



DANE

SEN^{2.0}
Sistema Estadístico
Nacional - Colombia

DIRECCIÓN DE REGULACIÓN, PLANEACIÓN, ESTANDARIZACIÓN Y NORMALIZACIÓN - DIRPEN

RECOMENDACIONES DE MECANISMOS, ESTÁNDARES Y PROTOCOLOS PARA LA INTEROPERABILIDAD EN EL INTERCAMBIO, USO Y REÚSO DE DATOS EN EL SISTEMA ESTADÍSTICO NACIONAL (SEN)

GRUPO INTERNO DE TRABAJO DE REGULACIÓN ESTADÍSTICA

MAYO de 2025

DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA (DANE)

B. Piedad Urdinola Contreras
Directora

Andrea Ramírez Pisco
Subdirectora

Álvaro Fernando Guzmán Lucero
Secretario general

DIRECCIONES TÉCNICAS

Julieth Alejandra Solano Villa
Dirección de Regulación, Planeación,
Estandarización y Normalización

Elkin Ernesto Ramírez Niño
Dirección de Geoestadística

César Mauricio López Alfonso
Dirección de Metodología y
Producción Estadística

Javier Sebastián Ruiz Santacruz
Dirección de Censos y Demografía

Juan Pablo Cardoso Torres
Dirección de Síntesis y Cuentas Nacionales

Diana María Bohórquez Losada
Dirección de Difusión y Cultura Estadística

Diana Rojas Orduz
Dirección de Recolección y Acopio

© DANE, 2025

Prohibida la reproducción total o parcial sin permiso o autorización del Departamento Administrativo Nacional de Estadística, Colombia.

**DIRECCIÓN DE REGULACIÓN, PLANEACIÓN,
ESTANDARIZACIÓN Y NORMALIZACIÓN (DIRPEN)**
Grupo Interno de Trabajo de Regulación Estadística.

Iván Rolando Castillo Prieto
Coordinador GIT Regulación Estadística

Equipo de trabajo:

Juliana Mosquera Dueñas
Profesional especializada GIT Regulación Estadística

Corrección de estilo
Sonia Marcela Naranjo Morales
Profesional universitaria GIT Regulación Estadística

CONTENIDO

Introducción	5
1. Antecedentes	6
2. Objetivo y alcance	10
2.1. Objetivo.....	10
2.2. Alcance	10
3. Marco legal y de interoperabilidad	10
3.1. Marco legal	10
3.2. Marco de interoperabilidad	12
4. Conceptos clave	13
5. Referentes internacionales y nacionales	17
5.1. Referentes internacionales.....	17
5.1.1. <i>Global Partnership for Sustainable Development Data (GPSDD) (Asociación Mundial de Datos para el Desarrollo Sostenible).....</i>	<i>17</i>
5.1.2. <i>Data to Policy Navigator</i>	<i>18</i>
5.1.3. <i>Organización para la Cooperación y el Desarrollo Económico (OCDE).....</i>	<i>18</i>
5.1.4. <i>Comisión Económica de las Naciones Unidas para Europa (UNECE)</i>	<i>19</i>
5.1.5. <i>Portal de Datos Europeo</i>	<i>21</i>
5.1.6. <i>Conclusiones y recomendaciones a partir de la revisión de referentes internacionales.....</i>	<i>22</i>
5.2. Referentes nacionales.....	23
5.2.1. <i>Plan Nacional de Infraestructura de Datos (PNID).....</i>	<i>23</i>
6. Recomendaciones de interoperabilidad en el intercambio, el uso y el reúso de datos en el SEN	24
6.1. Pasos recomendados para el acceso, el uso y el reúso de datos.....	25
6.2. Sugerencias de mecanismos, estándares y protocolos para el intercambio, el uso y el reúso de datos organizados por dimensión de interoperabilidad	28
7. Bibliografía	42
Anexo A. Vocabularios, métodos, formatos, marcos, lenguajes, flujos de trabajo y modelos de datos estandarizados	43

LISTA DE CUADROS

Cuadro 1. Recomendaciones para la interoperabilidad del Portal de datos europeo.....	21
Cuadro 2. Retos en la gestión de datos, gobernanza e interoperabilidad	35
Cuadro 3. Retos en modelos de datos y de metadatos	36
Cuadro 4. Retos en clasificaciones y vocabularios	38
Cuadro 5. Retos en formatos de datos abiertos e interfaces estandarizadas	40
Cuadro 6. Retos en datos enlazados	41

LISTA DE TABLAS

Tabla 1. Resumen de instrumentos normativos asociados a la interoperabilidad, la reserva estadística y al intercambio de información estadística en el SEN	10
Tabla 2. Resumen de artículos Ley 2335 de 2023 asociados a la interoperabilidad, la reserva estadística y al intercambio de información estadística en el SEN	11
Tabla 3. Dominios del Marco de Interoperabilidad de MinTIC	13
Tabla 4. Conceptos clave asociados al intercambio de información estadística.....	13
Tabla 5. Dimensiones clave para fortalecer la interoperabilidad según GPSDD	17
Tabla 6. Pasos del Data to Policy Navigator para formulación de políticas públicas.....	18
Tabla 7. Desafíos en la interoperabilidad en el proceso de producción estadística según el GSBPM	20
Tabla 8. Algunas directrices de MinTIC y de la SIC para la implementación del PNID asociadas a la interoperabilidad, el aprovechamiento y la seguridad de los datos	23
Tabla 9. Ejemplos de métricas para monitorear la interoperabilidad estadística en organizaciones	28

Introducción

El Departamento Administrativo Nacional de Estadística (DANE), como autoridad estadística en Colombia y ente rector del Sistema Estadístico Nacional (SEN), coordina y regula la producción y la difusión de estadísticas oficiales con plena independencia técnica y en articulación con las entidades productoras. Orienta a los miembros del sistema en la aplicación de metodologías y estándares en materia estadística (Ley 2335 de 2023), con el fin de mejorar la transparencia, la pertinencia, la interoperabilidad, el acceso, la oportunidad y la coherencia de las estadísticas producidas para la toma de decisiones en el ámbito nacional y territorial.

Por ello, el DANE elabora este documento en cumplimiento del artículo 17 de la Ley 2335, que le asigna definir mecanismos, estándares y protocolos para el intercambio de información estadística respetando la confidencialidad y la reserva estadística. El objetivo es brindar al SEN recomendaciones que fortalezcan la interoperabilidad para facilitar el intercambio eficiente de información estadística, así como el uso y reúso de datos estadísticos bajo condiciones de privacidad, seguridad y calidad.

La interoperabilidad optimiza el intercambio y uso de información estadística al estandarizar conceptos, metadatos y procesos, lo que mejora la integración, el análisis y el acceso. Para los usuarios, reduce errores interpretativos y facilita la automatización de sistemas; para los productores, fomenta un intercambio eficiente de información, impulsa la reutilización de herramientas y metodologías, automatiza procesos, incrementa la calidad estadística y reduce costos operativos.

Este documento¹ se basa en los lineamientos de los principales referentes internacionales en la materia como el *Data Governance Framework for Statistical Interoperability* de la Comisión Económica de las Naciones Unidas para Europa (UNECE), la guía *Data Interoperability: A Practitioner's Guide to Joining Up Data in the Development Sector* de la Global Partnership for Sustainable Development Data (GPSDD) y el *Data to Policy Navigator* de Data4Policy de la Cooperación Técnica Alemana (GIZ) y el Programa de las Naciones Unidas para el Desarrollo (PNUD).

También, considera referentes nacionales como el Plan Nacional de Infraestructura de Datos (PNID), el Marco de interoperabilidad del Ministerio de Tecnologías de la Información y las Comunicaciones (MinTIC) y los lineamientos, normas y estándares del DANE en la materia.

¹ Hace parte de la estrategia 6 del Plan Estadístico Nacional 2022-2027 que busca fortalecer la capacidad del SEN para aprovechar los registros administrativos. Disponible en: https://www.sen.gov.co/sites/default/files/noticias-files/2023-12/plan_estadistico_naciona_2023-2027_0.pdf

El documento presenta los antecedentes, el marco legal, los conceptos y los referentes clave. Aborda los pasos para el acceso, el uso y el reúso de información estadística y las recomendaciones de mecanismos, estándares y protocolos según las cinco dimensiones clave de la interoperabilidad de la GPSDD. Finalmente, incluye un anexo con ejemplos de vocabularios, métodos, formatos, marcos y lenguajes del *Data Governance Framework for Statistical Interoperability* (DAFI) de la UNECE.

1. Antecedentes

La interoperabilidad en los sistemas estadísticos es esencial para garantizar la integración, el intercambio y el aprovechamiento eficiente de la información en un entorno caracterizado por la rápida evolución tecnológica y la creciente interconexión de sistemas de información. La interoperabilidad es la capacidad de las organizaciones para intercambiar información y conocimiento en el marco de sus procesos de negocio², facilitando la interacción hacia objetivos mutuamente beneficiosos, como la entrega de servicios en línea a ciudadanos, empresas y otras entidades mediante el intercambio de datos entre sus sistemas (MinTIC, 2019).

En el ámbito de las organizaciones estadísticas la interoperabilidad es la capacidad de compartir y utilizar información con poca o ninguna comunicación previa (UNECE, 2023). Esta capacidad es esencial para garantizar que los datos puedan ser integrados, intercambiados, usados y reusados eficientemente, maximizando su valor para los productores y los usuarios de información.

La interoperabilidad facilita la localización e integración de datos, reduce errores interpretativos y mejora la comprensión de la información para los usuarios. Para los productores, permite la reutilización de herramientas, optimiza procesos de gestión de datos, eleva la calidad estadística y reduce costos operativos (UNECE, 2023; UNSD y GPSDD, 2018).

La comunidad estadística internacional ha reconocido la interoperabilidad como un pilar esencial que mejora la calidad, el uso y el intercambio de información estadística. En el Primer Foro Mundial de Datos de las Naciones Unidas (UNWDF) en 2017, se estableció la *Collaborative on SDG Data Interoperability*, liderada por UNSD y GPSDD con el fin de fortalecer la cooperación basada en datos y desarrollar guías sobre interoperabilidad, destacando su rol en la gestión de datos y la formulación de políticas.

² Un proceso de negocio es un conjunto de funciones, en una secuencia específica, que entrega valor a un cliente interno o externo de una organización. (Kirchmer, 2017).

En este contexto, la GPSDD elaboró la guía *Data Interoperability: A Practitioner's Guide to Joining Up Data in the Development Sector (2018)*³ encaminada a mejorar la integración y el reúso de datos. La guía aborda cinco áreas fundamentales para fortalecer la interoperabilidad: gestión de datos y gobernanza e interoperabilidad; modelos canónicos de datos y metadatos; clasificaciones y vocabularios; interfaces estandarizadas; y datos enlazados.

Por su parte, el informe de la OCDE *Enhancing Access to and Sharing of Data*⁴ de 2019 promueve el acceso y la compartición de datos para maximizar beneficios sociales y científicos. Propone el uso de API y *data sandboxes* y protocolos para la interoperabilidad sintáctica y semántica con estándares como OWL y Dublin Core y normas como ISO/IEC 19941:2017 y 20889:2018 para garantizar la portabilidad, la desidentificación y la protección de datos.

En 2021 Data.europa.eu creó el documento *Data Quality Guidelines*⁵ que promueve la calidad e interoperabilidad de los datos abiertos de la UE mediante principios FAIR (encontrabilidad, accesibilidad, interoperabilidad y reutilización). Destacan estándares como DCAT-AP, ISO 8601 y UTF-8, vocabularios controlados y esquemas JSON/XML para integración semántica y planes de gestión de datos.

También, en 2021, la UNECE creó el *Data Governance Framework for Statistical Interoperability (DAFI)*⁶ donde propone mecanismos para abordar los desafíos de interoperabilidad en el proceso estadístico resaltando la gobernanza de los datos para lograr interoperabilidad semántica, estructural y sintáctica mediante estándares abiertos y tecnologías para la integración y el intercambio.

Igualmente, el *Data to Policy Navigator*⁷ (2023) de Data4Policy de la GIZ y el PNUD, promueve el acceso, uso e intercambio de información para transformarla en políticas públicas. El proceso consta de cuatro pasos: definir el problema; acceder a los datos; procesarlos y analizarlos; y diseñar políticas, integrando aspectos como ética, gobernanza de datos, interoperabilidad, desarrollo de capacidades y mecanismos institucionales.

³ Disponible en: https://www.data4sdgs.org/sites/default/files/services_files/Interoperability%20-%20A%20practitioner%E2%80%99s%20guide%20to%20joining-up%20data%20in%20the%20development%20sector.pdf

⁴ disponible en: <https://www.oecd-ilibrary.org/docserver/276aaca8-en.pdf?expires=1727731397&id=id&accname=guest&checksum=9AA5276F9C5CF55C3BCABA7CC0E05CAF>

⁵ Disponible en: <https://op.europa.eu/en/publication-detail/-/publication/023ce8e4-50c8-11ec-91ac-01aa75ed71a1>

⁶ Disponible en: https://unece.org/sites/default/files/2024-03/HLG2023%20DAFI%20Final_0.pdf

⁷ Disponible en: <https://www.datatopolicy.org/navigator>

Colombia ha desarrollado e implementado diversos lineamientos⁸ que potencian el valor social y económico de los datos como el CONPES 3920 de 2018, el CONPES 3975 de 2019, el CONPES 4023 de 2021 y la Ley 2335 de 2023. Estos promueven la disponibilidad e interoperabilidad de datos, fomentan una cultura de datos, garantizan principios éticos en su uso y fortalecen la economía del intercambio de datos. (MinTIC, 2019).

Además, en el marco del CONPES 4023 de 2021 surgió el Plan Nacional de Infraestructura de Datos (PNID), liderado por el MinTIC, el Departamento Nacional de Planeación y la Presidencia de la República, que creó una hoja de ruta para gestionar la infraestructura de datos del Estado, fomentando la gobernanza, la interoperabilidad y la integración de datos abiertos y transaccionales.

Asimismo, la Ley 2335 de 2023 asignó al DANE definir mecanismos, estándares y protocolos para el intercambio eficiente de información estadística, protegiendo la privacidad y la confidencialidad. Este intercambio debe seguir el Marco de Interoperabilidad⁹ que brinda directrices para facilitar el intercambio entre entidades y promueve el uso de estándares abiertos garantizando la interoperabilidad semántica y técnica, así como datos consistentes y reutilizables.

Igualmente, la Ley establece las instancias de coordinación del SEN como el Comité de Administración de Datos (CAD) y los Comités Estadísticos Sectoriales. El CAD facilita la interacción entre entidades, promueve la interoperabilidad, el uso ético y el intercambio de información, asegurando la reserva estadística¹⁰. Los Comités, gestionan necesidades estadísticas, planes de acción y estándares para el intercambio de datos.

Con el fin de fortalecer el intercambio de información y promover la interoperabilidad en el SEN, el DANE ha desarrollado lineamientos y diversos instrumentos técnicos¹¹. Por ejemplo, el *Marco de Aseguramiento de la Calidad Estadística (MAC) (2022)*¹² que promueve el uso de estándares abiertos, la modernización tecnológica y la coordinación entre entidades, asegurando datos consistentes y reutilizables. El *Código Nacional de Buenas Prácticas del Sistema Estadístico Nacional (CNBP) (2022)*¹³

⁸ Estas políticas y la Ley 2335 de 2023 se describen en el marco legal del documento.

⁹ Disponible en: https://lenguaje.mintic.gov.co/sites/default/files/archivos/marco_de_interoperabilidad_para_gobierno_digital.pdf

¹⁰ La Política de gobernanza del CAD incluye los mecanismos de articulación, colaboración e interlocución con el SEN y otras iniciativas de política del Estado colombiano. Disponible en: https://www.sen.gov.co/sites/default/files/pagina-migraciones-files/2024-02/politica_gob_cad_051223.pdf

¹¹ Estos lineamientos, normas, estándares y recomendaciones pueden consultarse en: <https://www.sen.gov.co/normatividad/lineamientos-guias-manuales>

¹² Disponible en: <https://www.sen.gov.co/sites/default/files/migracion-files/normatividad/calidad/Marco-Aseguramiento-Calidad-SEN-Colombia-2022.pdf>

¹³ Disponible en: <https://www.sen.gov.co/sites/default/files/migracion-files/BuenasPracticas/Codigo-Nacional-de-Buenas-Practicas-2022.pdf>

fomenta la interoperabilidad mediante la adopción de estándares globales, la coordinación interinstitucional, el acceso y reúso de registros administrativos, la coherencia y comparabilidad de la información y el uso de tecnologías interoperables.

Los *Lineamientos para el Proceso Estadístico V2 (2020)*¹⁴ brindan una estructura estándar para la producción y difusión estadística, describiendo las fases, los subprocesos y las actividades fundamentales en la producción estadística mejorando la calidad y la consistencia de la información. La *Norma Técnica de Calidad del Proceso Estadístico (NTC PE 1000:2020)*¹⁵ establece los requisitos de calidad en los procesos estadísticos y brinda un marco de evaluación y mejora para que la información sea consistente, confiable y útil para su intercambio y análisis.

Además, las guías para el uso de estándares como SDMX¹⁶, DDI y Dublin Core¹⁷, la *Guía de anonimización de datos estructurados*¹⁸, la *Guía de metadatos de registros administrativos*¹⁹ y la *Guía para realizar intercambios de información*²⁰, entre otros lineamientos²¹, facilitan la documentación, catalogación e intercambio de información, asegurando su trazabilidad, reúso y privacidad. También la *Guía para la estandarización de conceptos*²² y las clasificaciones estadísticas adaptadas para Colombia²³ contribuyen a la coherencia semántica en la información estadística.

