumplimiento con su propósito o, si es necesario hacer ajustes, para mejorar su utilidad. Analizar su impacto ayuda a validar su relevancia, identificar oportunidades de optimización y justificar la inversión en su mantenimiento y actualización.

La subetapa 6.2. Evaluación de impacto se articula directamente con la fase de Evaluación del Proceso Estadístico, ya que permite analizar en qué medida el Diccionario de Datos (DD) ha contribuido a mejorar la gestión de datos, facilitar el trabajo técnico y fortalecer la calidad de la información producida. Esta evaluación va más allá del monitoreo operativo, enfocándose en los efectos concretos que ha tenido el DD en la eficiencia, transparencia y trazabilidad del proceso estadístico. Medir su impacto permite validar su utilidad institucional, identificar áreas de mejora, y fundamentar la toma de decisiones respecto a su mantenimiento, actualización o rediseño. Así, el DD deja de ser solo un instrumento documental, y se posiciona como un activo estratégico dentro de la infraestructura estadística nacional.

La siguiente ilustración muestra algunas directrices y pautas que se han establecido al estructurar la subetapa para la Evaluación de impacto. Las instituciones productoras deben tomar en cuenta implementar:

Ilustración 18. Directrices y pautas para la Evaluación de impacto del DD.

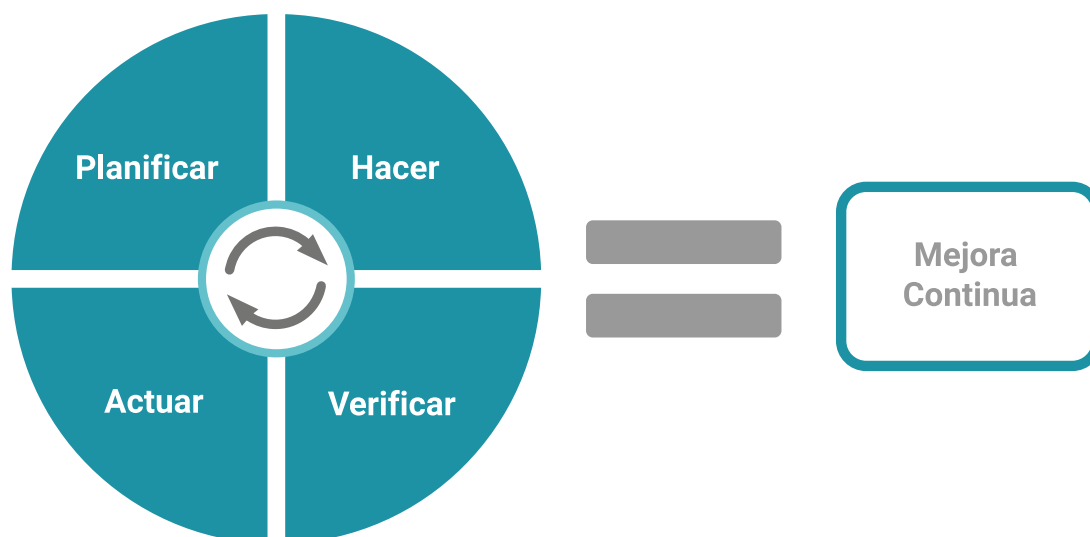


6.3. Mejora continua

El DD no puede ser un documento que se cree una sola vez y se deje en el olvido. Para que este sea útil y siga respondiendo a las necesidades de la institución productora, debe evolucionar constantemente con el tiempo. La mejora continua se encarga de que se revise constantemente el DD, detectando así, cualquier oportunidad de mejora que se pueda optimizar y hacer los ajustes necesarios que permitan que su contenido sea claro, actualizado y relevante.

Esta subetapa permite que el DD se adapte a cambios en los sistemas, nuevas normativas y a las nuevas necesidades existentes por parte de entidades y personas usuarias. Existen unos cuatro puntos clave que las instituciones productoras deben tomar en cuenta para mantener un DD útil, oportuno y listo para la toma de decisiones informadas, la siguiente ilustración hace referencia a estos puntos clave.

Ilustración 19. Aplicación de ciclo Deming¹⁷ para la mejora continua de un DD.



