

ACUERDO MINISTERIAL No. 035-2020

EL MINISTRO DE TELECOMUNICACIONES Y DE LA SOCIEDAD DE LA INFORMACIÓN

CONSIDERANDO:

Que, el numeral 1 del artículo 154 de la Constitución de la República confiere a las ministras y ministros de Estado, además de las atribuciones establecidas en la ley, la rectoría de las políticas públicas del área a su cargo; así como la facultad de expedir los acuerdos y resoluciones administrativas que requiera su gestión;

Que, el numeral 2 del artículo 18 de la Constitución de la República reconoce como derecho de todas las personas: *“2. Acceder libremente a la información generada en entidades públicas, o en las privadas que manejen fondos del Estado o realicen funciones públicas. No existirá reserva de información excepto en los casos expresamente establecidos en la ley. En caso de violación a los derechos humanos, ninguna entidad pública negará la información*

Que, el artículo 226 de la Constitución de la República indica que: *“Las instituciones del Estado, sus organismos, dependencias, las servidoras o servidores públicos y las personas que actúen en virtud de una potestad estatal ejercerán solamente las competencias y facultades que les sean atribuidas en la Constitución y la ley. Tendrán el deber de coordinar acciones para el cumplimiento de sus fines y hacer efectivo el goce y ejercicio de los derechos reconocidos en la Constitución”*;

Que, el artículo 227 ibídem dispone que: *“La administración pública constituye un servicio a la colectividad que se rige por los principios de eficacia, eficiencia, calidad, jerarquía, desconcentración, descentralización, coordinación, participación, planificación, transparencia y evaluación”*;

Que, el 10 de noviembre de 2007 se adoptó la Carta Iberoamericana de Gobierno Electrónico (CLAD) estableciéndose en el literal b) del numeral 2 como uno de sus objetivos *“Incrementar la calidad de los servicios y productos públicos que el Estado tiene que suministrar a los ciudadanos al mejorar la eficiencia, la eficacia y una mayor transparencia de la gestión pública, aprovechando la utilización de las TIC en el Gobierno y en la Administración Pública”*. Asimismo, promueve los principios del gobierno electrónico entre los que está la *transparencia*.

Que, el 29 de octubre de 2016, se adoptó la Carta Iberoamericana Gobierno Abierto CIGA cuyo apartado C señala: *“Datos públicos para el desarrollo incluyente y sostenible. En materia de apertura de datos, los gobiernos deberían diseñar, implementar y desarrollar portales de datos abiertos y elaborar normativas y/o pautas metodológicas para su adecuada categorización, uso y reutilización por parte de la ciudadanía y otros actores del ecosistema del gobierno abierto. Para ello, los países podrán apoyarse en la Carta Internacional de Datos Abiertos”*.

Que, el 14 de abril del 2018, el Ecuador suscribió el Compromiso de Lima en la VIII Cumbre de las Américas, denominado *“Gobernabilidad Democrática frente la Corrupción”*,

compro-metiéndose en el numeral 20 a: “Impulsar el establecimiento de un Programa Interamericano de Datos Abiertos, en el marco de la OEA, con el objetivo de fortalecer las políticas de apertura de información, e incrementar la capacidad de los gobiernos y ciudadanos en la prevención y el combate a la corrupción, teniendo en cuenta los importantes trabajos realizados en *el ámbito interamericano en esta materia y otras iniciativas regionales y mundiales*”.

Que, el artículo 1 de la Ley Orgánica de Transparencia y Acceso a la Información Pública determina: “*Ley tiene por objeto disponer la optimización de trámites administrativos, regular su simplificación y reducir sus costos de gestión, con el fin de facilitar la relación entre las y los administrados y la Administración Pública y entre las entidades que la componen (...)*”y, en su artículo 7, consta la lista de información mínima actualizada que las entidades de gobierno deben difundir desde sus sitios web, siendo este el principal insumo en la actualidad para la disponibilidad libre de información del sector público;

Que, el artículo 3 de la Ley Orgánica del Sistema de Registro de Datos Públicos indica que: “*Los datos públicos registrales deben ser completos, accesibles, en formatos libres, sin licencia alrededor de los mismos, no discriminatorios, veraces, verificables y pertinentes, en relación al ámbito y fines de su inscripción*”;