A pesar del avance en normatividad y lineamientos, el SEN enfrenta retos de interoperabilidad que limitan el uso, el intercambio y el reúso de información como la fragmentación de sistemas, la adopción inconsistente de estándares, las barreras organizacionales y la necesidad de proteger los datos personales garantizando el acceso. Por ello, es crucial definir mecanismos, estándares y protocolos que fortalezcan la interoperabilidad, optimicen recursos, garanticen la privacidad y mejoren la confianza en las estadísticas.

¹⁴ Disponible en: https://www.sen.gov.co/sites/default/files/migracion-files/sen/normatividad/Lineamientos_Proceso_Estad%C3%ADstico_v2.pdf

¹⁵ Disponible en: <https://www.dane.gov.co/files/sen/normatividad/NTC-Proceso-Estadistico-PE-1000-2020.pdf>

¹⁶ Disponible en: <https://www.sen.gov.co/normatividad/sdmx>

¹⁷ Disponible en: https://www.sen.gov.co/sites/default/files/migracion-files/guias/Guia_para_implementar_DDI_y_DC.pdf

¹⁸ Disponible en: <https://www.sen.gov.co/servicios/fortalecimiento-registros-administrativos>

¹⁹ Disponible en: https://www.sen.gov.co/sites/default/files/migracion-files/RegistrosAdministrativos/Gu%C3%ADa_Metadatos_RRAA_30072018.pdf

²⁰ Disponible en: https://www.sen.gov.co/sites/default/files/migracion-files/sen/lineamientos/Gu%C3%ADa_Protocolos_de_Intercambio.pdf

²¹ Consulte las guías, normas y estándares en: <https://www.sen.gov.co/normatividad/lineamientos-guias-manuales>

²² Consulte: <https://www.sen.gov.co/sites/default/files/pagina-migraciones-files/2024-02/guia-para-la-estandarizacion-de-conceptos-del-sistema-estadistico-nacional.pdf>

²³ Consulte las clasificaciones y tablas correlativas: <https://www.dane.gov.co/index.php/sistema-estadistico-nacional-sen/normas-y-estandares/nomenclaturas-y-clasificaciones>

2. Objetivo y alcance

2.1. Objetivo

Brindar a los miembros del SEN recomendaciones de mecanismos, estándares y protocolos que fortalezcan la interoperabilidad para promover el intercambio, el uso y el reúso eficiente y seguro de información estadística.

2.2. Alcance

Productores y usuarios de estadísticas del SEN. Se centra en la información estadística²⁴ sin considerar otros tipos de información (e.g. recursos humanos, aspectos financieros o normativos) gestionada por los productores y usuarios para apoyar la actividad estadística. No prioriza las recomendaciones de acuerdo con su impacto en la interoperabilidad ni analiza los costos y beneficios asociados a la implementación de estas recomendaciones por parte de las entidades del SEN.

3. Marco legal y de interoperabilidad

3.1. Marco legal

Este apartado presenta un resumen de los instrumentos normativos (Tabla 1) y de los artículos de la Ley 2335 de 2023 que están relacionados con la interoperabilidad, el intercambio y la reserva estadística (Tabla 2).

Tabla 1. Resumen de instrumentos normativos asociados a la interoperabilidad, la reserva estadística y al intercambio de información estadística en el SEN

Instrumento Legal	Nombre	Resumen
CONPES 3920 de 2018	"Política Nacional de Explotación de Datos y Big Data"	Establece lineamientos para la explotación estratégica de datos y Big Data, promoviendo su uso como motor de valor social y económico. Impulsa la interoperabilidad y la gestión eficiente de datos en sectores clave.
CONPES 3975 de 2019	"Política Nacional de Transformación Digital e Inteligencia Artificial"	Define estrategias para la adopción de tecnologías digitales y de inteligencia artificial, enfocándose en la transformación digital del sector público y privado, así como en la ética en el manejo de datos.

²⁴ Se define como los resultados obtenidos a partir de operaciones estadísticas, junto con la documentación que los respalda, los cuales describen o representan características de elementos, fenómenos u objetos de estudio (Ley 2335 de 2023).

RECOMENDACIONES DE MECANISMOS, ESTÁNDARES Y PROTOCOLOS PARA LA INTEROPERABILIDAD EN EL INTERCAMBIO, EL USO Y EL REÚSO DE DATOS EN EL SEN / REGULACIÓN ESTADÍSTICA

Instrumento Legal	Nombre	Resumen
CONPES 4023 de 2021	"Política para la reactivación y el crecimiento sostenible e incluyente"	Plantea acciones para la reactivación económica sostenible e inclusiva, integrando el uso de datos para fortalecer políticas públicas y promover el desarrollo con sostenibilidad ambiental y social. Da origen al <i>Plan Nacional de Infraestructura de Datos (PNID)</i> que establece hoja de ruta para gestionar la infraestructura de datos del Estado, fomentando la interoperabilidad, la integración de datos abiertos y transaccionales, y la gobernanza de datos.
Ley 2335 de 2023	"Por la cual se expiden disposiciones sobre las estadísticas oficiales en el país"	Regula la producción y difusión de estadísticas oficiales en Colombia, promoviendo estándares de calidad, confidencialidad, interoperabilidad y acceso público para garantizar datos confiables y oportunos.
Ley 1581 de 2012	"Por la cual se dictan disposiciones generales para la protección de datos personales"	Establece disposiciones para la protección de datos personales y busca desarrollar el derecho constitucional de todas las personas a conocer, actualizar y rectificar la información almacenada sobre ellas en bases de datos o archivos. Protege otros derechos, libertades y garantías establecidos en el artículo 15 de la Constitución Política, junto con el derecho a la información reconocido en el artículo 20.
Ley 1712 de 2014	"Por medio de la cual se crea la Ley de Transparencia y del Derecho de Acceso a la Información Pública Nacional y se dictan otras disposiciones"	Regula el derecho de los ciudadanos a acceder a la información pública. Define los procedimientos para ejercer y garantizar este derecho, estableciendo que toda la información en poder, control o custodia de los sujetos obligados es pública, salvo excepciones previstas por la Constitución o la ley.
Decreto 1389 de 2022	"Por el cual se adiciona el Título 24 a la Parte 2 del Libro 2 del Decreto Único 1078 de 2015, Reglamentario del Sector de Tecnologías de la Información y las Comunicaciones, con el fin de establecer los lineamientos generales para la gobernanza en la infraestructura de datos, y se crea el Modelo de gobernanza de la infraestructura de datos"	Establece el Modelo de Gobernanza de la Infraestructura de Datos en Colombia, con el fin de fortalecer la interoperabilidad entre entidades públicas, privadas y la sociedad civil, mediante estándares comunes que permitan el intercambio seguro y eficiente de datos. Promueve el uso y reúso de los datos como activos estratégicos para la toma de decisiones, la formulación de políticas públicas y la creación de valor social. El modelo se estructura en tres niveles de gobernanza (estratégico, táctico y operativo), articulándose con la Política de Gobierno Digital y fomentando una cultura organizacional basada en la gestión responsable, abierta y reutilizable de la información.

Fuente: elaboración propia

Tabla 2. Resumen de artículos Ley 2335 de 2023 asociados a la interoperabilidad, la reserva estadística y al intercambio de información estadística en el SEN

Artículo	Descripción
Artículo 4	Establece principios de coherencia, comparabilidad, transparencia y accesibilidad, esenciales para asegurar la interoperabilidad y la calidad de las estadísticas oficiales.
Artículo 7	Funciones del DANE como autoridad para orientar, coordinar, y regular la producción de estadísticas, garantizando la aplicación de metodologías y estándares.
Artículo 10	Define la finalidad del SEN, estableciendo un esquema de coordinación que promueva la interoperabilidad y la coherencia en la producción de estadísticas.
Artículo 11	Objetivos del SEN, destacando la importancia de lenguajes y procedimientos comunes, así como la interoperabilidad y la calidad de las estadísticas.
Artículo 13	Obligaciones de los integrantes del SEN para implementar estándares, normas técnicas, y garantizar la producción y difusión de estadísticas oficiales, incluidas las estrategias y acciones establecidas en el Plan Estadístico Nacional.

RECOMENDACIONES DE MECANISMOS, ESTÁNDARES Y PROTOCOLOS PARA LA INTEROPERABILIDAD EN EL INTERCAMBIO, EL USO Y EL REÚSO DE DATOS EN EL SEN / REGULACIÓN ESTADÍSTICA

Artículo	Descripción
Artículo 14	Establece instancias de coordinación del SEN: CASEN, Comité de Seguimiento a Estadísticas Estratégicas, CAD, Comités Estadísticos Sectoriales y Mesas Estadísticas Sectoriales. El CAD promueve el intercambio de bases de datos a nivel de microdato en un ambiente seguro, garantizando un uso ético de los datos. El DANE regula su funcionamiento, asegurando independencia técnica, calidad y cumplimiento de normas de protección de datos.
Artículo 17	Establece que los integrantes del SEN pueden intercambiar datos agregados y microdatos de forma gratuita y oportuna, siguiendo las Leyes Estatutarias 1266 de 2008 y 1581 de 2012. Se requiere que el intercambio cumpla con los estándares tecnológicos y el marco de interoperabilidad del Ministerio TIC, asegurando la protección de la identidad. El DANE define los mecanismos, estándares y protocolos, respetando la reserva estadística.
Artículo 18	Regula el aprovechamiento de registros administrativos para la producción de estadísticas oficiales, permitiendo el intercambio de microdatos respetando la reserva estadística. Se hace referencia a las Leyes Estatutarias 1266 de 2008 y 1581 de 2012, que regulan la protección y tratamiento de datos personales.
Artículo 23	Indica que las entidades públicas deben proporcionar datos al DANE sin oponer reservas legales, garantizando así el acceso a la información necesaria para la producción estadística. Se hace referencia a las Leyes Estatutarias 1266 de 2008 y 1581 de 2012, que regulan la protección y tratamiento de datos personales.
Artículo 36	Establece que los datos individuales que permitan la identificación directa o indirecta deben mantenerse bajo reserva estadística, protegiendo la privacidad de personas y entidades. Se hace referencia a las Leyes Estatutarias 1266 de 2008 y 1581 de 2012, que regulan la protección y tratamiento de datos personales.
Artículo 37	Los datos recogidos para fines estadísticos no podrán ser utilizados para otros propósitos, garantizando la confidencialidad y evitando usos indebidos. Se hace referencia a las Leyes Estatutarias 1266 de 2008 y 1581 de 2012, que regulan la protección y tratamiento de datos personales.
Artículo 38	Proporciona lineamientos para la seguridad en el procesamiento y almacenamiento de datos individuales, asegurando que el acceso esté restringido a personal autorizado. Se hace referencia a las Leyes Estatutarias 1266 de 2008 y 1581 de 2012, que regulan la protección y tratamiento de datos personales.
Artículo 39	Restringe la divulgación de datos individuales por los miembros del SEN, excepto cuando los datos hayan sido anonimizados para uso público.
Artículo 40	Obliga a que los datos sean procesados de manera que se eliminen los identificadores para evitar la identificación de personas o entidades, asegurando la privacidad en la publicación de datos.
Artículo 41	Autoriza al DANE a recibir datos individuales bajo reserva, asegurando la protección de datos durante su custodia y procesamiento.
Artículo 42	Establece el compromiso de confidencialidad para todas las personas que participen en la producción de estadísticas oficiales, asegurando que la reserva estadística se mantenga incluso después de concluir sus funciones. Se hace referencia a las Leyes Estatutarias 1266 de 2008 y 1581 de 2012, que regulan la protección y tratamiento de datos personales.

Fuente: elaboración propia con base en Ley 2335 de 2023.

3.2. Marco de interoperabilidad

El *Marco de Interoperabilidad*²⁵ del MinTIC es una guía técnica para facilitar el intercambio seguro y eficiente de información entre entidades públicas. Promueve servicios digitales ágiles, efectivos y centrados en el ciudadano, así como la importancia de la interoperabilidad para mejorar la calidad de los servicios públicos. Se organiza en cuatro dominios (Tabla 3).

²⁵ Disponible en: https://lenguaje.mintic.gov.co/sites/default/files/archivos/marco_de_interoperabilidad_para_gobierno_digital.pdf

Tabla 3. Dominios del Marco de Interoperabilidad de MinTIC

Dominio	Descripción
Político-legal	Asegura que el intercambio de información se realice dentro de un marco normativo adecuado, ajustado a las políticas y leyes vigentes. Acciones y recomendaciones: 1) Definir instrumentos legales claros para formalizar acuerdos de intercambio, asegurando cumplimiento de obligaciones legales y de privacidad. 2) Garantizar el cumplimiento normativo de las leyes de protección de datos, priorizando la seguridad y confidencialidad de la información.
Organizacional	Facilita la interoperabilidad mediante la integración de procesos administrativos y la cooperación entre entidades. Acciones y recomendaciones: 1) Simplificar y adaptar procesos administrativos para el intercambio digital. 2) Fomentar la colaboración interinstitucional a través de mesas de interoperabilidad. 3) Capacitar equipos internos en competencias relacionadas con la interoperabilidad para asegurar una implementación efectiva.
Semántico	Asegura que la información intercambiada sea uniforme y precisa mediante la estandarización de términos y un lenguaje común. Acciones y recomendaciones: 1) Implementar un lenguaje común de intercambio administrado por el Ministerio TIC para garantizar comprensión entre entidades. 2) Adoptar estándares internacionales para asegurar consistencia semántica. 3) Gestionar los datos de forma unificada, asegurando su interpretabilidad y reutilización.
Técnico	Aborda aspectos técnicos como la definición de arquitecturas y estándares que faciliten el intercambio eficiente, seguro y adaptable. Acciones y recomendaciones: 1) Definir arquitecturas estándar que permitan la conexión entre sistemas. 2) Implementar medidas de seguridad para proteger la integridad y disponibilidad de los datos durante la transmisión. 3) Reutilizar componentes tecnológicos para reducir costos y garantizar la neutralidad tecnológica.

Fuente: elaboración propia con base en Marco de Interoperabilidad de MinTIC.

Además, presenta un modelo de madurez que permite a las entidades evaluar su nivel de avance en la interoperabilidad que establece una hoja de ruta para alcanzar una interoperabilidad plena y principios que guíen su implementación:

- Priorizar servicios centrados en el ciudadano que mejoren la experiencia del usuario.
- Adaptar soluciones a las necesidades de cada entidad, garantizando cobertura y proporcionalidad.
- Proteger datos personales que permitan un intercambio seguro y confiable.
- Garantizar accesibilidad y adaptabilidad tecnológica en múltiples plataformas.

4. Conceptos clave

La siguiente tabla presenta los principales conceptos relacionados con el intercambio de información estadística. Estos términos son clave para garantizar la interoperabilidad, la protección de los datos y el cumplimiento de principios técnicos y normativos en el marco del SEN.

Tabla 4. Conceptos clave asociados al intercambio de información estadística

Concepto	Definición	Fuente
Anonimización de microdatos	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.	Decreto 1743 de 2016. Por el cual se reglamenta el artículo 160 de la ley 1753 de 2015 y se adiciona el título 3 a la parte 2 del libro 2 del Decreto 1170 de 2015 Único del Sector Administrativo de Información Estadística. Diario Oficial No. 50044 del 01 noviembre de 2016. https://www.suin-

RECOMENDACIONES DE MECANISMOS, ESTÁNDARES Y PROTOCOLOS PARA LA INTEROPERABILIDAD EN EL INTERCAMBIO, EL USO Y EL REÚSO DE DATOS EN EL SEN / REGULACIÓN ESTADÍSTICA