Aplicar el ciclo Deming o PHVA al DD asegura que este sea actualizado de manera constante, que esté alineado a los procesos institucionales más actualizados, optimiza la calidad y confiabilidad de la información documentada, facilita la identificación y corrección de errores o inconsistencias y mejora la experiencia y accesibilidad de entidades y personas usuarias a los datos.

Recomendaciones generales

Para garantizar que el Diccionario de Datos (DD) sea una herramienta eficiente y de gran valor para las instituciones productoras del Sistema Estadístico Nacional, es fundamental, seguir al pie de la letra los lineamientos que se han establecido en este documento, las directrices que se han venido desarrollando están relacionadas a los principios de calidad y buenas prácticas establecidas en el Código Nacional de Buenas Prácticas para las Estadísticas Oficiales y permiten robustecer el proceso estadístico llevado a cabo.

Mantener una documentación clara y accesible del DD para las entidades y personas usuarias es de vital importancia, evitar tecnicismos innecesarios y asegurar que la información esté organizada de manera intuitiva, facilitar la consulta y su uso.

Asegurar su actualización periódica es fundamental para su eficiencia al igual que promover su integración con distintos sistemas. Además, cumplir con los aspectos relacionados a seguridad y control de calidad permitirá que el DD sea una herramienta esencial para resguardar la confidencialidad e integridad de los datos y el mantener la precisión y coherencia.

Fomentar el uso adecuado del DD y su cumplimiento en los procesos internos de la institución, promoviendo su consulta activa entre los diferentes equipos de trabajo, para esto, se recomienda socializar su importancia como herramienta clave en la toma de decisiones y la gestión efectiva de datos.

Desarrollar y promover un plan de capacitaciones para las entidades y personas usuarias involucradas en la gestión y consulta del DD es una estrategia esencial que garantiza su correcta aplicación.

Por último, definir indicadores que permitan medir el desempeño de su uso, impacto, y nivel de cumplimiento, asegurará que el DD evolucione a las nuevas necesidades institucionales.

¹⁷ El Ciclo de Deming, es la metodología más usada para solucionar problemas y ejecutar sistemas de mejora continua. Su aplicación ayuda a que las organizaciones mejoren su rendimiento y aumenten su productividad.

Bibliografía

Código Nacional de Buenas Prácticas para las Estadísticas Oficiales

Lineamientos para el Proceso Estadístico en el Sistema Estadístico Nacional – ONE

Instructivo para la creación de Diccionario de Datos - Ideca

ISO/IEC 11179-1-2023 (4ta edición) - Information technology – Metadata

registries (MDR)

Diccionario de Datos Abiertos – Agencia Nacional de Contrataciones Públicas Colombia Compra Eficiente, 2022

Guía práctica para la elaboración de Diccionarios de Datos, versión 1.0

Ciclo de Deming: Metodología de mejora continua | PDCA – PHVA. <https://www.ingenieriadecalidad.com/2020/02/ciclo-de-deming.html>

Norma Técnica de la Calidad del Proceso Estadístico 1000 – 2020 – DANE

FORMATO EJEMPLO DE DICCIONARIO DE DATOS – DANE

Lineamiento para la Difusión de Resultados Estadísticos en el Sistema Estadístico Nacional – ONE

Lineamientos para la Realización de Prueba Experimental en el Proceso de Diseño de una Operación Estadística – ONE

Exploring Data-Driven Innovation as a New Source of Growth – OECD

Resolución núm. 073-2025, que establece los lineamientos para el desarrollo e

implementación del Marco Nacional de Interoperabilidad y Gobernanza de

Datos de la Administración Pública.

Anexos

Anexo 1. Definición de estándares de datos

1. Nomenclatura y formato de campo

- Definir una convención clara de nombres (ejemplo: camelCase, snake_case, uso de prefijos o sufijos).
- Especificar longitud máxima y mínima de campos.
- Tipos de datos permitidos (texto, numérico, booleano, fecha).

2. Reglas de documentación

- Metadatos obligatorios por campo: nombre técnico, descripción, unidad de medida (si aplica), tipo de dato, dominio permitido, periodicidad, confidencialidad, fuente.

3. Estandarización de valores

- Identificación de catálogos oficiales (ejemplo: clasificaciones CIIU, COICOP, códigos de país).
- Inclusión de dominios válidos y validaciones asociadas.