Que, el artículo 116 del Código Orgánico de la Economía Social de los Conocimientos, la Creatividad y la Innovación señala que: que “*La información y el contenido de las bases de datos producto de las investigaciones financiadas con recursos públicos serán de acceso abierto. Las instituciones o entidades responsables de tales investigaciones deberán poner a disposición dicha información a través de las tecnologías de la información*”;

Que, el artículo 142 del Código Orgánico de la Economía Social de los Conocimientos, la Creatividad y la Innovación establece: “*(...) Los estándares abiertos son formas de manejo y almacenamiento de los datos en los que se conoce su estructura y se permite su modificación y acceso no imponiéndose ninguna restricción para su uso; además que los datos almacenados en formatos de estándares abiertos no requieren de software propietario para ser utilizados (...)*”;

Que, el artículo 90 de la Ley Orgánica de Participación Ciudadana obliga a: “*Las autoridades del Estado, electas o de libre remoción, representantes legales de las empresas públicas o personas jurídicas del sector privado que manejen fondos públicos o desarrollen actividades de interés público, los medios de comunicación social, a través de sus representantes legales, están obligados a rendir cuentas, sin perjuicio de las responsabilidades que tienen las servidoras y los servidores públicos sobre sus actos y omisiones (...)*” ;

Que, de conformidad con lo dispuesto en el artículo 17 del Estatuto del Régimen Jurídico Administrativo de la Función Ejecutiva los Ministros de Estado son competentes para el despacho de todos los asuntos inherentes a sus ministerios sin necesidad de autorización alguna del Presidente de la República, salvo los casos expresamente señalados en leyes especiales;

Que, el Decreto Ejecutivo No. 371, de 19 de abril de 2018, señala en su artículo 1. “*Declarar como política pública del Gobierno Nacional, la adopción de la Agenda 2030*

para el Desarrollo Sostenible, orientada al cumplimiento de sus objetivos y metas en el marco de su alineación a la planificación y desarrollo nacional”. El Objetivo 16 de la Agenda 2030, entre otras metas plantea: “16.5 Reducir considerablemente la corrupción y el soborno en todas sus formas; 16.6 Crear a todos los niveles instituciones eficaces y transparentes que rindan cuentas.; 16.7 Garantizar la adopción en todos los niveles de decisiones inclusivas, participativas y representativas que respondan a las necesidades; (...) 16.10 Garantizar el acceso público a la información y proteger las libertades fundamentales, de conformidad con las leyes nacionales y los acuerdos internacionales”.

Que, mediante Decreto Ejecutivo No. 8, publicado en el Registro Oficial No. 10 de 24 de agosto de 2009, el Presidente de la República del Ecuador creó el Ministerio de Telecomunicaciones y de la Sociedad de la Información;

Que, mediante Decreto Ejecutivo No. 784, de 04 de junio de 2019, el Presidente de la República del Ecuador designó al licenciado Andrés Michelena Ayala como Ministro de Telecomunicaciones y de la Sociedad de la Información;

Que, el artículo 2 del Decreto Ejecutivo No. 981, de 28 de enero de 2020, establece: “*El Ministerio de Telecomunicaciones y de la Sociedad de la Información será la entidad rectora en gobierno electrónico de la Función Ejecutiva. Para la correcta implementación del gobierno electrónico ejercerá las siguientes atribuciones y responsabilidades: “(...) 4. Emitir políticas, directrices, acuerdos, convenios y desarrollar proyectos respecto a datos abiertos de la Función Ejecutiva” , “5. Articular y coordinar con las demás entidades de la Función Ejecutiva, así como con las otras Funciones del Estado y demás actores públicos y privados que directa o indirectamente coadyuvan a la aplicación del presente Decreto; y 6. Ejercer la representación oficial del Estado ante organismos y entidades internacionales en temas de gobierno electrónico”;*

Que, el artículo 13 de la Resolución No. 007-DPE-CGAJ de la Defensoría del Pueblo determina como parámetros para cumplimiento del artículo 7 de la Ley Orgánica de Transparencia y Acceso a la Información Pública: “*(...) los contenidos a publicar en el link de transparencia de los sitios web institucionales en las respectivas matrices homologadas en formato PDF, con los enlaces (hipervínculos) y los documentos para descargar la información que corresponda, sin perjuicio de que dicha información adicionalmente se publique en formato de dato abierto, para aportar a los procesos de transparencia, participación y colaboración ciudadana, innovación y emprendimiento para el desarrollo del país”;*