Concepto	Definición	Fuente
		juriscol.gov.co/viewDocument.asp?ruta=Decretos/30027053
Clasificación estadística (nomenclaturas y clasificaciones)	Se refiere a un conjunto de observaciones homogéneas, exhaustivas y mutuamente excluyentes que pueden ser asignadas a una o más variables a medir en la recopilación y/o la presentación de los datos, elaborada para la recolección y presentación de datos numéricos recolectados sistemáticamente (es decir, estadísticas). La utilidad de una clasificación estadística es perfeccionada si se basa en o representa una clasificación uniforme.	Organización de las Naciones Unidas (ONU). Glossary of statistical terms. https://unstats.un.org/unsd/classifications/bestpractices/glossary_short.pdf
Concepto estandarizado	Definición unificada basada en estándares estadísticos internacionales fundamentales para la armonización estadística, con el fin de proporcionar los elementos necesarios para la construcción y fortalecimiento de los marcos conceptuales. Este concepto debe estar armonizado en estándares estadísticos internacionales fundamentales para la armonización estadística, con el fin de proporcionar los elementos necesarios para la construcción y el fortalecimiento de los marcos conceptuales.	Departamento Administrativo Nacional de Estadística (DANE). (2023). Guía para la elaboración del documento metodológico de operaciones estadísticas. https://www.dane.gov.co/files/sen/lineamientos/Guia-elaboracion-documento-metodologico.pdf
Dato sensible	Información que afecta la intimidad de las personas o cuyo uso indebido puede generar discriminación.	Ministerio de Tecnologías de la Información y las Comunicaciones. Marco de interoperabilidad para Gobierno Digital.2019. http://lenguaje.mintic.gov.co/sites/default/files/archivos/marco_de_interoperabilidad_para_gobierno_digital.pdf
Datos abiertos	Todos aquellos datos primarios o sin procesar, que se encuentran en formatos estándar e interoperables que facilitan su acceso y reutilización, los cuales están bajo la custodia de las entidades públicas o privadas que cumplen con funciones públicas y que son puestos a disposición de cualquier ciudadano, de forma libre y sin restricciones, con el fin de que terceros puedan reutilizarlos y crear servicios derivados de los mismos.	Ley 1712 de 2014. Por medio de la cual se crea la Ley de Transparencia y del Derecho de Acceso a la Información Pública Nacional y se dictan otras disposiciones. Diario Oficial 49.084 de marzo 6 de 2014. https://www.suin-juriscol.gov.co/viewDocument.asp?ruta=Leyes/1687091
Datos estadísticos	Datos que son recolectados y/o generados por estadísticas ya sea en el proceso de las observaciones estadísticas, o por procesamiento de datos estadísticos.	European Statistical Office (Eurostat). (s.f.). Statistics explained glossary. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Data / United Nations Statistical Commission. (2000). (Terminology on Statistical Metadata). https://unece.org/DAM/stats/publications/53metadaterminology.pdf
Ecosistema de datos	Conjunto de actores, productores y los subsistemas departamental, municipal y de otras entidades territoriales y usuarios que presentan demandas y que deben asumir un rol activo en la producción de datos e interactuar en torno a los datos que se articulan para la producción, la accesibilidad y el uso de la información en un contexto institucional.	Departamento Administrativo Nacional de Estadística (DANE). (2020). <i>Metodología para el desarrollo de planes estadísticos</i> . https://www.dane.gov.co/files/sen/planificacion/metodologia/metodologia-desarrollo-planes-estadisticos.pdf
Estándar	Documento establecido por consenso y aprobado por un organismo reconocido que brinda, para uso común y repetido, un conjunto de reglas, líneas guía o características para las actividades o sus resultados, con el fin de lograr el grado óptimo de organización en un contexto.	ISO/IEC Guide 2:2004 y Ministerio de Tecnologías de la Información y las Comunicaciones. Marco de interoperabilidad para Gobierno Digital.2019. https://lenguaje.mintic.gov.co/sites/default/files/archivos/marco_de_interoperabilidad_para_gobierno_digital.pdf

RECOMENDACIONES DE MECANISMOS, ESTÁNDARES Y PROTOCOLOS PARA LA INTEROPERABILIDAD EN EL INTERCAMBIO, EL USO Y EL REÚSO DE DATOS EN EL SEN / REGULACIÓN ESTADÍSTICA

Concepto	Definición	Fuente
Gobernanza y administración de datos	Es la función a través de la cual el DANE articula las necesidades de información estadística asociadas a las políticas públicas de los miembros del SEN, partiendo de criterios técnicos para el análisis de los requerimientos, el respeto por la autonomía de las entidades, un marco ético para el uso e intercambio de información y garantías de protección de datos.	Ley 2335 de 2023. Por la cual se expiden disposiciones sobre las estadísticas oficiales en el país. Diario Oficial No. 52539 del 03 de octubre de 2023. https://www.suin-juriscol.gov.co/viewDocument.asp?ruta=Le yes/30050183#:~:text=Es%20la%20aplicaci%C3%B3n%20del%20conjunto.de%20inter%C3%A9s%20nacional%20o%20territorial.
Gobierno de datos	Autoridad sobre los datos a lo largo de todo su ciclo de generación de valor del proceso estadístico.	Departamento Administrativo Nacional de Estadística (DANE), Equipo técnico. 2021
Información estadística	Es el 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.	Ley 2335 de 2023. Por la cual se expiden disposiciones sobre las estadísticas oficiales en el país. Diario Oficial No. 52539 del 03 de octubre de 2023. https://www.suin-juriscol.gov.co/viewDocument.asp?ruta=Le yes/30050183#:~:text=Es%20la%20aplicaci%C3%B3n%20del%20conjunto.de%20inter%C3%A9s%20nacional%20o%20territorial.
Intercambio de datos	Cualquier mecanismo técnico y legal a través del cual una entidad (individuo u organización) puede acceder e intercambiar de manera efectiva los datos digitales, con el fin de favorecer la reutilización de estos últimos.	OECD. (2019). Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use Across Societies. OECD Publishing. https://doi.org/10.1787/276aaca8-en
Interoperabilidad	Capacidad de las organizaciones para intercambiar información y conocimiento en el marco de sus procesos de negocio para interactuar hacia objetivos mutuamente beneficiosos, con el propósito de facilitar la entrega de servicios en línea a ciudadanos, empresas y a otras entidades, mediante el intercambio de datos entre sus sistemas.	Ministerio de Tecnologías de la Información y las Comunicaciones. Marco de interoperabilidad para Gobierno Digital. 2019. http://lenguaje.mintic.gov.co/sites/default/files/archivos/marco_de_interoperabilidad_para_gobierno_digital.pdf
Mecanismo ²⁶	Manera de producir o de realizar una actividad, una función o un proceso, de forma planificada, que puede incluir una tecnología o un método.	Adaptado del Diccionario de Cambridge. https://dictionary.cambridge.org/dictionary/spanish-english/mecanismo
Metadatos	Información necesaria para el uso e interpretación de las estadísticas. Los metadatos describen la conceptualización, la calidad, la generación, el cálculo y las características de un conjunto de datos estadísticos.	Ley 2335 de 2023. Por la cual se expiden disposiciones sobre las estadísticas oficiales en el país. Diario Oficial No. 52539 del 03 de octubre de 2023. https://www.suin-juriscol.gov.co/viewDocument.asp?ruta=Le yes/30050183#:~:text=Es%20la%20aplicaci%C3%B3n%20del%20conjunto.de%20inter%C3%A9s%20nacional%20o%20territorial.
Microdatos	Corresponde a los datos sobre las características asociadas a las unidades de observación que se encuentran consolidadas en una base de datos.	Ley 2335 de 2023. Por la cual se expiden disposiciones sobre las estadísticas oficiales en el país. Diario Oficial No. 52539 del 03 de octubre de 2023. https://www.suin-juriscol.gov.co/viewDocument.asp?ruta=Le yes/30050183#:~:text=Es%20la%20aplicaci

²⁶ En el contexto de intercambio de información, los mecanismos y estándares están intrínsecamente relacionados, ya que los primeros se refieren a los procedimientos o métodos utilizados para facilitar el flujo de datos entre entidades, mientras que los estándares proporcionan un marco de referencia que asegura la comparabilidad y consistencia en dicho intercambio. Por ejemplo, un mecanismo es el uso de interfaces de programación de aplicaciones (APIs) para la transferencia de datos, mientras que un estándar como SDMX garantiza que los datos compartidos a través de estas interfaces sean comprensibles y utilizables por todas las partes involucradas.

RECOMENDACIONES DE MECANISMOS, ESTÁNDARES Y PROTOCOLOS PARA LA INTEROPERABILIDAD EN EL INTERCAMBIO, EL USO Y EL REÚSO DE DATOS EN EL SEN / REGULACIÓN ESTADÍSTICA

Concepto	Definición	Fuente
		https://www.suin-juriscal.gov.co/viewDocument.asp?ruta=Leyes/30050183#:~:text=Es%20la%20aplicaci%C3%B3n%20del%20conjunto.de%20inter%C3%A9s%20nacional%20o%20territorial.
Proceso estadístico	Conjunto sistemático de actividades encaminadas a la producción de estadísticas, entre las cuales están comprendidas: la detección de necesidades de información, el diseño, la construcción, la recolección, el procesamiento, el análisis, la difusión y la evaluación.	Ley 2335 de 2023. Por la cual se expiden disposiciones sobre las estadísticas oficiales en el país. Diario Oficial No. 52539 del 03 de octubre de 2023. https://www.suin-juriscal.gov.co/viewDocument.asp?ruta=Leyes/30050183#:~:text=Es%20la%20aplicaci%C3%B3n%20del%20conjunto.de%20inter%C3%A9s%20nacional%20o%20territorial.
Protocolo ²⁷	Constructo habilitante que describe las condiciones, el entorno y los métodos que permitan que la información fluya entre sistemas informáticos.	Northwestern University. (2016). Protocol for Exchange and Shared Responsibility for Institutional Data. https://www.it.northwestern.edu/docs/ExchangeSharedResponsibilityData.pdf
Reúso de datos	Utilizar datos para fines distintos a aquellos para los que fueron recolectados originalmente.	ELIXIR Research Data Management Kit. (s.f.). What is data reuse?. Recuperado el 26 de diciembre de 2024: https://rdmkit.elixir-europe.org/reusing
Reserva estadística ²⁸	Es la obligación legal del DANE y de las entidades del SEN, en el marco de la producción estadística, de garantizar que los datos que impliquen la identificación directa o por deducción de las fuentes primarias o secundarias de personas naturales o jurídicas estén restringidos al público en general, a las entidades públicas y privadas, a los organismos oficiales y a las autoridades públicas; dichos datos únicamente serán difundidos en resúmenes numéricos o microdatos anonimizados que no expongan información alguna de carácter individual que pudiera utilizarse con fines comerciales, de tributación fiscal, de investigación judicial o cualquier otro diferente del propiamente estadístico.	Ley 2335 de 2023. Por la cual se expiden disposiciones sobre las estadísticas oficiales en el país. Diario Oficial No. 52539 del 03 de octubre de 2023. https://www.suin-juriscal.gov.co/viewDocument.asp?ruta=Leyes/30050183#:~:text=Es%20la%20aplicaci%C3%B3n%20del%20conjunto.de%20inter%C3%A9s%20nacional%20o%20territorial.

Fuente: elaboración propia con base en sistema de consulta de conceptos estandarizados (https://conceptos.dane.gov.co/conceptos/area_tematica) y otros referentes.

²⁷ El protocolo también implica un contrato entre administradores en el que se comprometen a respetar el protocolo y ser responsables de su cumplimiento.

²⁸ La reserva estadística no aplica a la información y datos que las entidades del SEN deben suministrar al DANE. Asimismo, los productores de estadísticas oficiales que recopilan u obtengan datos individuales que se refieran a personas naturales o jurídicas deberán mantenerlos reservados y asegurar su reserva, conforme a las Leyes Estatutarias 1266 de 2008 y 1581 de 2012 y las normas que las modifiquen o sustituyan.

5. Referentes internacionales y nacionales

5.1. Referentes internacionales

5.1.1. Global Partnership for Sustainable Development Data (GPSDD) (Asociación Mundial de Datos para el Desarrollo Sostenible)

La guía *Interoperability: A practitioner's guide to joining-up data in the development sector* (Interoperabilidad: una guía para profesionales para unir datos en el sector del desarrollo)²⁹ ofrece recomendaciones que facilitan y mejoran la integración y la reusabilidad de los datos y los sistemas de datos. Estas se estructuran en torno a cinco dimensiones (Tabla 5).

Tabla 5. Dimensiones clave para fortalecer la interoperabilidad según GPSDD

Dimensión clave	Descripción	Recomendaciones y buenas prácticas
Interoperabilidad, Gestión de datos y Gobernanza	Establece un marco sólido para gestionar datos de manera coherente y responsable, garantizando la interoperabilidad a largo plazo.	Recomendaciones: desarrollar políticas claras que definan roles en la gestión de datos; crear comités de interoperabilidad. Buenas prácticas: gobernanza transparente (modelos centralizado, federado, colaborativo); políticas de acceso basadas en roles; capacitación continua.
Modelos de datos y de metadatos	Promueve el uso de modelos estandarizados para datos y metadatos, asegurando que sean comprensibles y utilizables uniformemente.	Recomendaciones: implementar estándares internacionales (SDMX, ISO 19115); documentar metadatos con esquemas como Dublin Core. Buenas prácticas: establecer modelos de datos consistentes y reutilizables; definir claramente los aspectos relevantes en los metadatos.
Clasificaciones y vocabularios	Enfatiza el uso de clasificaciones y vocabularios estandarizados para integrar datos de diversas fuentes, facilitando la interoperabilidad.	Recomendaciones: usar vocabularios controlados; adoptar esquemas SKOS y directrices SDMX. Buenas prácticas: usar vocabularios comunes proporcionados por organizaciones internacionales; asegurar consistencia y precisión mediante clasificaciones estándar.
Formatos de datos abiertos e interfaces estandarizadas	Se centra en el desarrollo de interfaces estandarizadas para la comunicación eficaz entre sistemas, facilitando el acceso y la integración de datos.	Recomendaciones: desarrollar API basadas en RESTful; usar formatos como CSV, JSON, XML; emplear serializaciones SDMX. Buenas prácticas: diseñar interfaces con medidas de seguridad como autenticación y cifrado; implementar API escalables y flexibles para acceso a datos.
Datos enlazados	Fomenta el uso de datos enlazados para mejorar la interoperabilidad conectando y enriqueciendo datos a través de múltiples fuentes.	Recomendaciones: adoptar estándares como RDF y SPARQL; usar repositorios de datos enlazados. Buenas prácticas: implementar técnicas de datos enlazados para enriquecer y conectar datos existentes; usar estándares como RDF, OWL, JSON-LD para serialización; emplear protocolos como OAI-PMH para recolección de metadatos.

Fuente: elaboración propia con base en GPSDD (2018).

²⁹ Ver documento disponible en: https://www.data4sdgs.org/sites/default/files/services_files/Interoperability%20-%20A%20practitioner%E2%80%99s%20guide%20to%20joining-up%20data%20in%20the%20development%20sector.pdf Wiki en línea: <https://unstats.un.org/wiki/display/InteropGuide/Home>

5.1.2. Data to Policy Navigator

El Navegador de Datos para la Formulación de Políticas³⁰ aborda la falta de capacidades en ciencia de datos entre los responsables de políticas públicas. Consta de cuatro pasos (Tabla 6) y ofrece guías detalladas para integrar datos en cada fase del ciclo de formulación de políticas.

Tabla 6. Pasos del Data to Policy Navigator para formulación de políticas públicas

Paso	Descripción
Paso 1. Definir el problema y necesidades de información	Identificar preguntas clave, comprender el problema, mapear el ecosistema de datos y detectar brechas que puedan limitar el potencial de las decisiones basadas en datos.
Paso 2. Acceder a los datos	Identificar fuentes confiables, recopilar datos considerando marcos legales y recursos disponibles, y establecer asociaciones efectivas para garantizar el acceso adecuado a los datos necesarios.
Paso 3. Analizar los datos	Limpiar, procesar y estructurar los datos para análisis precisos; validar e interpretar la información utilizando enfoques descriptivos, inferenciales y predictivos, y visualizar los datos para extraer conocimientos útiles.
Paso 4. Diseñar y evaluar la política	Convertir los conocimientos en acciones concretas, monitorear la implementación y evaluar el impacto de las políticas para garantizar resultados efectivos y basados en evidencia.

Fuente: elaboración propia con base en Data to Policy navigator 2023.

Consideraciones transversales

- Asegurar una ética y gobernanza de datos adecuada, promoviendo su uso reflexivo y ético.
- Desarrollar capacidades en datos, identificando roles, capacidades y brechas de habilidades.
- Fomentar una cultura de uso de datos en la formulación de políticas, implementando medidas para generar el cambio cultural necesario.
- Desarrollar mecanismos institucionales para el uso, reúso e intercambio continuo de datos, fortaleciendo su integración en la toma de decisiones a nivel regional, nacional y local.

5.1.3. Organización para la Cooperación y el Desarrollo Económico (OCDE)

Enhancing Access to and Sharing of Data (Mejorando el acceso y compartición de datos)³¹2019

Se centra en mejorar el acceso y la compartición de datos para maximizar los beneficios sociales, económicos y científicos, proteger derechos individuales y organizacionales y fomentar un entorno de datos más confiable y accesible.

Mecanismos de intercambio y el reúso de datos

³⁰ Disponible en: <https://www.datatopolicy.org/navigator>

³¹ Ver documento disponible en: <https://www.oecd-ilibrary.org/docserver/276aaca8-en.pdf?expires=1727731397&id=id&accname=quest&checksum=9AA5276F9C5CF55C3BCABA7CC0E05CAF>

- API (interfaces de programación de aplicaciones) que permiten el acceso y la transferencia de datos de manera segura y eficiente, manteniendo control sobre el acceso y uso de los datos. Los entornos controlados o “data sandboxes” ofrecen espacios seguros para pruebas y análisis sin comprometer información sensible para fomentar la innovación en las organizaciones.

Protocolos de interoperabilidad³² de datos

- Protocolos de transferencia de datos estandarizados que aseguran que los datos se puedan interpretar y utilizar independientemente de la plataforma de origen. Incluyen medidas que garantizan la inclusión de metadatos para facilitar la comprensión y reúso de datos.
- Ontologías y esquemas de metadatos como OWL y Dublin Core, asegurando que los datos puedan transferirse, reutilizarse y entenderse eficazmente en diversos sistemas y aplicaciones.

Estándares para el intercambio y reúso de datos

Adopción de estándares internacionales de clasificación y metadatos, como los ISO:

- ISO/IEC 19941:2017 asegura la interoperabilidad y portabilidad en servicios en la nube, permitiendo la transferencia y reúso de datos entre plataformas al clasificar los datos según su grado de identificabilidad, desde completamente identificados hasta anonimizados.
- ISO/IEC 20889:2018 establece técnicas de desidentificación, como enmascaramiento y generalización, para minimizar el riesgo de reidentificación, facilitando un intercambio seguro y responsable de datos que protege la privacidad.

5.1.4. Comisión Económica de las Naciones Unidas para Europa (UNECE)

Data Governance Framework for Statistical Interoperability (DAFI) (Marco de gobernanza de datos para la interoperabilidad estadística)³³ 2023

Brinda directrices que mejoran la interoperabilidad en sistemas estadísticos nacionales que incluye aspectos semánticos, estructurales, sintácticos y técnicos, garantizando coherencia y calidad en el intercambio de datos.