4. Requisitos de calidad y validación

- Definir reglas para asegurar:
- Consistencia: validación cruzada entre campos.
- Precisión: exactitud del valor documentado.
- Integridad: obligatoriedad o completitud del campo.
- Actualización: fecha de última modificación y responsable.

5. Uso de formatos abiertos

- JSON, XML, CSV: justificar su uso para facilitar interoperabilidad, reutilización y lectura automatizada.

6. Cumplimiento con normativas

- Normas para intercambio: SDMX, DCAT, API RESTful, etc.
- Reglas de compatibilidad tecnológica (versión de sistema, codificación UTF-8, estructura jerárquica, etc.)

Anexo 2. Ejemplo de cómo estructurar la documentación del DD para una Operación Estadística específica.

Metadatos Técnicos	
Elemento	Descripción
1. ID Tabla	Es un identificador único que permite hacer referencia a cada tabla en la base de datos.
2. Nombre de la Tabla	El nombre de la tabla debe ser descriptivo y claro para indicar qué tipo de información contiene.
3. ID Variable	Un identificador único para cada campo dentro de la tabla
4. Nombre de la Variable	El nombre del campo debe ser claro y reflejar el tipo de información que almacena.
5. Tipo de Llave	Indica si el campo es una clave primaria (PK) o clave foránea (FK).
6. Condición de la Variable	Establece si un campo es obligatorio (NOT NULL) o si puede ser nulo (NULL).
7. Longitud de la Variable	Indica el número máximo de caracteres o la longitud de los datos que puede almacenar el campo.
8. ID Base de Datos	Es un identificador único que permite hacer referencia a la Base de Datos.
9. Tipo de Dato	Define el tipo de dato del campo, por ejemplo, texto, numérico, fecha, etc.
9. Regla de Validación	Define las reglas que el dato debe cumplir, como ser único, estar dentro de un rango, etc.
Metadatos Semánticos	
Elemento	Descripción
1. Descripción de la Tabla	Breve resumen que explica el propósito de la tabla dentro del sistema.
2. Etiqueta de la Variable	Una versión amigable o legible del nombre técnico de la variable.
3. Categoría	Especifica el grupo conceptual al que pertenece la variable (por ejemplo, "Demográfico").
4. Código de la Categoría	Identificador único de la categoría (por ejemplo, DEMO01, FIN02).
5. Descripción de la Variable	Una descripción que explica qué representa el campo y cómo se usa.
5. Unidad de Medida	Especifica unidad de medida asociada al campo, si aplica (por ejemplo, kg, m, días, etc.).
5. Dominios/Valores Permitidos	Especifica valores o rangos de valores que el campo puede aceptar.
5. Etiqueta de Valor	Describe, de manera legible cada valor permitido en la variable, facilitando su interpretación.
5. Comentarios Adicionales	Información adicional del campo, como las dependencias con otros campos o restricciones adicionales.

Fuente: elaborado a partir de la captura de información de las áreas productoras, ONE 2025.

Nota 1. La estructura antes dispuesta se debe dar para cada una de las variables establecidas en el DD.

Nota 2. Metadatos Técnicos son aquellos campos que definen la estructura técnica de la base de datos, sus restricciones y formatos.

Nota 3. Metadatos Semánticos son aquellos campos que describen el significado, uso e interpretación del dato, estos son útiles para las entidades, personas usuarias y analistas.

Anexo 3. Ejemplo de un DD para una Base de Datos sobre Comercio Exterior.

Base de Datos: "Comercio_Exterior_EXP" Metadatos técnicos de la tabla exportaciones	
Elemento	Descripción
ID Tabla	1
Nombre de la Tabla	exportaciones
ID Variable	1
Nombre de la Variable	id_exportacion
Tipo de Llave	PK
Condición de la Variable	NOT NULL
Longitud de la Variable	10
ID Base de Datos	1
Tipo de Dato	Númérico (Entero)
Regla de Validación	Único, No repetido
Descripción de la Tabla	Información sobre las exportaciones realizadas, incluyendo su ID, fecha, valor, etc.

Elemento	Descripción
ID Tabla	1
Nombre de la Tabla	exportaciones
ID Variable	2
Nombre de la Variable	fecha_exportacion
Tipo de Llave	NULL
Condición de la Variable	NOT NULL
Longitud de la Variable	10
ID Base de Datos	1
Tipo de Dato	Fecha
Regla de Validación	Formato válido de fecha (YYYY-MM-DD)
Descripción de la Tabla	Fecha en la que se realizó la exportación.