Que, mediante Acuerdo No. 1184, de 20 de julio de 2015, se emite la Norma Técnica en la Gestión Pública, dentro de lo cual se estipula que “*La disponibilidad de los datos abiertos entre actores internos y externos de la organización, promueve las relaciones con otros grupos de interés relevantes, la generación de círculos de innovación, y favorece la participación de personas, embajadores de innovación y organizaciones de distinta índole, con lo cual se dinamiza la interacción de los tres sectores de un sistema nacional de innovación, es decir, público, privado y académico; entendiendo que el eje de los tres sectores es la ciudadanía en general”;*

Que, mediante Acuerdo Ministerial No. 015-2019, de 18 de julio de 2019, el Ministro de Telecomunicaciones y de la Sociedad de la Información aprobó la Política Ecuador Digital

que tiene como objetivo transformar al país hacia una economía basada en tecnologías digitales, mediante la disminución de la brecha digital, el desarrollo de la Sociedad de la Información y del Conocimiento, el Gobierno Digital, la eficiencia de la administración pública y la adopción digital en los sectores sociales y económicos;

Que, el artículo 4 de la Política Ecuador Digital establece: *“La implementación de la presente política se realizará con base en tres ejes de acción: Conectividad, Eficiencia y Seguridad de la Información, e Innovación y Competitividad. (...) 4.2. Eficiencia y Seguridad de la Información.- Garantizar la participación ciudadana, la democratización de los servicios públicos, la simplificación de trámites, la gestión estatal eficiente, la publicación y utilización de datos abiertos, la gestión de la seguridad de la información, y la protección de datos personales (...);”*

Que, con Acuerdo Ministerial No. 011-2020, publicado en el Registro Oficial No. 190 el 24 de abril de 2020, el Ministerio de Telecomunicaciones y de la Sociedad de la Información emitió la Política de Datos Abiertos, con el cual dispone la implementación de los datos abiertos en la Función Ejecutiva para fortalecer la participación ciudadana, la transparencia gubernamental, mejorar la eficiencia en la gestión pública, promover la investigación, el emprendimiento y la innovación en la sociedad. La Primera Disposición Transitoria señala: *“Para implementación de la presente Política el ente rector de Gobierno Electrónico emitirá la Guía de Datos Abiertos, en el plazo de noventa (90) días contados desde la publicación del presente Acuerdo en el Registro Oficial”*.

Que, el 27 de diciembre de 2019, se suscribió entre el Ministerio de Telecomunicaciones y de la Sociedad de la Información y la Red Ecuatoriana de Datos Abiertos y Metadatos (REDAM), un Acta Compromiso de Plan de Acción de Gobierno Abierto Ecuador, con el objetivo de validar el compromiso denominado *“Elaboración colaborativa de la política de datos abiertos”*, tanto por parte de la máxima autoridad de la Institución Pública responsable de su implementación como por la contra-parte.

Que, con memorando Nro. SENADI-DNDA-2020-0037-M, de 20 de julio de 2020, el Servicio Nacional de Derechos Intelectuales emitió su pronunciamiento respecto al uso de licencias, en el que indicó: *“(...) Para la aplicación de licencias de datos abiertos, se deberá en principio verificar la efectiva existencia de una obra protegida por propiedad intelectual, para en función de la obra y en atención de las circunstancias determinar el tipo de licencia que más convenga aplicar, considerando al efecto el justo equilibrio entre el bienestar colectivo, el reconocimiento al trabajo realizado por el creador y el incentivo a la inversión”*.

Que, mediante sumilla inserta en memorando No. MINTEL-SGERC-2020-0235-M, de 11 de noviembre de 2020, el Viceministro de Tecnologías de la Información y Comunicación aprobó el informe técnico que sustenta la emisión de la Guía de Datos abiertos de aplicación para la Administración Pública Central;

En ejercicio de las atribuciones que le confiere el artículo 154 numeral 1 de la Constitución de la República y artículo 17 del Estatuto del Régimen Jurídico y Administrativo de la Función Ejecutiva;

ACUERDA:

Art. 1.- Emitir la Guía de Datos Abiertos de aplicación en la Administración Pública Central, que consta como Anexo al presente Acuerdo Ministerial.