³² La OCDE resalta dos niveles clave: la sintáctica, que permite transferir datos entre sistemas en formatos compatibles para su decodificación, y la semántica, que garantiza la comprensión del significado de los datos en su contexto de uso.

³³ Disponible en: <https://www.efta.int/sites/default/files/images/publications/guide%20to%20sharing%20economic%20data.pdf>

Componentes clave para la interoperabilidad estadística

- Roles y gobernanza: definir roles específicos (como oficiales de datos y expertos en TI) para gestionar la interoperabilidad dentro de las organizaciones.
- Políticas legales y empresariales: implementar leyes y políticas que promuevan el uso de estándares comunes y el intercambio de datos seguros y eficientes.
- Estándares y herramientas: usar estándares internacionales como SDMX, DDI y GSIM para estructurar y gestionar datos y metadatos.

Facetas de la interoperabilidad

1. Semántica: uso de conceptos y clasificaciones comunes que aseguran una comprensión coherente.
2. Estructural: uso de formatos estándar que facilitan la identificación y uso de datos.
3. Sintáctica: protocolos de intercambio estandarizados que permiten la comunicación consistente.
4. Técnica: conectividad y operación entre sistemas usando estándares técnicos.

Desafíos en la interoperabilidad en el proceso de producción estadística según el GSBPM

La interoperabilidad enfrenta desafíos específicos en cada fase del proceso de producción estadística. La Tabla 7 resume estos retos conforme al modelo GSBPM, evidenciando los principales obstáculos técnicos y conceptuales que afectan la integración eficiente de datos.

Tabla 7. Desafíos en la interoperabilidad en el proceso de producción estadística según el GSBPM

Fase del Proceso Estadístico	Desafío de interoperabilidad
Detección y Análisis de Necesidades	La falta de alineación en conceptos primarios y herramientas de datos limita las oportunidades de interoperabilidad desde el inicio del proceso.
Diseño	La creación de clasificaciones y variables no estandarizadas impacta negativamente la integración de datos en fases posteriores.
Construcción	Las decisiones técnicas, como el uso de formatos incompatibles, dificultan la conectividad entre sistemas.
Recolección/Acopia	Las fuentes externas no controladas, como macrodatos, incrementan los riesgos de inconsistencias y reducen la calidad de los datos interoperables.
Procesamiento	La preparación de datos enfrenta problemas de interoperabilidad al carecer de conceptos y clasificaciones comunes.
Análisis	La falta de estándares compartidos obstaculiza la comparación de datos entre ciclos o con otras fuentes.
Difusión	La ausencia de clasificaciones comunes dificulta la publicación y reutilización de datos interoperables.

Fuente: elaboración propia con base *Data Governance Framework for Statistical Interoperability*

Recomendaciones

- Desarrollar una estrategia de interoperabilidad: definir objetivos claros y métricas para monitorear la implementación.
- Adoptar estándares abiertos: usar estándares internacionales que faciliten el intercambio de datos.
- Fomentar el cambio cultural: capacitar al personal para que adopten prácticas de interoperabilidad y colaboren efectivamente.

5.1.5. Portal de Datos Europeo

*Data.europa.eu Data Quality Guidelines*³⁴

Proporciona recomendaciones de calidad e interoperabilidad para los datos publicados en el portal de datos abiertos de la UE (Cuadro 1). Las recomendaciones están estructuradas en torno a los principios FAIR (Findability, Accessibility, Interoperability, Reusability)³⁵ que mejoran la calidad y utilidad de los datos y fomentan su intercambio y reutilización en múltiples plataformas y sistemas.

Los principios FAIR buscan garantizar que los datos sean encontrables (Findability) mediante el uso de metadatos claros y sistemas de identificación únicos; accesibles (Accessibility), asegurando que los datos estén disponibles a través de mecanismos abiertos y seguros; interoperables (Interoperability), permitiendo que los datos puedan ser integrados y utilizados con otros conjuntos de datos mediante el uso de formatos y vocabularios estándar; y reusables (Reusability), asegurando que los datos estén bien documentados para que puedan ser reutilizados fácilmente en diferentes contextos.

Cuadro 1. Recomendaciones para la interoperabilidad del Portal de datos europeo

- Uso de formatos estandarizados y específicos, como ISO 8601 para fechas y UTF-8 para caracteres, garantizando el procesamiento sin errores en sistemas diversos.
- Empleo de estándares internacionales como DCAT-AP para metadatos, facilitando la compatibilidad internacional y la integración con otras fuentes de datos.
- Uso de vocabularios controlados y estándares de la UE para asegurar la consistencia y claridad de los datos.
- Uso de identificadores y etiquetas en XML, así como URLs en RDF para denotar recursos y mejorar la interoperabilidad mediante el enlace de datos.
- Armonización de tablas y documentación de estructuras de datos utilizando esquemas JSON o XML.

³⁴ Directrices de calidad de datos de Data.europa.eu de 2021. Disponible en: <https://op.europa.eu/en/publication-detail/-/publication/023ce8e4-50c8-11ec-91ac-01aa75ed71a1>

³⁵ Disponible en: <https://www.go-fair.org/fair-principles/>

- Implementación de planes de gestión de datos (DMP) para definir estrategias de publicación, almacenamiento y descripción de datos y metadatos.
- Documentación de versiones de conjuntos de datos, diferenciando entre versiones mayores y menores, y manteniendo un historial de cambios claro para garantizar trazabilidad e interpretación a largo plazo.
- Eliminación de restricciones de acceso a los datos, promoviendo su descarga sin necesidad de registro.
- Provisión de enlaces de descarga directos y funcionales para fomentar la apertura y facilitar el intercambio y reutilización de los datos.

Fuente: elaboración propia con base en Data Quality Guidelines 2021

5.1.6. Conclusiones y recomendaciones a partir de la revisión de referentes internacionales

- La gobernanza de datos es fundamental para la interoperabilidad sostenible, requiriendo marcos institucionales robustos que definan roles, responsabilidades y mecanismos de coordinación entre las entidades del SEN.
- La adopción de estándares internacionales (SDMX, DDI, ISO/IEC 11179, Dublin Core) resulta crucial en la interoperabilidad semántica y técnica, facilitando el intercambio de información estadística entre sistemas heterogéneos.
- La implementación de los principios FAIR es un marco de referencia esencial que asegura que los datos estadísticos sean encontrables, accesibles, interoperables y reutilizables.
- La infraestructura tecnológica debe fundamentarse en protocolos estandarizados (REST, SOAP, HTTPS) y formatos abiertos (CSV, JSON, XML) que faciliten el intercambio eficiente de información entre las entidades del SEN.

Algunas recomendaciones para fortalecer la interoperabilidad en el intercambio de información:

- Establecer comités técnicos y mesas de interoperabilidad que coordinen esfuerzos entre instituciones y definir políticas claras de gestión de datos con roles y responsabilidades.
- Adoptar estándares internacionales como el SDMX y protocolos seguros con APIs estandarizadas, asegurando compatibilidad con formatos abiertos y un sistema unificado de metadatos.

- Desarrollar una plataforma centralizada para intercambio de datos estadísticos con "data sandboxes", implementando protocolos de seguridad y técnicas de desidentificación.
- Implementar procesos automatizados de validación de datos, vocabularios controlados, clasificaciones estandarizadas y aplicar los principios FAIR en la gestión de datos.
- Promover capacitación en interoperabilidad, comunidades de práctica, cultura de uso de datos y recursos accesibles sobre gestión de datos.
- Diseñar una hoja de ruta con indicadores medibles, garantizar recursos para la infraestructura, actualizar estándares continuamente y fortalecer la gestión del conocimiento.

5.2. Referentes nacionales

5.2.1. Plan Nacional de Infraestructura de Datos (PNID)

El PNID³⁶ es un documento técnico que establece las bases para el desarrollo y la gestión de la infraestructura de datos del Estado colombiano que optimice la toma de decisiones en el sector público.

Propone una hoja de ruta para su implementación y sostenibilidad mediante herramientas y lineamientos de distintas entidades (Tabla 8) incluido el DANE³⁷. Enfatiza la integración de diversas fuentes de datos, incluyendo datos abiertos y transaccionales, con un enfoque en la interoperabilidad y gobernanza de datos.

Tabla 8. Algunas directrices de MinTIC y de la SIC para la implementación del PNID³⁸ asociadas a la interoperabilidad, el aprovechamiento y la seguridad de los datos

Documento	Alcance
Marco de Interoperabilidad	Documento de MinTIC que establece principios y directrices para el intercambio seguro y eficiente de información entre entidades del Estado, mejorando la relación entre gobierno y ciudadanía. https://lenguaje.mintic.gov.co/sites/default/files/archivos/marco_de_interoperabilidad_para_gobierno_digital.pdf
Guía Uso Lenguaje Común de Intercambio de Datos	Guía de MinTIC que asegura que los datos compartidos entre entidades cumplan con principios de calidad, reutilización y estandarización, facilitando la interoperabilidad. https://estrategia.gobiernoenlinea.gov.co/623/articles-8240_Guia_Lenguaje.pdf

³⁶ Disponible en: https://infraestructuradatos.gov.co/798/articles-195195_recurso_1.pdf

³⁷ En el apartado de antecedentes se mostraron las principales instrumentos del DANE para promover la interoperabilidad así como la implementación del PNID.

³⁸ Consulte toda la información sobre los documentos normativos y técnicos del PNID: <https://infraestructuradatos.gov.co/798/w3-propertyvalue-379024.html>

RECOMENDACIONES DE MECANISMOS, ESTÁNDARES Y PROTOCOLOS PARA LA INTEROPERABILIDAD EN EL INTERCAMBIO, EL USO Y EL REÚSO DE DATOS EN EL SEN / REGULACIÓN ESTADÍSTICA

Documento	Alcance
Guía para el Uso y Aprovechamiento de Datos Abiertos	Guía de MinTIC que proporciona orientaciones para desarrollar estrategias de apertura y reutilización de datos abiertos, generando valor en diversos ámbitos sociales y económicos. https://herramientas.datos.gov.co/sites/default/files/2021-08/Guia%20de%20Datos%20Abiertos%20de%20Colombia.pdf
Guía de Estándares de Calidad e Interoperabilidad de Datos Abiertos	Guía de MinTIC que proporciona orientaciones y buenas prácticas para el desarrollo de estrategias de apertura de datos que cumplan con los principales estándares de calidad e interoperabilidad de datos abiertos. https://gobiernodigital.mintic.gov.co/692/articles-179118_recurso_5.pdf
Guía para el tratamiento de datos personales mediante servicio computación en la nube	Guía de la SIC que establece los principales aspectos a tener presente al usar servicios de computación en la nube considerando la regulación sobre tratamiento de datos personales. https://www.sic.gov.co/sites/default/files/files/2021/Guia%20cloud%20computing%202021.pdf
Guía sobre el tratamiento de datos personales en las entidades estatales	Guía de la SIC que presenta algunas sugerencias a las entidades estatales y a los servidores públicos que recolectan o tratan Datos Personales para que cumplan correctamente la regulación sobre el tratamiento de los mismos. https://www.sic.gov.co/sites/default/files/files/2021/SIC%20(2021)%20Gu%C3%ADa%20sobre%20el%20tratamiento%20de%20datos%20personales%20en%20las%20entidades%20estatales.pdf
Guía para la gestión de incidentes en seguridad de datos personales	Guía de la SIC que presenta algunas sugerencias a los responsables del tratamiento de Datos Personales para afrontar los incidentes de seguridad que afecten los datos personales bajo su custodia. Cualquier acción en este sentido debe centrarse en mitigar su impacto sobre los titulares de la información y sus datos. https://www.sic.gov.co/sites/default/files/files/Publicaciones/Guia_gestion_incidentes_dic21_2020.pdf
Modelo de seguridad y protección de información (MSPI)	Lineamientos de MinTIC para las entidades públicas en la implementación y adopción de buenas prácticas, basadas en estándares internacionales para la gestión adecuada del ciclo de vida de la seguridad de la información (Planeación, Implementación, Evaluación, Mejora Continua), permitiendo habilitar la adecuada ejecución de la Política de Gobierno Digital. https://www.mintic.gov.co/gestionti/615/articles-5482_Modelo_de_Seguridad_Privacidad.pdf

Fuente: elaboración propia con base en PNID.

6. Recomendaciones de interoperabilidad en el intercambio, el uso y el reúso de datos³⁹ en el SEN

Este apartado tiene dos secciones: el numeral 6.1 ofrece un marco metodológico basado en el *Data to Policy Navigator* para guiar al SEN en la gestión estratégica de datos, desde la identificación de necesidades hasta su transformación en políticas concretas. También ayuda a detectar brechas en el ecosistema de datos, promover el acceso a fuentes confiables y facilitar el reúso de información. El numeral 6.2. brinda herramientas para implementar este marco mediante las recomendaciones de mecanismos, estándares y protocolos para asegurar que los datos y la información generada a partir de estos sea accesible, interoperable, reutilizable y segura. Se reconoce que las condiciones para seleccionar mecanismos, estándares y protocolos pueden variar entre entidades, dependiendo de sus objetivos, capacidades técnicas y necesidades operativas.

³⁹ El reúso de datos significa utilizar datos para fines distintos a aquellos para los que fueron recolectados originalmente (ELIXIR Research Data Management Kit, s.f.)

6.1. Pasos recomendados para el acceso, el uso y el reúso de datos

PASO 1. Definir el problema y las necesidades de información.

- Definir el problema⁴⁰: identificar las preguntas clave de política, comprender profundamente el problema y determinar las necesidades de información.
- Mapear el ecosistema de datos⁴¹: identificar las fuentes de datos, los responsables de los datos, los usuarios y cómo interactúan para abordar las necesidades de información. Se sugiere que las entidades del SEN usen el Sistema de Identificación y Caracterización de Oferta y Demanda Estadística del SEN (SICODE) donde se puede conocer las características de los Registros Administrativos (RR.AA.) y Operaciones Estadísticas (OO.EE.) producidas por las entidades del sistema en: https://inventariosen.dane.gov.co/ooee/consulta_ooee
- Identificar brechas de datos (lagunas de datos)⁴²: detectar las posibles carencias de datos específicas del problema que podrían impedir que las políticas alcancen su máximo potencial. Además del SICODE se sugiere acudir a las mesas estadísticas sectoriales⁴³, que son espacios de articulación entre productores y usuarios de información estadística para identificar brechas de información y aprovechar registros administrativos.

PASO 2. Acceder a los datos.

- Identificar fuentes de datos confiables⁴⁴: explorar nuevas fuentes de datos que puedan ayudar a abordar las brechas identificadas y garantizar su confiabilidad. Se sugiere que las entidades del SEN usen el SICODE donde se puede conocer las características de los Registros Administrativos y Operaciones Estadísticas producidas por las entidades del sistema en: https://inventariosen.dane.gov.co/ooee/consulta_ooee
- Acceder a los datos y recopilarlos⁴⁵: obtener información sobre cómo realizar solicitudes significativas a los proveedores de datos, considerar aspectos presupuestarios y recursos, y entender

⁴⁰ 1. Para conocer más sobre cómo definir el problema consulte: <https://www.datatopolicy.org/navigator/1-1-define-the-problem#how-to-get-started> 2. Consulte las Recomendaciones del DANE para la identificación y el análisis de necesidades de información: <https://www.sen.gov.co/sites/default/files/pagina-migraciones-files/2024-05/recomendaciones-para-la-identificacion-y-el-analisis-de-necesidades-de-informacion.pdf>

⁴¹ Para conocer más sobre cómo identificar su ecosistema de datos consulte: <https://www.datatopolicy.org/navigator/map-your-data-ecosystem>

⁴² Para saber si ha identificado correctamente las lagunas de datos consulte: <https://www.datatopolicy.org/navigator/identify-data-gaps#how-will-i-know-i-have-successfully-identified-the-data-gaps>

⁴³ Consulte: <https://www.sen.gov.co/conozca-el-sen/instancias/comites-estadisticos-sectoriales>

⁴⁴ Para saber si hay identificado correctamente las fuentes de datos consulte: <https://www.datatopolicy.org/navigator/identify-data-sources-ensure-reliability>

⁴⁵ Para conocer cómo formular una solicitud de datos consulte: <https://www.datatopolicy.org/navigator/access-collect-data#formulating-a-data-request>

los marcos legislativos pertinentes. El CAD ayuda a gestionar con las entidades del SEN los requerimientos de microdatos que se necesiten para investigación, de forma ética, confiable y segura⁴⁶.

- Construir asociaciones de datos⁴⁷: explorar diferentes tipos de asociaciones de datos y aprender qué considerar al establecer una colaboración. Además, el DANE dispone de la *Guía para realizar intercambios de información (protocolos de intercambio)*⁴⁸ que detalla las etapas clave para implementar un proyecto de intercambio de datos.

PASO 3. Procesar y analizar los datos.

- Preparar y procesar los datos⁴⁹: limpiar y estructurar los datos para garantizar análisis precisos. Los responsables de políticas deben entender su importancia para colaborar con analistas.
- Validar, analizar e interpretar los datos⁵⁰: obtener una visión general de las principales consideraciones y errores comunes en el análisis y la interpretación de datos, guiando en el uso de análisis descriptivo, inferencial y predictivo para fundamentar decisiones con información fiable y comprensible.
- Documentar los procesos de transformación e integración de los datos realizados para un tema específico, a fin de asegurar la trazabilidad y transparencia del uso de la información en todas las etapas del análisis.
- Visualizar los datos⁵¹: aprender sobre diferentes formas de visualizar los datos para obtener mayores conocimientos e impacto. Una visualización correcta es precisa (datos actualizados y cálculos exactos), pertinente (responde a las preguntas del análisis) y clara (fácil de entender con etiquetas, títulos y leyendas bien definidos).
- Aplicar el marco ético del SETE (Sistema de Ética del Tratamiento Estadístico)⁵²: asegurar que durante el procesamiento y análisis de los datos se respeten principios como la no maleficencia, la transparencia, la proporcionalidad y la justicia, promoviendo un uso responsable de la información que minimice riesgos y genere valor público.