Elemento	Descripción
ID Tabla	1
Nombre de la Tabla	exportaciones
ID Variable	3
Nombre de la Variable	valor_exportacion
Tipo de Llave	NULL
Condición de la Variable	NOT NULL
Longitud de la Variable	15
ID Base de Datos	1
Tipo de Dato	Decimal
Regla de Validación	Valores numéricos positivos, con dos decimales
Descripción de la Tabla	Valor total de la exportación en la moneda de referencia.

Anexo 3. Ejemplo de un DD para una Base de Datos sobre Comercio Exterior, continuación...

Elemento	Descripción
ID Tabla	1
Nombre de la Tabla	exportaciones
ID Variable	4
Nombre de la Variable	id_pais_destino
Tipo de Llave	FK
Condición de la Variable	NOT NULL
Longitud de la Variable	10
ID Base de Datos	1
Tipo de Dato	Numérico (Entero)
Regla de Validación	Relacionado con la tabla mercado_destino (FK)
Descripción de la Tabla	País al que se realiza la exportación.

Elemento	Descripción
ID Tabla	1
Nombre de la Tabla	exportaciones
ID Variable	5
Nombre de la Variable	producto
Tipo de Llave	NULL
Condición de la Variable	NOT NULL
Longitud de la Variable	100
ID Base de Datos	1
Tipo de Dato	Texto
Regla de Validación	No aplica
Descripción de la Tabla	Descripción del producto exportado.

Metadatos semánticos de la tabla exportaciones

Elemento	Descripción
Etiqueta de la Variable	ID Exportación
Categoría	Comercio Exterior
Código de la Categoría	COMEX01
Descripción de la Variable	Identificador único de una exportación específica.
Unidad de Medida	N/A
Dominios/Valores Permitidos	Números enteros positivos de 1 a 9999999999
Etiqueta de Valor	Número único que identifica la exportación.
Comentarios Adicionales	Este campo es clave primaria para las tablas relacionadas con exportaciones.

Anexo 3. Ejemplo de un DD para una Base de Datos sobre Comercio Exterior, continuación...

Elemento	Descripción
Etiqueta de la Variable	Fecha Exportación
Categoría	Comercio Exterior
Código de la Categoría	COMEX02
Descripción de la Variable	Fecha en la que se realiza la exportación.
Unidad de Medida	N/A
Dominios/Valores Permitidos	Fecha en formato "YYYY-MM-DD"
Etiqueta de Valor	Fecha en la que se completa la exportación.
Comentarios Adicionales	Es importante para la temporalidad de los informes.