⁴⁶ Se pueden postular las necesidades de datos a través del siguiente enlace https://www.sen.gov.co/sites/default/files/pagina-migraciones-files/2023-10/formulario_de_requerimientos_de_informacion_CAD_v2.xlsx y enviarlo al correo sen@dane.gov.co

⁴⁷ Para conocer cómo establecer alianzas en materia de datos consulte: <https://www.datatopolicy.org/navigator/build-data-partnerships>

⁴⁸ Disponible en: https://www.sen.gov.co/sites/default/files/migracion-files/sen/lineamientos/Gu%C3%ADa_Protocolos_de_Intercambio.pdf

⁴⁹ Para conocer los pasos relevantes en el preprocesamiento de datos consulte: <https://www.datatopolicy.org/navigator/prepare-process-the-data#relevant-steps-of-data-pre-processing>

⁵⁰ Para saber si ha analizado e interpretado correctamente los datos consulte: <https://www.datatopolicy.org/navigator/validate-analyze-interpret-the-data>

⁵¹ Para ver un catálogo de visualizaciones consulte: <https://datavizcatalogue.com/ES/>

⁵² Disponible en <https://www.sen.gov.co/sites/default/files/pagina-migraciones-files/2024-02/marco-etico-de-los-datos-vf.pdf>

- Garantizar que los datos y metadatos que se difundan cumplan con las normas de confidencialidad, reserva estadística y protección de datos personales⁵³: controlar el riesgo de identificación de la fuente mediante la anonimización⁵⁴ y asegurar la confidencialidad⁵⁵.

PASO 4. Diseñar y evaluar el uso y reúso de datos.

- Transformar datos en acciones⁵⁶: utilizar los conocimientos derivados de los datos para informar decisiones concretas, ya sea en políticas, proyectos, servicios o estrategias de cualquier ámbito.
- Monitorear el impacto del uso de datos⁵⁷: implementar herramientas y metodologías para evaluar cómo los datos están siendo utilizados y el valor que generan en el contexto de su aplicación, como mejoras en procesos, resultados tangibles o impacto social.
- Evaluar resultados y aprendizaje⁵⁸: realizar evaluaciones periódicas del impacto del uso y reúso de datos, considerando tanto las métricas directas (eficiencia, calidad, alcance) como los aprendizajes obtenidos para optimizar su aplicación futura.

⁵³ 1. Consulte la política de tratamiento y protección de datos personales del DANE: <https://www.dane.gov.co/files/images/ventana-unica/documentos/politicadedatosdane.pdf> 2. Consulte la Política general de seguridad y privacidad del DANE: <https://www.dane.gov.co/files/acerca/Normatividad/resoluciones/2020/politica-de-seguridad-y-privacidad-DANE-FONDANE.pdf> 3. Consulte el modelo de MinTIC y las guías de la SIC sobre protección de datos personales: Modelo de seguridad y protección de información (MSPI) de MinTIC: https://www.mintic.gov.co/gestionti/615/articles-5482_Modelo_de_Seguridad_Privacidad.pdf
Guía SIC para el tratamiento de datos personales mediante servicio computación en la nube: <https://www.sic.gov.co/sites/default/files/files/2021/Guia%20cloud%20computing%202021.pdf>
Guía sobre el tratamiento de datos personales en las entidades estatales: [https://www.sic.gov.co/sites/default/files/files/2021/SIC%20\(2021\)%20Gu%C3%ADa%20sobre%20el%20tratamiento%20de%20datos%20personales%20en%20las%20entidades%20estatales.pdf](https://www.sic.gov.co/sites/default/files/files/2021/SIC%20(2021)%20Gu%C3%ADa%20sobre%20el%20tratamiento%20de%20datos%20personales%20en%20las%20entidades%20estatales.pdf)
Guía SIC para la gestión de incidentes en seguridad de datos personales: https://www.sic.gov.co/sites/default/files/files/Publicaciones/Guia_gestion_incidentes_dic21_2020.pdf

⁵⁴ Revisar las Guías de anonimización de datos estructurados del DANE <https://www.sen.gov.co/sites/default/files/pagina-migraciones-files/2024-07/guia-de-anonimizacion-de-datos-estructurados.pdf> y del Archivo General de la Nación https://www.archivogeneral.gov.co/sites/default/files/Estructura_Web/5_Consulte/Recursos/Publicaciones/2021-06-25_Guia_de_Anonimizacion.pdf

⁵⁵ Algunas leyes claves son: Ley 2335 de 2023 (Por la cual se expiden disposiciones sobre las estadísticas oficiales en el país); Ley estatutaria 1266 de 2008 (Por la cual se dictan las disposiciones generales del hábeas data y se regula el manejo de la información contenida en bases de datos personales, en especial la financiera, crediticia, comercial, de servicios y la proveniente de terceros países y se dictan otras disposiciones), Ley estatutaria 1581 de 2012 (Por la cual se dictan disposiciones generales para la protección de datos personales) y Ley 1712 de 2014 (Por medio de la cual se crea la Ley de Transparencia y del Derecho de Acceso a la Información Pública Nacional y se dictan otras disposiciones) y las normas que las modifiquen o sustituyan.

⁵⁶ Ejemplo de cómo transformar los datos en políticas: <https://www.datatopolicy.org/navigator/turn-data-into-policy>

⁵⁷ Ejemplo de cómo supervisar la aplicación de políticas: <https://www.datatopolicy.org/navigator/monitor-policy-implementation>

⁵⁸ Ejemplo de cómo evaluar el impacto de políticas: <https://www.datatopolicy.org/navigator/evaluate-policy-impact>

6.2. Sugerencias de mecanismos, estándares y protocolos para el intercambio, el uso y el reúso de datos organizados por dimensión de interoperabilidad

Presenta algunas sugerencias de mecanismos, estándares y protocolos considerando los referentes internacionales y nacionales, así como los lineamientos DANE y de algunas entidades del SEN. Las sugerencias se organizan de acuerdo con las cinco dimensiones clave propuestas por la GPSDD para fortalecer la interoperabilidad estadística y el intercambio de información: Gestión de datos, Gobernanza e Interoperabilidad; Modelos de datos y metadatos; Clasificaciones y vocabularios; Formatos de datos abiertos e interfaces estandarizadas, y Datos enlazados. También, [el Anexo A⁵⁹](#) de este documento muestra algunos vocabularios, métodos, formatos, marcos, lenguajes, flujos de trabajo y modelos de datos estandarizados para uso de las entidades del SEN.

Dimensión 1. Gestión de datos (seguridad, privacidad y confidencialidad), Gobernanza e Interoperabilidad

Mecanismos

- Definición de una política de gobernanza de datos que incluya roles y responsabilidades para asegurar que cada etapa del ciclo de vida de los datos sea supervisada y gestionada correctamente. Esta política debe contemplar aspectos de interoperabilidad y uso ético.
- Desarrollo de una estrategia de interoperabilidad y monitoreo: incluye definir la visión y los beneficios, identificar grupos de interés, establecer claramente los roles y responsabilidades, implementar estándares de interoperabilidad y realizar un monitoreo continuo (Tabla 7). Se puede iniciar con un proyecto piloto para facilitar el aprendizaje y la adaptación.

Tabla 9. Ejemplos de métricas para monitorear la interoperabilidad estadística en organizaciones

Indicador	Descripción
Porcentaje de procesos estadísticos que usan estándares y definiciones comunes.	Grado en que se utilizan estándares comunes para recolectar, procesar y difundir datos.
Porcentaje de procesos estadísticos que utilizan las mismas herramientas de software en subprocesos específicos.	Uso de software generalizado en lugar de software ad-hoc (por ejemplo, partiendo de herramientas estándar para el intercambio de datos como SDMX).
Porcentaje de procesos estadísticos que utilizan un sistema de metadatos estándar:	Nivel de uso de un sistema de metadatos estandarizado.
Porcentaje de procesos estadísticos cuyos datos se difunden como Datos Abiertos Enlazados (LOD)	Grado en que los datos estadísticos están integrados semánticamente con otros datos.
Porcentaje de procesos estadísticos que pueden compartir datos entre sí sin problemas	Grado en el que diferentes procesos son capaces de compartir datos entre ellos sin necesidad de intervención manual.

⁵⁹ Este anexo es tomado del documento de la UNECE *Data Governance Framework for Statistical Interoperability (DAFI)* disponible en: https://unece.org/sites/default/files/2024-03/HLG2023%20DAFI%20Final_0.pdf

RECOMENDACIONES DE MECANISMOS, ESTÁNDARES Y PROTOCOLOS PARA LA INTEROPERABILIDAD EN EL INTERCAMBIO, EL USO Y EL REÚSO DE DATOS EN EL SEN / REGULACIÓN ESTADÍSTICA

Indicador	Descripción
Porcentaje de procesos estadísticos que están automatizados	Grado en el que la interoperabilidad permite la automatización de los procesos estadísticos. Una interoperabilidad eficiente debería reducir las intervenciones manuales y optimizar los flujos de trabajo.
Satisfacción de los grupos de interés y usuarios con la interoperabilidad de los procesos estadísticos.	Información recopilada a través de encuestas u otros mecanismos de retroalimentación para evaluar cuán satisfechos están los usuarios con la capacidad de acceder y utilizar los datos desde múltiples fuentes.

Fuente: elaboración propia con base en UNECE (2022) (DAFI).

- Implementación de estándares abiertos: para crear un ecosistema de datos interconectado y eficiente, facilitando la interoperabilidad y fomentando la colaboración entre entidades. El DANE orienta en la implementación de estos estándares a través de lineamientos, guías, manuales, así como cursos virtuales y capacitaciones disponibles para las entidades del SEN⁶⁰. Consulte también [el Anexo A](#) de este documento.
- Identificación de leyes y regulaciones aplicables⁶¹: revisar y documentar las normativas que rigen la producción y difusión de estadísticas oficiales, y las relacionadas con la gestión, protección y confidencialidad de los datos.
- Aplicación de controles de divulgación de la información estadística: para asegurar que los datos y metadatos que serán difundidos cumplan con las reglas de confidencialidad, reserva estadística y protección de datos personales.

Consulte: Política de seguridad y privacidad del DANE⁶², Política de tratamiento y protección de datos personales el DANE⁶³, así como otros documentos SEN: *Modelo de seguridad y protección de información* (MSPI) de MinTIC⁶⁴; la *Circular única título V protección de datos personales*⁶⁵ y las guías de la SIC: *Guía para el tratamiento de datos personales mediante servicio computación en la nube*⁶⁶,

⁶⁰ Disponibles en: <https://www.sen.gov.co/normatividad/lineamientos-guias-manuales> Cursos virtuales: <https://www.sen.gov.co/servicios/aula-digital> y capacitaciones: <https://www.sen.gov.co/servicios/plan-capacitacion> Nomenclaturas y clasificaciones estadísticas: <https://www.dane.gov.co/index.php/sistema-estadistico-nacional-sen/normas-y-estandares/nomenclaturas-y-clasificaciones>

⁶¹ Algunas leyes claves son: Ley 2335 de 2023 (Por la cual se expiden disposiciones sobre las estadísticas oficiales en el país); Ley estatutaria 1266 de 2008 (Por la cual se dictan las disposiciones generales del hábeas data y se regula el manejo de la información contenida en bases de datos personales, en especial la financiera, crediticia, comercial, de servicios y la proveniente de terceros países y se dictan otras disposiciones), Ley estatutaria 1581 de 2012 (Por la cual se dictan disposiciones generales para la protección de datos personales) y Ley 1712 de 2014 (Por medio de la cual se crea la Ley de Transparencia y del Derecho de Acceso a la Información Pública Nacional y se dictan otras disposiciones) y las normas que las modifiquen o sustituyan. Además, el intercambio a nivel tecnológico debe ceñirse al Marco para la interoperabilidad del Ministerio de las Tecnologías de la Información y Comunicaciones (https://lenguaje.mintic.gov.co/sites/default/files/archivos/marco_de_interoperabilidad_para_gobierno_digital.pdf) y a los preceptos constitucionales y legales sobre protección de datos personales.

⁶² Disponible en: <https://www.dane.gov.co/files/acerca/Normatividad/resoluciones/2020/politica-de-seguridad-y-privacidad-DANE-FONDANE.pdf>

⁶³ Disponible en: <https://www.dane.gov.co/files/images/ventana-unica/documentos/politicadedatosdane.pdf>

⁶⁴ Disponible en: https://www.mintic.gov.co/gestionti/615/articles-5482_Modelo_de_Seguridad_Privacidad.pdf

⁶⁵ Disponible en: <https://www.sic.gov.co/sites/default/files/normatividad/092022/T%C3%ADtulo%20V%20Versi%C3%B3n%2029-09-2022.pdf>

⁶⁶ Disponible en: <https://www.sic.gov.co/sites/default/files/files/2021/Guia%20cloud%20computing%202021.pdf>

*Guía sobre el tratamiento de datos personales en las entidades estatales*⁶⁷, *Guía para la gestión de incidentes en seguridad de datos personales*⁶⁸.

- Anonimización de datos personales mediante herramientas y metodologías que permiten eliminar o modificar información identificable en los conjuntos de datos. Se sugiere consultar la *Guía para la Anonimización de Datos Estructurados*⁶⁹ del DANE, así como el *Caso práctico de anonimización de bases de datos en el SEN*⁷⁰. También la *Guía de anonimización de datos estructurados* del Archivo General de la Nación⁷¹ y revisar los estándares ISO/IEC 19941:2017 e ISO/IEC 20889:2018.
- Implementación de los elementos tecnológicos mínimos para el intercambio, integración y gestión de microdatos de acuerdo con el Marco de Interoperabilidad. Se sugiere considerar también la *Guía Uso Lenguaje Común de Intercambio de Datos*⁷² de MinTIC que asegura que los datos compartidos entre entidades cumplan con principios de calidad, reutilización y estandarización, facilitando la interoperabilidad.
- Postulación de las necesidades de datos al Comité de Administración de Datos: el CAD es mediador en el intercambio y ayuda a gestionar con las entidades del SEN los requerimientos de microdatos que se necesiten para investigación, de forma ética, confiable y segura. Se debe completar este formulario: https://www.sen.gov.co/sites/default/files/pagina-migraciones-files/2023-10/formulario_de_requerimientos_de_informacion_CAD_v2.xlsx y enviarlo a sen@dane.gov.co
- Uso de modelos de madurez o IMM (por sus siglas en inglés) para medir la interoperabilidad que definen niveles para caracterizar el estado de campos o áreas específicas. En el ámbito nacional el MinTIC cuenta con el modelo de madurez del Marco de interoperabilidad⁷³ mientras que en el ámbito internacional algunos ejemplos son: Programa ISA de la Comisión Europea – IMM⁷⁴, IMM de la Oficina de Información Científica y Técnica (OSTI) del Departamento de Energía de los EE. UU.

⁶⁷ Disponible en:

[https://www.sic.gov.co/sites/default/files/files/2021/SIC%20\(2021\)%20Gu%C3%ADa%20sobre%20el%20tratamiento%20de%20datos%20personales%20en%20las%20entidades%20estatales.pdf](https://www.sic.gov.co/sites/default/files/files/2021/SIC%20(2021)%20Gu%C3%ADa%20sobre%20el%20tratamiento%20de%20datos%20personales%20en%20las%20entidades%20estatales.pdf)

⁶⁸ Disponible en: https://www.sic.gov.co/sites/default/files/files/Publicaciones/Guia_gestion_incidentes_dic21_2020.pdf

⁶⁹ Disponible en: <https://www.sen.gov.co/sites/default/files/pagina-migraciones-files/2024-07/guia-de-anonimizacion-de-datos-estructurados.pdf>

⁷⁰ Disponible en: https://www.sen.gov.co/sites/default/files/migracion-files/RegistrosAdministrativos/CASO_PR%C3%81CTICO_GU%C3%8DA_PARA_LA_ANONIMIZACI%C3%93N_DE_BASES_DE_DATOS_EN_EL_SISTEMA_ESTAD%C3%8DSTICO_NACIONAL.xlsx

⁷¹ Disponible en:

https://www.archivogeneral.gov.co/sites/default/files/Estructura_Web/5_Consulte/Recursos/Publicaciones/Guia_de_Anonimizacion-min.pdf

⁷² Disponible en: https://estrategia.gobiernoenlinea.gov.co/623/articles-8240_Guia_Lenguaje.pdf

⁷³ Disponible en: https://lenguaje.mintic.gov.co/sites/default/files/archivos/marco_de_interoperabilidad_para_gobierno_digital.pdf

⁷⁴ Mide cómo las administraciones públicas interactúan con entidades externas para proporcionar servicios públicos de manera eficiente. Se enfoca en tres dominios de interoperabilidad: entrega, consumo y gestión de servicios, con un modelo de cinco etapas para evaluar la madurez. Disponible en: <https://interoperable-europe.ec.europa.eu/collection/semic-support-centre/document/interoperability-maturity-model>

(DOE)⁷⁵, IMM de Datos de los Archivos Nacionales de Australia⁷⁶ y la Evaluación de Madurez de GPSDD-UNSD⁷⁷.

Implementación de modelos y estructuras de procesos: El DANE ha puesto a disposición del SEN los *Lineamientos para el Proceso Estadístico en el SEN*⁷⁸ que está basado en el GSBPM 5.1 publicado por la UNECE, que proporciona las directrices para la producción estadística por parte de las entidades que conforman el SEN.

- Uso de mecanismos de evaluación y control de calidad: el DANE ha puesto a disposición del SEN la Norma Técnica de la Calidad del proceso Estadístico NTC PE 1000:2020⁷⁹ que establece los requisitos que se evalúan para garantizar la calidad del proceso estadístico, para las entidades del SEN que producen y difunden estadísticas.

Otros mecanismos incluyen el uso de herramientas automáticas para perfilado, validación y limpieza de datos antes de su publicación. Para estándares abiertos en perfilado y validación, se pueden emplear esquemas como JSON Schema o XML Schema (XSD).

- Uso de Plataformas y API estandarizadas: herramientas como SDMX-API y REST API permiten el acceso y la reutilización de datos en formatos abiertos como JSON y XML, facilitando la interoperabilidad técnica. Estas API aseguran que los datos sean accesibles y reutilizables por múltiples sistemas y usuarios.
- Uso de Formatos Accionables por Máquinas (MAF): formatos diseñados para que las máquinas accedan, compartan y utilicen los datos de manera automática, ayudando a automatizar tareas como la extracción y limpieza de datos, mejorando la gobernanza y la seguridad y liberando recursos humanos para tareas estratégicas.

⁷⁵ Dirigido a integración tecnológica, evalúa seis categorías principales: configuración, seguridad, operación, organizacional, informacional y técnica. También incluye una categoría adicional de "comunidad" para abordar cambios culturales y colaboración, basándose en el Capability Maturity Model Integration (CMMI) (Integración de modelos de madurez de capacidades). Disponible en: <https://www.osti.gov/biblio/1804457>

⁷⁶ Evalúa la gobernanza de datos en cinco temas clave (negocio, seguridad, legal, semántico, técnico) mediante un modelo de cinco pasos que describe comportamientos y procesos comunes de interoperabilidad. Disponible en: <https://www.naa.gov.au/information-management/build-data-interoperability/data-interoperability-maturity-model>

⁷⁷ Diseñada para estadísticos y profesionales del desarrollo sostenible, evalúa cuatro capas de interoperabilidad (organizacional, humana, de datos y tecnológica), cada una con sus dimensiones y niveles de madurez. Disponible en: <https://www.data4sdgs.org/resources/joined-data-maturity-assessment>

⁷⁸ Disponible en: https://www.sen.gov.co/sites/default/files/migracion-files/sen/normatividad/Lineamientos_Proceso_Estad%C3%ADstico_v2.pdf

⁷⁹ Disponible en: <https://www.sen.gov.co/sites/default/files/migracion-files/RegulacionEstadistica/NTC%20PE%201000-2020.pdf>

- Desarrollo de mecanismos institucionales para el uso continuado de los datos que implique el establecimiento de procesos claros, la creación de capacidades y el fomento de una cultura de toma de decisiones basada en pruebas⁸⁰.
- Fomento del cambio cultural y superación de la resistencia organizacional hacia la adopción de la interoperabilidad: mediante capacitación al personal en tecnologías, ajuste de flujos de trabajo para colaboración, apoyo de la alta dirección e involucramiento de grupos de interés.
- Uso de la inteligencia artificial (IA) y el aprendizaje automático (ML) para mejorar la calidad de los datos y automatizar tareas en los procesos estadísticos. Estas tecnologías corrigen errores mediante imputación predictiva y detección de anomalías, permiten análisis en tiempo real de grandes volúmenes de datos, reducen costos y optimizan la accesibilidad con interfaces intuitivas.⁸¹
- Uso de tecnologías emergentes como grafos, blockchain e inteligencia artificial federada para fortalecer la integridad, transparencia y seguridad en la gestión de datos. Los grafos permiten analizar relaciones complejas de forma eficiente, blockchain asegura la inmutabilidad y auditabilidad de los datos, y la inteligencia artificial federada protege la privacidad al minimizar la transferencia de datos sensibles. Estas herramientas mejoran la trazabilidad, la gobernanza y el cumplimiento normativo en privacidad y protección de datos⁸¹.

Estándares

- Normas técnicas y marcos de calidad: la *Norma Técnica de la Calidad del Proceso Estadístico (NTC PE 1000:2020)* establece los requisitos para evaluar y garantizar la calidad en el proceso estadístico, alineándose con los estándares internacionales de interoperabilidad.
- Modelo Genérico del Proceso Estadístico (GSBPM): un estándar internacionalmente reconocido para estructurar y estandarizar los procesos de producción estadística, mejorando la interoperabilidad entre las entidades. El DANE tomó como base este modelo para elaborar los *Lineamientos del Proceso Estadístico en el SEN* para orientar la producción y difusión estadística en el SEN.
- ISO 38500: directrices para la gobernanza de TI y datos en organizaciones estadísticas.

⁸⁰ Para conocer más sobre estos mecanismos consulte: <https://www.datatopolicy.org/considerations/develop-institutional-mechanisms-for-continued-use-of-data>

⁸¹ Consulte: *Marco de clasificación de sistemas de IA de la OCDE* (2022) disponible en: https://www.oecd.org/en/publications/oecd-framework-for-the-classification-of-ai-systems_cb6d9eca-en.html

Sinergias de IA, gobernanza de datos y privacidad y áreas de cooperación internacional (2024) disponible en https://www.oecd.org/en/publications/ai-data-governance-and-privacy_2476b1a4-en.html

- ISO/IEC 19941:2017 (anonimización): asegura la interoperabilidad y la portabilidad en servicios en la nube, permitiendo la transferencia y reúso de datos entre plataformas al clasificar los datos, según su grado de identificabilidad, desde completamente identificados hasta anonimizados.
- ISO/IEC 20889:2018 (anonimización): establece técnicas de desidentificación, como enmascaramiento y generalización, para minimizar el riesgo de reidentificación y así facilitar un intercambio seguro y responsable de datos que protege la privacidad.
- DCAT (Data Catalog Vocabulary) y Dublin Core: estándares que permiten la descripción y estructuración de catálogos de datos y metadatos, facilitando su accesibilidad y reutilización. El DANE cuenta con la *Guía para la implementación de los estándares estadísticos para la documentación de metadatos y microdatos (Data Documentation Initiative (DDI) y Dublin Core (DC))*⁸².
- RDF, SPARQL, OWL, SKOS: estándares del World Wide Web Consortium (W3C) para la interoperabilidad semántica, utilizados para representar y compartir datos en la Web Semántica.
- SDMX (Statistical Data and Metadata Exchange): estándar para el intercambio estructurado de datos y metadatos entre las instituciones estadísticas. El DANE cuenta con la *Guía para la implementación del estándar estadístico de intercambio de datos y metadatos (SDMX)*⁸³.

Protocolos

- Memorandos de Entendimiento (MOU) y Acuerdos de Nivel de Servicio (SLA): formalizan acuerdos entre entidades para compartir datos bajo reglas comunes garantizando una gestión efectiva y segura de los datos estadísticos.
- Protocolo de intercambio de información estadística (artículo 17 Ley 2335 de 2023): los miembros del SEN podrán intercambiar información estadística correspondiente a datos agregados y a microdatos de forma gratuita y oportuna en desarrollo de los objetivos del SEN, en los términos del artículo 27 del Código de Procedimiento Administrativo y de lo Contencioso Administrativo y demás normas concordantes, así como de las Leyes Estatutarias 1266 de 2008 y 1581 de 2012 y las normas que las modifiquen o sustituyan.
- Políticas de tratamiento y protección de datos personales: donde se definen lineamientos internos y responsabilidades institucionales conforme a la Ley 1581 de 2012, y se comunica a los interesados los procedimientos, recursos y canales para garantizar sus derechos sobre el tratamiento de datos personales. Consulte la *Política de tratamiento y protección de datos personales del DANE y FONDANE*⁸⁴.

⁸² Disponible en: https://www.sen.gov.co/sites/default/files/migracion-files/guias/Guia_para_implementar_DDI_y_DC.pdf

⁸³ Disponible en: https://www.sen.gov.co/sites/default/files/migracion-files/sdmx/Gu%C3%ADa_Implementaci%C3%B3n_SDMX.pdf

⁸⁴ Disponible en: <https://www.dane.gov.co/files/images/ventana-unica/documentos/politicadedatosdane.pdf>

- Políticas de seguridad y privacidad de la información: que definen lineamientos y responsabilidades institucionales para proteger la información, asegurando su confidencialidad, integridad y disponibilidad, cumpliendo con las normativas para evitar sanciones y daños a la imagen institucional, y minimizando los riesgos de seguridad. Consulte la *Política de seguridad y privacidad de la información del DANE y FONDANE*⁸⁵.
- Protocolos de anonimización: Se sugiere consultar la Guía para la Anonimización de Datos Estructurados⁸⁶ del DANE, el Caso práctico de anonimización de bases de datos en el SEN⁸⁷. También la Guía de anonimización de datos estructurados del Archivo General de la Nación⁸⁸ así como revisar los estándares ISO/IEC 19941:2017 e ISO/IEC 20889:2018.
- Protocolos de autenticación y seguridad (HTTPS, OAuth): aseguran la transmisión segura de datos y la autenticación de usuarios, protegiendo la integridad y la confidencialidad de los datos durante su intercambio.
- Protocolos de intercambio estandarizado (SEP): protocolos como SDMX, JSON, y CSV, que definen cómo deben intercambiarse los datos entre diferentes sistemas, reduciendo los costos y mejorando la calidad y la trazabilidad de los datos. Consulte la *Guía para realizar intercambios de información (protocolos de intercambio)* del DANE⁸⁹ en el contexto del SDMX.
- Políticas de datos abiertos: establecen las condiciones para que los datos pueden ser compartidos y reutilizados, garantizando la confidencialidad y la privacidad. Revise la *Guía para el uso y el aprovechamiento de Datos Abiertos en Colombia* de MinTIC⁹⁰ así como la *Guía de Estándares de Calidad e Interoperabilidad de Datos Abiertos*⁹¹.

⁸⁵ Disponible en: <https://www.dane.gov.co/files/acerca/Normatividad/resoluciones/2020/politica-de-seguridad-y-privacidad-DANE-FONDANE.pdf>

⁸⁶ Disponible en: <https://www.sen.gov.co/sites/default/files/pagina-migraciones-files/2024-07/guia-de-anonimizacion-de-datos-estructurados.pdf>

⁸⁷ Disponible en: https://www.sen.gov.co/sites/default/files/migracion-files/RegistrosAdministrativos/CASO_PR%C3%81CTICO_GU%C3%8DA_PARA_LA_ANONIMIZACI%C3%93N_DE_BASES_DE_DATOS_EN_EL_SISTEMA_ESTAD%C3%8DSTICO_NACIONAL.xlsx

⁸⁸ Disponible en: https://www.archivogeneral.gov.co/sites/default/files/Estructura_Web/5_Consulte/Recursos/Publicaciones/Guia_de_Anonimizacion-min.pdf

⁸⁹ Disponible en: https://www.sen.gov.co/sites/default/files/migracion-files/sen/lineamientos/Gu%C3%ADa_Protocolos_de_Intercambio.pdf

⁹⁰ Disponible en: <https://herramientas.datos.gov.co/sites/default/files/Guia%20de%20Datos%20Abiertos%20de%20Colombia.pdf>

⁹¹ Disponible en: https://gobiernodigital.mintic.gov.co/692/articles-179118_recurso_5.pdf

- Políticas de calidad de datos: directrices internas que definen procedimientos y responsabilidades para asegurar que los datos cumplan con los estándares de calidad. Consulte el *Marco de Aseguramiento de la Calidad (MAC)*⁹² y el *Código Nacional de Buenas Prácticas del SEN*⁹³.
- API estandarizadas (REST, SDMX-API, CKAN API): protocolos que permiten la automatización del acceso a los datos mediante interfaces estándar y así asegurar que múltiples sistemas puedan interactuar con los datos de manera consistente.

Cuadro 2. Retos en la gestión de datos, gobernanza e interoperabilidad

- No adoptar un enfoque organizacional para la gestión y la gobernanza de datos relegando dichos temas exclusivamente al departamento de TI, en lugar de abordarlos desde una perspectiva organizacional más amplia.
- Falta de una cadena clara de responsabilidad al no desarrollar ni aplicar una estructura de responsabilidad clara que especifique roles y funciones entre los distintos departamentos para garantizar una gobernanza de datos eficaz entre y dentro de las organizaciones.
- Ignorar la interoperabilidad al actualizar o adquirir nuevos sistemas de TI, lo que resulta en la creación de silos de datos internos y en la gestión de datos en formatos y esquemas incompatibles.
- No aprovechar las herramientas legales y regulatorias para crear un entorno seguro y estructurado para compartir e integrar datos, respetando la privacidad, la protección de datos y las consideraciones de seguridad.

Fuente: elaborado con base en *Data Interoperability: A Practitioner's Guide to Joining up data in the development sector*.

Dimensión 2. Modelos de datos y de metadatos

Mecanismos

- Usar modelos multidimensionales (Data Cube)⁹⁴: permiten la organización de datos en diversas dimensiones como tiempo, localización o categoría, facilitando el análisis estructurado.
- Implementar catálogos de datos⁹⁵: permiten gestionar y compartir metadatos de manera estandarizada, mejorando la accesibilidad.
- Implementar esquemas de Metadatos Estandarizados (DCAT-AP): herramientas para estructurar y describir conjuntos de datos, asegurando que sean fáciles de encontrar y reutilizables. Consulte los

⁹² Disponible en: <https://www.sen.gov.co/sites/default/files/migracion-files/normatividad/calidad/Marco-Aseguramiento-Calidad-SEN-Colombia-2022.pdf>

⁹³ Disponible en: <https://www.sen.gov.co/sites/default/files/migracion-files/BuenasPracticas/Codigo-Nacional-de-Buenas-Practicas-2022.pdf>

⁹⁴ Consulte: https://www.youtube.com/watch?v=A0pl_MJzc y <https://ayudaleyprotecciondatos.es/bases-de-datos/multidimensionales/>

⁹⁵ Vea el Archivo Nacional de Datos del DANE: <https://microdatos.dane.gov.co/index.php/catalog/central/about>

documentos DANE *Guía de metadatos de registros administrativos*⁹⁶ y *Recomendaciones para la documentación de metadatos de registros administrativos*⁹⁷.

Estándares

- DCAT (Data Catalog Vocabulary): facilita la descripción estandarizada de catálogos de datos, asegurando que los metadatos sean consistentes y reutilizables.
- Dublin Core⁹⁸: estándar para describir recursos y garantizar la interoperabilidad de los metadatos.
- DDI (Data Documentation Initiative)⁹⁸: proporciona un marco estandarizado para la documentación de encuestas y archivos de datos estadísticos.
- Uso de UTF-8 para codificación de caracteres: adoptar UTF-8 como estándar para codificar caracteres, garantizando que los datos sean interpretables y compatibles a nivel internacional.

Protocolos

- Data Structure Definition (DSD) en SDMX⁹⁹: define la estructura de los datos y los metadatos para asegurar su intercambio y reutilización entre instituciones.
- Documentación completa de API: una documentación clara y detallada de las API permite que los usuarios comprendan y utilicen eficazmente las interfaces de programación, mejorando la interoperabilidad técnica.
- Formatos de Estructuración de Datos (JSON Schema, XML Schema): reglas definidas para estructurar datos que aseguren consistencia y claridad.

Cuadro 3. Retos en modelos de datos y de metadatos

- Un error común al modelar conjuntos de datos para su intercambio y difusión es intentar replicar las estructuras de datos internas a partir de los sistemas de bases de datos operativos. El enfoque debería estar en producir conjuntos de datos simples y autónomos que sean fáciles de entender y manipular por los usuarios y las aplicaciones cliente.
- La ausencia de modelos estandarizados puede generar confusión y dificultad para integrar datos provenientes de distintas fuentes.

⁹⁶ Disponible en: https://www.sen.gov.co/sites/default/files/migracion-files/RegistrosAdministrativos/Gu%C3%ADa_Metadatos_RRAA_30072018.pdf

⁹⁷ Disponible en: https://www.sen.gov.co/sites/default/files/migracion-files/guias/Recomendaciones_para_la_documentacion_de_metadatos_de_registros_administrativos.pdf

⁹⁸ Consulte: Guía del DANE para la implementación de los estándares Data Documentation initiative (DDI) y Dublin Core (DC)

⁹⁹ Consulte: 1. Guía del DANE para la implementación del SDMX: https://www.sen.gov.co/sites/default/files/migracion-files/sdmx/Gu%C3%ADa_Implementaci%C3%B3n_SDMX.pdf 2. Ficha institucional de metadatos para indicadores SDMX: https://www.sen.gov.co/sites/default/files/migracion-files/sdmx/Gu%C3%ADa_Metadato_SDMX_MSD.xlsx

Cuadro 3. Retos en modelos de datos y de metadatos

- La falta de metadatos completos y precisos hace que los datos sean difíciles de interpretar y reutilizar en contextos variados.
- Mantener actualizados los metadatos y catalogarlos adecuadamente representa un desafío para las organizaciones con recursos limitados.