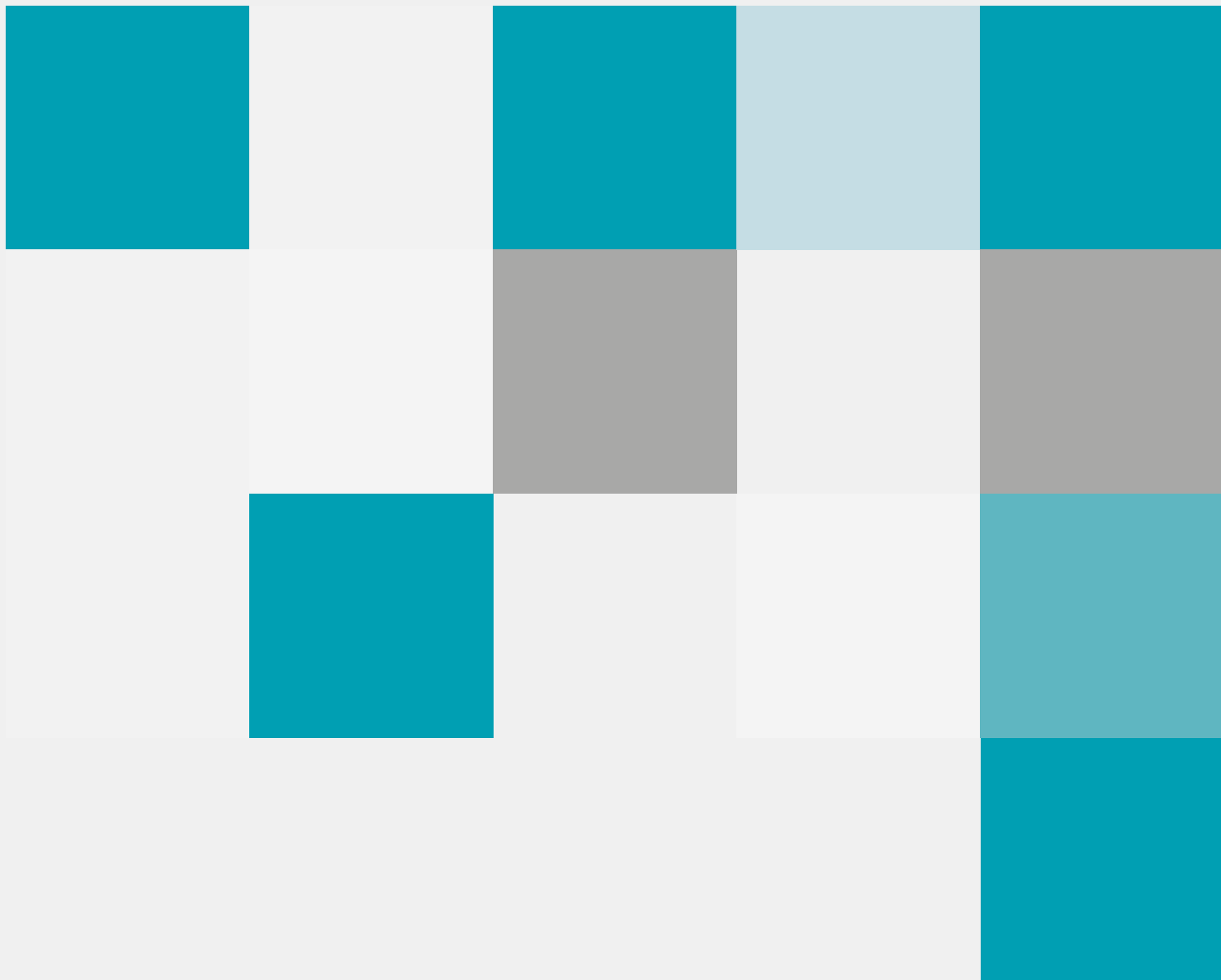
Elemento	Descripción
Etiqueta de la Variable	Valor Exportación
Categoría	Comercio Exterior
Código de la Categoría	COMEX03
Descripción de la Variable	Valor total de la exportación en la moneda de referencia.
Unidad de Medida	USD o la moneda de referencia
Dominios/Valores Permitidos	Números decimales con dos dígitos, valores positivos.
Etiqueta de Valor	Monto total de la exportación
Comentarios Adicionales	Puede estar relacionado con la tasa de cambio para conversiones.

Elemento	Descripción
Etiqueta de la Variable	ID País Destino
Categoría	Comercio Exterior
Código de la Categoría	COMEX04
Descripción de la Variable	País al cual se realiza la exportación.
Unidad de Medida	N/A
Dominios/Valores Permitidos	Lista de países válidos para el comercio internacional.
Etiqueta de Valor	Países o regiones a los que se destina el producto exportado.
Comentarios Adicionales	Este campo es clave foránea hacia la tabla mercado_destino.

Elemento	Descripción
Etiqueta de la Variable	Producto
Categoría	Comercio Exterior
Código de la Categoría	COMEX05
Descripción de la Variable	Tipo de producto exportado.
Unidad de Medida	N/A
Dominios/Valores Permitidos	Texto libre que describe el producto exportado.
Etiqueta de Valor	Nombre del producto exportado.
Comentarios Adicionales	Puede ser una lista extensa dependiendo de la diversidad de productos exportados.

Ejemplo de Datos en la Tabla exportaciones

ID Exportación	Fecha Exportación	Valor Exportación	ID País Destino	Producto
1	01-04-2025	50000.00	1	Automóviles
2	02-04-2025	120000.00	2	Café



Oficina Nacional de Estadística
Av. México esq. Leopoldo Navarro
Edificio de Oficinas Gubernamentales
Juan Pablo Duarte, pisos 8 y 9
Tel.: 809-682-7777 • Correo: info@one.gob.do