Fuente: elaborado con base en *Data Interoperability: A Practitioner's Guide to Joining up data in the development sector*

Dimensión 3. Clasificaciones y vocabularios

Mecanismos

- Uso de repositorios de vocabularios controlados (SKOS, XKOS)¹⁰⁰: facilitan el uso de términos estandarizados para describir datos, asegurando la coherencia semántica entre los actores.
- Implementación de Ontologías y Sistemas de Organización del Conocimiento (KOS): proporcionan marcos estructurados para categorizar y organizar datos, mejorando la integración y el análisis.
- Uso de vocabularios controlados¹⁰¹ y terminología unificada¹⁰²: sistemas para mantener terminología uniforme y precisa en los datos, asegurando claridad y coherencia.
- Implementación de clasificaciones estadísticas: son instrumentos técnicos que sirven para ordenar y categorizar características de la realidad económica, social y ambiental. El DANE pone a disposición de los usuarios internos las clasificaciones adoptadas y adaptadas por la entidad¹⁰³.
- Uso de conceptos estandarizados: hace posible el intercambio de información mediante un lenguaje coherente, que permite la comparación en el ámbito nacional e internacional de los resultados estadísticos. Consulte la *Guía de estandarización de conceptos del SEN*¹⁰⁴

Estándares

- Clasificaciones estadísticas¹⁰⁰¹⁰³: clasificaciones internacionales adoptadas y adaptadas al contexto nacional para garantizar la coherencia y la comparabilidad de los datos entre países y sectores.

¹⁰⁰ Consulte el Sistema de consulta de conceptos estandarizados del SEN: https://conceptos.dane.gov.co/conceptos/area_tematica

¹⁰¹ Los vocabularios controlados son disposiciones estandarizadas y organizadas de palabras y frases que proporcionan una forma coherente de describir los datos. (University of Texas Library, 10 de septiembre de 2024.).

¹⁰² Consulte las Recomendaciones para la implementación de la terminología unificada para la difusión estadística: <https://www.sen.gov.co/sites/default/files/pagina-migraciones-files/2024-12/recomendaciones-para-la-implementacion-de-la-terminologia-unificada-para-la-difusion-de-estad%C3%ADsticas-oficiales-en-el-SEN.pdf>

¹⁰³ Disponibles en: <https://www.dane.gov.co/index.php/sistema-estadistico-nacional-sen/normas-y-estandares/nomenclaturas-y-clasificaciones/clasificaciones>

¹⁰⁴ Disponible en: <https://www.sen.gov.co/sites/default/files/pagina-migraciones-files/2024-02/guia-para-la-estandarizacion-de-conceptos-del-sistema-estadistico-nacional.pdf>

- Modelo genérico del proceso estadístico (GSBPM)¹⁰⁵ y el Modelo Genérico de Información Estadística (GSIM)¹⁰⁶: GSBPM describe el proceso de producción estadística, propone una estructura de procesos y subprocesos para la producción de estadísticas. GSIM proporciona un marco conceptual para describir los objetos de información utilizados en la producción de estadísticas.
- Conceptos estandarizados¹⁰⁴: “definición unificada basada en estándares estadísticos internacionales fundamentales para la armonización estadística, con el fin de proporcionar los elementos necesarios para la construcción y fortalecimiento de los marcos conceptuales.” (DANE, 2023).
- Formatos de Fecha y Hora estandarizados (ISO 8601): aplicación de ISO 8601 para evitar confusiones y asegurar consistencia en el manejo de fechas y tiempos en datos.

Protocolos

- Terminología unificada para la difusión estadística¹⁰²: garantiza que los términos usados en diferentes programas estadísticos sean coherentes y permite la interoperabilidad semántica.
- Proceso de estandarización de conceptos¹⁰⁴: aseguran que las definiciones y las variables estadísticas sean consistentes y comparables entre diferentes organizaciones y así se facilita el intercambio y la comparación de datos entre distintos sistemas y entidades.
- Proceso de estandarización de campos de datos (ejemplo nombres, fechas): directrices que especifican cómo deben ser representados y formateados ciertos tipos de datos.

Cuadro 4. Retos en clasificaciones y vocabularios

- A medida que los sistemas de información se vuelven más grandes, complejos e interconectados, surge una tensión creciente entre la necesidad de utilizar clasificaciones y vocabularios estándar para mejorar la interoperabilidad y la necesidad de diseñar clasificaciones especializadas para grupos de usuarios específicos. Es necesario reflexionar sobre cómo se establece este equilibrio y realizar esfuerzos para conectar las nuevas clasificaciones con las existentes, asegurando continuidad e interoperabilidad a largo plazo.
- La falta de clasificaciones y vocabularios estandarizados dificulta la comparabilidad y el análisis de datos entre organizaciones.
- Los términos y clasificaciones pueden no estar alineados entre países o regiones, afectando los esfuerzos de cooperación internacional.

¹⁰⁵ Consulte: https://www.sen.gov.co/sites/default/files/migracion-files/sen/normatividad/Lineamientos_Proceso_Estad%C3%ADstico_v2.pdf

¹⁰⁶ <https://unstats.un.org/unsd/classifications/expertgroup/egm2015/ac289-22.PDF>

Cuadro 4. Retos en clasificaciones y vocabularios

- Las clasificaciones y los vocabularios controlados no deben considerarse estáticos; deben ser flexibles y actualizarse continuamente para adaptarse a los intereses diversos y cambiantes de los productores de datos, los gestores de datos y los usuarios de datos. La necesidad de adaptar y actualizar clasificaciones para reflejar cambios sociales y económicos puede ser complejo.

Fuente: elaborado con base en *Data Interoperability: A Practitioner's Guide to Joining up data in the development sector*

Dimensión 4. Formatos de datos abiertos e interfaces estandarizadas

Mecanismos

- Uso de API estándar: las API que proporcionan acceso a datos en formatos abiertos como CSV, JSON y XML permiten que los datos sean utilizados por diferentes plataformas y aplicaciones de manera eficiente.
- Uso de Formatos Accionables por Máquinas (MAF): facilitan la extracción automática, la limpieza y el análisis de datos, lo que reduce la intervención manual y mejora la eficiencia en el manejo de datos estadísticos.

Estándares

- DCAT-AP: una extensión de DCAT utilizada para asegurar la interoperabilidad entre portales de datos abiertos y que permite que los datos publicados sean coherentes y fácilmente accesibles.
- JSON-LD: un formato utilizado para representar datos enlazados en JSON, asegurando que los datos sean legibles, tanto por humanos como por máquinas.
- CSV, JSON, SDMX: estos formatos estandarizados aseguran que los datos puedan ser compartidos, procesados y reutilizados sin la necesidad de convertir los formatos, lo que facilita su integración en diferentes sistemas.
- OpenAPI: estándar para definir interfaces de programación que faciliten la interoperabilidad técnica.

Protocolos

- Políticas de datos abiertos¹⁰⁷: establecen condiciones para la publicación de datos en formatos abiertos, garantizando su accesibilidad y reutilización bajo licencias como Creative Commons.

¹⁰⁷ Revisar las Guías de MinTIC: Guía para el uso y el aprovechamiento de Datos Abiertos en Colombia: <https://herramientas.datos.gov.co/sites/default/files/2021-08/Guia%20de%20Datos%20Abiertos%20de%20Colombia.pdf> y la Guía de Estándares de Calidad e Interoperabilidad de Datos Abiertos: https://gobiernodigital.mintic.gov.co/692/articles-179118_recurso_5.pdf

- SEP (Protocolos de Intercambio Estandarizado): protocolos como SDMX, JSON, y CSV permiten una interoperabilidad técnica fluida, asegurando que los datos sean compartidos y reutilizados sin errores ni duplicaciones.
- Uso de Formatos Abiertos (CSV, XML, JSON): reglas que especifican cómo los datos deben ser compartidos en formatos que sean accesibles y reutilizables. La adopción de JSON-LD en API públicas permite que los datos sean fácilmente integrados en aplicaciones de terceros, lo que facilita su análisis y reutilización en múltiples plataformas.

Cuadro 5. Retos en formatos de datos abiertos e interfaces estandarizadas

- Las interfaces de sistemas deben priorizar la interoperabilidad y la flexibilidad sobre la especificidad y la optimización. Debe encontrarse un equilibrio entre las necesidades de grupos de usuarios específicos y la usabilidad general.
- Todas las API de una organización deben gestionarse como un único producto, asegurándose de que todas las API componentes estén estandarizadas, cuenten con documentación y funcionalidades mutuamente consistentes, y apliquen patrones de diseño comunes basados en componentes reutilizables.
- La adopción de formatos de datos propietarios o cerrados limita la accesibilidad y la reutilización de los datos por parte de otros usuarios.
- El diseño de interfaces de usuario poco intuitivas o restrictivas puede desalentar el uso de datos, afectando la eficacia del intercambio de información.
- La falta de API bien documentadas y consistentes dificulta la automatización y el acceso controlado a los datos.

Fuente: elaborado con base en *Data Interoperability: A Practitioner's Guide to Joining up data in the development sector*

Dimensión 5. Datos enlazados¹⁰⁸

Mecanismos

- Uso de Identificadores únicos (URI): facilitan la vinculación de datos entre diferentes fuentes, permitiendo la integración y la consulta de datos a través de diversos sistemas de manera precisa y coherente.
- Implementación de Tecnologías de Datos Enlazados (RDF, SKOS): herramientas que permiten conectar datos de diferentes fuentes mediante relaciones semánticas, mejorando su integración y reutilización.

¹⁰⁸ Un conjunto de principios de diseño para publicar datos estructurados en un formato legible por máquinas que permite enlazarlos con otros datos (Data.europa.eu, 1 de agosto de 2022).

Estándares

- RDF (Resource Description Framework) y OWL (Web Ontology Language): facilitan el uso de datos enlazados mediante la representación de relaciones estructuradas entre diferentes elementos de datos. Esto mejora la interoperabilidad semántica y la integración de datos.
- DCAT y SDMX: facilitan la descripción y el intercambio de datos enlazados y sus metadatos, asegurando que los datos sean comprensibles y reutilizables por diferentes sistemas.

Protocolos

- SPARQL: un protocolo que permite la consulta y la extracción de datos enlazados a través de diferentes fuentes, mejorando la capacidad de las ONE para integrar datos de diversas fuentes de manera efectiva.
- Implementación de URI consistentes: directrices para el uso de identificadores únicos y permanentes que aseguren la conexión estable entre datos de diferentes fuentes. El uso de RDF y SPARQL permite que los datos estadísticos provenientes de censos y encuestas sean integrados y consultados para obtener un análisis más completo, facilitando la comparación de información entre diferentes sistemas.
- Modelos Comunes de Intercambio (CEM): formatos estandarizados para el intercambio de datos entre diferentes sistemas estadísticos lo que aseguran que los datos puedan ser interpretados y comparados de manera eficiente entre distintas agencias.

Cuadro 6. Retos en datos enlazados

- Los casos de uso, objetivos y metas a corto plazo pueden dificultar una visión integral. La implementación de enfoques basados en datos enlazados exige inversión y tiempo, y sus beneficios no siempre son inmediatos. No obstante, a largo plazo, este enfoque tiene el potencial de aumentar significativamente el valor y la usabilidad de los datos.
- Para aprovechar plenamente el potencial de los datos abiertos enlazados, es esencial adoptar enfoques colaborativos en la publicación de conjuntos de datos similares, tanto en ecosistemas de datos como entre distintas comunidades. La coordinación, una estructura bien definida y una gobernanza sólida son elementos clave para lograr este objetivo.
- La falta de identificadores únicos y persistentes para las entidades dificulta la integración de datos de diversas fuentes.
- Las ontologías inconsistentes o incompatibles entre sistemas pueden generar problemas de interpretación y conexión de los datos.
- La dependencia de fuentes de datos desactualizadas limita la utilidad de los datos enlazados y puede resultar en conclusiones erróneas.

Fuente: elaboración propia con base en *Data Interoperability: A Practitioner's Guide to Joining up data in the development sector*

7. Bibliografía

Data to Policy. (n.d.). *Data to Policy Navigator*. <https://www.datatopolicy.org/>

Departamento Administrativo Nacional de Estadística (DANE). (2020). *Lineamientos para el proceso estadístico en el Sistema Estadístico Nacional V2*. [https://www.sen.gov.co/sites/default/files/migracion-files/sen/normatividad/Lineamientos Proceso Estadístico v2.pdf](https://www.sen.gov.co/sites/default/files/migracion-files/sen/normatividad/Lineamientos%20Proceso%20Estadistico%20v2.pdf)

Departamento Administrativo Nacional de Estadístico (DANE). (2020). *Norma técnica de calidad proceso estadístico (NTC PE 1000:2020)*. [https://www.dane.gov.co/files/sen/normatividad/NTC Proceso Estadístico.pdf](https://www.dane.gov.co/files/sen/normatividad/NTC%20Proceso%20Estadistico.pdf)

Departamento Administrativo Nacional de Estadística. (2023). *Política de gobernanza del Comité de Administración de Datos (CAD)*. https://www.sen.gov.co/sites/default/files/pagina-migraciones-files/2024-02/politica_gob_cad_051223.pdf

Departamento Administrativo Nacional de Estadística (DANE). (2024). *Guía para la estandarización de conceptos estadísticos*. <https://www.dane.gov.co/files/sen/lineamientos/Guia-estandarizacion-conceptos.pdf>

Departamento Administrativo Nacional de Estadística (DANE). (2023). *Política de Gobernanza Comité de Administración de Datos*. https://www.sen.gov.co/sites/default/files/pagina-migraciones-files/2024-02/politica_gob_cad_051223.pdf

ELIXIR Research Data Management Kit. (s.f.). What is data reuse?. Recuperado el 26 de diciembre de 2024: <https://rdmkit.elixir-europe.org/reusing> González Morales, L. G., & Orrell, T. (2018). *Data Interoperability: A Practitioner's Guide to Joining-Up Data in the Development Sector*. Global Partnership for Sustainable Development Data. https://www.data4sdgs.org/sites/default/files/services_files/Interoperability%20-%20A%20practitioner%E2%80%99s%20guide%20to%20joining-up%20data%20in%20the%20development%20sector.pdf

Kirchmer, M. (2017). «"Chapter 1: Business Process Management: What Is It and Why Do You Need It?"». *High Performance Through Business Process Management: Strategy Execution in a Digital World* (en inglés). Springer. p. 1–28. ISBN 9783319512594.

Ley 2335 de 2023. *Por la cual se expiden disposiciones sobre las estadísticas oficiales en el país*. Diario Oficial No. 52539 del 03 de octubre de 2023. <https://www.suin-juriscal.gov.co/viewDocument.asp?ruta=Leyes/30050183#:~:text=Es%20la%20aplicaci%C3%B3n%20de%20conjunto,de%20inter%C3%A9s%20nacional%20o%20territorial.>

Ministerio de Tecnologías de la Información y las Comunicaciones. (2019). Marco de interoperabilidad para Gobierno Digital. [https://lenguaje.mintic.gov.co/sites/default/files/archivos/marco de interoperabilidad para gobierno digital.pdf](https://lenguaje.mintic.gov.co/sites/default/files/archivos/marco_de_interoperabilidad_para_gobierno_digital.pdf)

OECD. (2019). *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use Across Societies*. OECD Publishing. <https://doi.org/10.1787/276aaca8-en>

Publications Office of the European Union. (2021). *Data.europa.eu data quality guidelines*. Publications Office. <https://data.europa.eu/doi/10.2830/79367>.

Data.europa.eu (1 de agosto de 2022). *Linking data: what does it mean?* <https://data.europa.eu/doi/10.2830/79367>.

UNECE High-Level Group for the Modernisation of Official Statistics (HLG-MOS). (2023). *Data Governance Framework for Statistical Interoperability (DAFI)*. United Nations Economic Commission for Europe. https://unece.org/sites/default/files/2024-03/HLG2023%20DAFI%20Final_0.pdf

University of Texas Library. (10 de septiembre de 2024.). *Metadata basics*. <https://guides.lib.utexas.edu/metadata-basics/acknowledgements>

Anexo A. Vocabularios, métodos, formatos, marcos, lenguajes, flujos de trabajo y modelos de datos estandarizados¹⁰⁹

CATÁLOGOS DE (META)DATOS

Vocabularios

- [Schema.org](#): es un sitio web de referencia que publica documentación y guías para el uso de marcado de datos estructurados en páginas web (llamado microdatos). Es parte del proyecto de la web semántica.
- [DCAT \(Data Catalog Vocabulary\)](#): es un vocabulario RDF diseñado para facilitar la interoperabilidad entre catálogos de datos publicados en la Web. Se han creado y están en uso los siguientes perfiles de aplicación:
 - [DCAT-AP](#): es el perfil de aplicación de DCAT para portales de datos en Europa. Es una especificación basada en el vocabulario DCAT desarrollado por W3C. Este perfil de aplicación es una especificación para registros de metadatos que satisface necesidades específicas de los portales de datos europeos,

¹⁰⁹ Este anexo es tomado del documento de la UNECE *Data Governance Framework for Statistical Interoperability (DAFI)* disponible en: https://unece.org/sites/default/files/2024-03/HLG2023%20DAFI%20Final_0.pdf

proporcionando interoperabilidad semántica con otras aplicaciones mediante el uso de vocabularios controlados establecidos (e.g., EuroVoc) y mapeos con vocabularios de metadatos existentes (e.g., Dublin Core, SDMX, INSPIRE...).

- [StatDCAT-AP](#): es una extensión del Perfil de Aplicación DCAT para portales de datos en Europa, diseñada para mejorar la interoperabilidad entre descripciones de conjuntos de datos estadísticos dentro del dominio estadístico y entre datos estadísticos y portales de datos abiertos.
- [GeoDCAT-AP](#): es una extensión geoespacial del perfil de aplicación DCAT para portales de datos en Europa.
- [Dublin Core](#): también conocido como Conjunto de Elementos de Metadatos Dublin Core (DCMES), es un conjunto de quince elementos principales de metadatos para describir recursos digitales o físicos. Dublin Core ha sido formalmente estandarizado internacionalmente como ISO 15836 y como IETF RFC 5013.
- [Data Quality Vocabulary \(DQV\)](#): es un modelo de (meta)datos implementado como un vocabulario RDF, que extiende DCAT con propiedades y clases adecuadas para expresar la calidad de los conjuntos de datos.
- [PROV](#): es una especificación que proporciona un vocabulario para intercambiar información de procedencia. Los usuarios pueden marcar sus páginas web o hacer disponible información de procedencia expresada como datos enlazados.
- [DDI-RDF Discovery Vocabulary \(Disco\)](#): define un vocabulario RDF Schema que permite el descubrimiento de datos de investigación y encuestas en la Web. Se basa en los formatos XML de DDI Codebook y DDI Lifecycle.

CONCEPTOS

Vocabularios

- [DDI \(Data Documentation Initiative\)](#): es un estándar internacional gratuito para describir los datos producidos por encuestas y otros métodos observacionales en diferentes ciencias. DDI puede documentar y gestionar diferentes etapas del ciclo de vida de los datos de investigación.
- [SDMX \(Statistical Data and Metadata eXchange\)](#): es una iniciativa internacional que busca estandarizar y modernizar los mecanismos y procesos para el intercambio de datos y metadatos estadísticos entre organizaciones internacionales y sus países miembros.
- [SKOS/XKOS](#): SKOS (Simple Knowledge Organization System) desarrolla especificaciones y estándares para el uso de sistemas de organización del conocimiento, como tesauros y esquemas de clasificación. XKOS aprovecha SKOS para gestionar clasificaciones estadísticas y sistemas de gestión de conceptos.
- [NIEM \(National Information Exchange Model\)](#): es un vocabulario común que permite un intercambio de información eficiente entre diversas organizaciones públicas y privadas.

RECOMENDACIONES DE MECANISMOS, ESTÁNDARES Y PROTOCOLOS PARA LA INTEROPERABILIDAD EN EL INTERCAMBIO, EL USO Y EL REÚSO DE DATOS EN EL SEN / REGULACIÓN ESTADÍSTICA

- [XBRL \(eXtensible Business Reporting Language\)](#): es un estándar internacional abierto para reportes empresariales digitales, administrado por un consorcio global sin fines de lucro. Proporciona un lenguaje en el cual los términos de reporte (principalmente financieros) pueden ser definidos autoritariamente.
- [FIBO \(Financial Industry Business Ontology\)](#): es el estándar para las definiciones de conceptos empresariales en la industria de servicios financieros. Este diccionario permite identificar la terminología definida por el Vocabulario FIBO.
- [GML \(Geography Markup Language\)](#): es una especificación de implementación de OpenGIS diseñada para almacenar y transportar información geográfica. GML es un perfil (codificación) de XML.
- [RCC \(Region Connection Calculus\)](#): es un método utilizado en IA para representar y razonar sobre el espacio. Se basa en dividir el espacio en regiones y representar las relaciones entre estas regiones utilizando un conjunto de reglas de cálculo.
- [OWL-Time](#): es una ontología de conceptos temporales, para describir las propiedades temporales de recursos en el mundo o descritos en páginas web. La ontología proporciona un vocabulario para expresar hechos sobre relaciones entre instantes e intervalos e información sobre duraciones y posiciones temporales.
- [ORG](#): es una ontología central para estructuras organizativas, destinada a apoyar la publicación de datos enlazados de información organizacional en varios dominios.

INTERCAMBIO

Vocabularios

- [DDI \(Data Documentation Initiative\)](#): es un estándar internacional gratuito para describir los datos producidos por encuestas y otros métodos observacionales en diversas ciencias. DDI puede documentar y gestionar diferentes etapas del ciclo de vida de los datos de investigación.
- [SDMX \(Statistical Data and Metadata eXchange\)](#): es una iniciativa internacional que busca estandarizar y modernizar los mecanismos y los procesos para el intercambio de datos y metadatos estadísticos entre organizaciones internacionales y sus países miembros.
- [PMML \(Predictive Model Markup Language\)](#): es un formato de intercambio de modelos predictivos basado en XML. Proporciona una manera de describir e intercambiar modelos predictivos producidos por algoritmos de minería de datos y aprendizaje automático.
- [NIEM \(National Information Exchange Model\)](#): es un vocabulario común que permite un intercambio eficiente de información entre diversas organizaciones privadas y públicas.
- La especificación [OpenAPI](#): es un lenguaje de especificación para APIs HTTP que proporciona un medio estandarizado para definir tu API a terceros.

Métodos

- [REST \(Representational state transfer\)](#): es un estilo arquitectónico de software creado para guiar el diseño y el desarrollo de la arquitectura de la World Wide Web.
- [SOAP \(Simple Object Access Protocol\)](#): es un protocolo de mensajería para el intercambio de información estructurada en la implementación de servicios web en redes informáticas.
- [SPARQL \(SPARQL Protocol and RDF Query Language\)](#): es el lenguaje de consulta estándar y protocolo para Linked Open Data en la web o para almacenes triplestores RDF.
- [SHACL \(Shapes Constraint Language\)](#) is es un lenguaje estándar del W3C para describir gráficos RDF en el marco de Resource Description Framework.
- [GraphQL](#): es un lenguaje de consulta y manipulación de datos de código abierto para API y un motor de ejecución de consultas.
- [ODATA \(Open Data Protocol\)](#): es un protocolo abierto (estándar ISO) que permite la creación y el consumo de API REST interoperables de forma simple y estándar.
- [Protocol Buffers](#): es un formato de datos de código abierto y gratuito, neutral al lenguaje y la plataforma, que permite la serialización de datos estructurados.

Formatos

- [XML \(Extensible Markup Language\)](#): es un lenguaje de marcado para almacenar, transmitir y reconstruir datos arbitrarios. Define un conjunto de reglas para codificar documentos en un formato que es legible tanto por humanos como por máquina.
- [JSON \(JavaScript Object Notation\)](#) es un formato ligero de intercambio de datos, fácil de analizar y generar tanto para humanos como para máquinas. JSON es un formato de texto completamente independiente del lenguaje, pero utiliza convenciones familiares para los programadores de lenguajes de la familia C (C, C++, Java, Python, entre otros). Estas propiedades hacen de JSON un lenguaje ideal para el intercambio de datos.
- [YAML \(YAML Ain't Markup Language\)](#): es un lenguaje de serialización de datos legible por humanos para todos los lenguajes de programación. Se utiliza comúnmente para archivos de configuración y en aplicaciones donde se almacenan o transmiten datos.
- [HTML \(HyperText Markup Language\)](#): es el lenguaje estándar de marcado para documentos diseñados para mostrarse en un navegador web.
- [JSON-LD \(JSON for Linking Data\)](#): es un formato ligero de Datos Enlazados. Se basa en el exitoso formato JSON y proporciona una manera de ayudar a que los datos JSON interoperen a escala web. JSON-LD es un formato de datos ideal para servicios web REST y bases de datos no estructuradas como Apache CouchDB y MongoDB.
- [RDF/XML \(Resource Description Format/eXtensible Markup Language\)](#): es una sintaxis definida por el W3C para expresar un gráfico RDF como un documento XML.

- [Turtle \(Terse RDF Triple Language\)](#): es una sintaxis y formato de archivo para expresar datos en el modelo de datos RDF. La sintaxis Turtle es similar a la de SPARQL, un lenguaje de consulta RDF.
- [CSV \(Comma-Separated Values\)](#): es un archivo de texto delimitado que utiliza comas para separar valores. Cada línea del archivo es un registro de datos que consiste en uno o más campos separados por comas.
- [Text file](#): es un tipo de archivo informático estructurado como una secuencia de líneas de texto electrónico. Un archivo de texto se almacena como datos dentro de un sistema de archivos de computadora.

Marcos

- [GSBPM \(Generic Statistical Business Process Model\)](#): es un modelo que describe la producción de estadísticas de manera general y orientada a procesos. Se utiliza como base común para trabajar con la producción estadística en diferentes aspectos, como calidad, eficiencia y estandarización.
- [GAMSO \(Generic Activity Model for Statistical Organisations\)](#): describe y define las actividades que tienen lugar dentro de una organización típica que produce estadísticas oficiales. Extiende y complementa el GSBPM añadiendo actividades adicionales necesarias para respaldar la producción estadística.
- [GSIM \(Generic Statistical Information Model\)](#): es un marco de referencia de definiciones, atributos y relaciones internacionalmente acordados que describen las piezas de información utilizadas en la producción de estadísticas oficiales.
- [ESS EARF \(European Statistical System Enterprise Architecture Reference Framework\)](#): es un conjunto de documentos que contienen varios artefactos clave que pueden usarse en diferentes etapas de proyectos, así como en la gobernanza general de la realización de la Visión 2020 de Eurostat.
- [EIRA \(European Interoperability Reference Architecture\)](#): es una arquitectura de referencia de cuatro vistas para la prestación de servicios públicos digitales interoperables a través de fronteras y sectores. Define las capacidades necesarias para promover la interoperabilidad como un conjunto de bloques de construcción arquitectónicos
- [CSPA \(Common Statistical Production Architecture\)](#): es una arquitectura de referencia para la industria estadística, desarrollada y revisada por la comunidad estadística internacional. Proporciona un marco, incluidos principios, procesos y guías, para reducir los costos de desarrollo y mantenimiento de procesos y sistemas.
- [CSDA \(Common Statistical Data Architecture\)](#): es una arquitectura de datos desarrollada por la UNECE, centrada en capacidades relacionadas con datos y metadatos, considerados como "recursos de gestión de datos".
- [FAIR \(Find, Access, Interoperate, and Reuse\) Digital Objects](#): proporcionan un marco para desarrollar capacidades interdisciplinarias, gestionar el creciente volumen de datos, construir herramientas que ayuden a aumentar la confianza en los datos, crear mecanismos para operar de manera eficiente en el ámbito científico y promover la interoperabilidad de los datos.

- [International Open Data Charter](#): es un conjunto de principios y mejores prácticas para la publicación de datos abiertos gubernamentales, formalmente adoptados por muchos gobiernos.

NEGOCIO

Lenguajes

- [Java](#): es un lenguaje de programación de alto nivel, basado en clases y orientado a objetos, diseñado para tener la menor cantidad posible de dependencias de implementación.
- [Python](#): es un lenguaje de programación interpretado, orientado a objetos y de alto nivel con semántica dinámica.
- [Scala](#): combina programación orientada a objetos y funcional en un lenguaje de programación conciso y de alto nivel.
- [SAS](#): es un paquete de software estadístico desarrollado por el Instituto SAS para la gestión y el análisis estadístico de datos.
- [R](#): es un lenguaje y entorno extensible gratuito y de código abierto para la computación estadística y gráficos.
- [MDX](#): (MultiDimensional eXpressions) es un lenguaje de consulta utilizado para crear cálculos y agregaciones, mientras que DAX (Data Analysis eXpressions) es un lenguaje desarrollado por Microsoft para realizar cálculos y agregaciones.
- [SDTL \(Structured Data Transformation Language\)](#): es un lenguaje intermedio independiente para representar comandos de transformación de datos. SDTL, desarrollado por DDI Alliance, consiste en esquemas JSON para operaciones comunes.
- [VTL \(Validation and Transformation Language\)](#): es un lenguaje estándar para definir reglas de validación y transformación (conjunto de operadores, su sintaxis y semántica) para cualquier tipo de datos estadísticos.

Flujos de trabajo

- [DDI-CDI \(Cross Domain Integration\)](#): es una especificación diseñada para ayudar a los implementadores a integrar datos a través de límites de dominio e institucionales. DDI-CDI se centra en un enfoque uniforme para describir una gama de formatos de datos necesarios: datos tradicionales amplios/rectangulares, datos largos [eventos], datos multidimensionales y datos NoSQL/clave-valor.
- [BPMN \(Business Process Model and Notation\)](#): es un conjunto estándar de convenciones de diagramación para describir procesos empresariales. Representa visualmente una secuencia detallada de actividades comerciales y flujos de información necesarios para completar un proceso..
- [CMMN \(Case Management Model and Notation\)](#): es una notación gráfica utilizada para capturar métodos de trabajo basados en la gestión de casos que requieren diversas actividades que pueden realizarse en un orden impredecible en respuesta a situaciones cambiantes.

- [DMN \(Decision Model and Notation\)](#): es un lenguaje de modelado y notación para la especificación precisa de decisiones empresariales y reglas de negocio. DMN es fácilmente comprensible para los diferentes tipos de personas involucradas en la gestión de decisiones.
- [ProvONE](#): se define como una extensión del estándar PROV recomendado por el W3C, con el objetivo de capturar la información más relevante relacionada con los procesos computacionales de flujos de trabajo científicos..
- [CWL \(Common Workflow Language\)](#): es un estándar abierto para describir cómo ejecutar herramientas de línea de comandos y conectarlas para crear flujos de trabajo. Las herramientas y los flujos de trabajo descritos utilizando CWL son portátiles en diversas plataformas.

ESTRUCTURAS

Modelos de datos

- [DDI \(Data Documentation Initiative\)](#): es un estándar internacional gratuito para describir los datos producidos por encuestas y otros métodos observacionales en diversas ciencias. El DDI puede documentar y gestionar diferentes etapas del ciclo de vida de los datos de investigación.
- [SDMX \(Statistical Data and Metadata eXchange\)](#): es una iniciativa internacional que busca estandarizar y modernizar los mecanismos y los procesos para el intercambio de datos y metadatos estadísticos entre organizaciones internacionales y sus países miembros.
- [EDM \(Entity Data Model\)](#): es un conjunto de conceptos que describen la estructura de los datos, independientemente de su forma de almacenamiento. EDM se basa en el Modelo de Entidad-Relación descrito por Peter Chen en 1976, pero extiende sus usos tradicionales.
- [RDF DM \(Data Model\)](#): es un modelo estándar para el intercambio de datos en la Web. RDF tiene características que facilitan la combinación de datos y específicamente soporta la evolución de esquemas a lo largo del tiempo.
- [Directed Graph](#): como su nombre lo indica, es un grafo dirigido cuyas flechas tienen etiquetas. Un grafo dirigido (o digrafo) es un grafo compuesto por un conjunto de vértices conectados por aristas dirigidas, llamadas arcos.
- [HDF5 \(Hierarchical Data Format version 5\)](#): es un formato de archivo de código abierto que admite datos grandes, complejos y heterogéneos. HDF5 utiliza una estructura similar a un "directorio de archivos" que permite organizar datos dentro del archivo de muchas maneras estructuradas diferentes, como se haría con archivos en una computadora.

Vocabularios

- [DDI \(Data Documentation Initiative\)](#): es un estándar internacional gratuito para describir los datos producidos por encuestas y otros métodos de observación en diferentes ciencias. DDI puede documentar y gestionar diferentes etapas del ciclo de vida de los datos de investigación.
- [SDMX \(Statistical Data and Metadata eXchange\)](#): es una iniciativa internacional que tiene como objetivo estandarizar y modernizar los mecanismos y procesos para el intercambio de datos y metadatos estadísticos entre organizaciones internacionales y sus países miembros.
- [RDF Data Cube](#): proporciona un medio para publicar datos multidimensionales, como estadísticas, en la web de tal manera que se puedan vincular a conjuntos de datos y conceptos relacionados utilizando el estándar W3C [RDF \(Resource Description Framework\)](#) (Marco de descripción de recursos).
- [CSVW \(CSV en la web\)](#): es un método estándar para publicar y compartir datos almacenados en [Archivos CSV](#).

Formatos

- [XML \(Extensible Markup Language\)](#): es un lenguaje de marcado para almacenar, transmitir y reconstruir datos arbitrarios. Define un conjunto de reglas para codificar documentos en un formato que sea legible tanto por humanos como por máquinas.
- [JSON \(JavaScript Object Notation\)](#): es un formato ligero de intercambio de datos, fácil de analizar y generar tanto para humanos como para máquinas. JSON es un formato de texto completamente independiente del lenguaje, pero utiliza convenciones que son familiares para los programadores de la familia de lenguajes C (C, C++, Java, Python...). Estas propiedades hacen de JSON un lenguaje de intercambio de datos ideal.
- [RDF/XML \(Resource Description Format/eXtensible Markup Language\)](#): es una sintaxis, definida por el W3C, para expresar un gráfico RDF como un documento XML. A RDF/XML a veces se le llama simplemente RDF porque históricamente fue el primer formato de serialización RDF estándar del W3C.
- [Turtle \(Terse RDF Triple Language\)](#): es una sintaxis y un formato de archivo para expresar datos en el modelo de datos del Marco de descripción de recursos (RDF). La sintaxis de Turtle es como la de SPARQL, un lenguaje de consulta RDF.
- [JSON-LD \(JSON for Linking Data\)](#) (JSON para vincular datos): es un formato ligero de datos vinculados. Se basa en el ya exitoso formato JSON y proporciona una manera de ayudar a que los datos JSON interoperen a escala web. JSON-LD es un formato de datos ideal para servicios web REST y bases de datos no estructuradas como Apache CouchDB y MongoDB.
- [Protocol Buffers](#): es un formato de datos gratuito y de código abierto, multiplataforma, neutral en cuanto al lenguaje y la plataforma, que funciona como un mecanismo extensible para serializar datos estructurados.
- Archivo [CSV \(Comma-Separated Values\)](#): es un archivo de texto delimitado que utiliza una coma para separar valores. Cada línea del archivo es un registro de datos. Cada registro consta de uno o más campos, separados por comas.