Art. 2.- De la ejecución, monitoreo y seguimiento de la presente Guía encárguese a la Subsecretaria de Gobierno Electrónico y Registro Civil.


Art. 3.- El presente Acuerdo Ministerial entrará en vigencia a partir de su publicación en el Registro Oficial.

Dado en la ciudad de Quito, Distrito Metropolitano, a 11 de diciembre de 2020.



Firmado electrónicamente por:
**CARLOS ANDRES
MICHELENA AYALA**

Lcdo. Andrés Michelena Ayala
**MINISTRO DE TELECOMUNICACIONES
Y DE LA SOCIEDAD DE LA INFORMACIÓN**

Elaborado:	Ab. Cristina Puga Abogada DALDN	 Firmado electrónicamente por: CRISTINA ALEXANDRA PUGA CARRASCO
Revisado:	Ab. Ricardo Dávalos Director de Asesoría Legal y Desarrollo Normativo	 Firmado electrónicamente por: RICARDO ALEJANDRO DAVALOS GONZALEZ
	Ab. Gabriela Espinoza Coordinadora General Jurídica	MARJORIE GABRIELA ESPINOZA PLUA Firmado digitalmente por MARJORIE GABRIELA ESPINOZA PLUA Fecha: 2020.12.11 12:01:13 -05'00'

ANEXO

GUÍA DE DATOS ABIERTOS

1. ASPECTOS GENERALES

1.1. Objeto de la guía

Proporcionar criterios técnicos y metodológicos para planificar, abrir, publicar y promover la utilización de los datos abiertos gubernamentales.

1.2. Fines de la guía

- Orientar la apertura y publicación de datos de acuerdo con los principios de apertura establecidos en la Política de Datos Abiertos.
- Involucrar aliados estratégicos de los diferentes sectores de la sociedad, con criterios de corresponsabilidad en las diferentes etapas del ciclo de vida de los datos abiertos: conoce, planea, publica, promueve, evalúa y perfecciona.

1.3. Ámbito de aplicación

La presente Guía de Datos Abiertos instrumenta la implementación de la Política de Datos Abiertos para la Administración Pública Central, que según lo señalado en el Artículo 45 del Código Orgánico Administrativo, comprende:

- La Presidencia y Vicepresidencia de la República
- Los ministerios de Estado
- Las entidades adscritas o dependientes
- Las entidades del sector público cuyos órganos de dirección estén integrados, en la mitad o más, por delegados o representantes de organismos, autoridades, funcionarios o servidores de entidades que integran la administración pública central.

En el caso de otras entidades no dependientes de la Administración Pública Central, y otras funciones del Estado, podrán aplicar lo establecido en la presente Guía para la mejora de su gestión en materia de datos abiertos.

1.4. Términos y definiciones

Para efectos de aplicación del presente instrumento se definen los siguientes términos:

API: Siglas de Application Programming Interface, es una manera en que los programas informáticos se hablan entre ellos. Puede ser entendido en términos de cómo un programador envía instrucciones entre programas.

Anonimización: El proceso de adaptación de datos de manera que los individuos no pueden ser identificados a partir de ella.

Conjunto de datos: Es una colección de datos separados que son tratados como una unidad por un computador.

Datatón: Es una competencia en línea de un fin de semana en la que tiene el desafío de trabajar en un caso real desde diferentes áreas de aprendizaje automático, inteligencia artificial y ciencia de datos.¹

Hackatón: Es un evento en el que un gran número de personas se reúnen para escribir o mejorar programas de computadora.²

Datos abiertos: Son datos digitales que son puestos a disposición con las características técnicas y jurídicas necesarias para que puedan ser usados, reutilizados y redistribuidos libremente por cualquier persona, en cualquier momento y en cualquier lugar.

Datos personales: es la información que permite identificar o volver identificable a una persona natural.

Digital: Es un medio por el que crea, presenta, transporta o almacena información mediante la combinación de bits.

Gamificación: Es el empleo de mecánicas de juego en entornos y aplicaciones no lúdicas con el fin de potenciar la concentración, motivar la acción, promover el aprendizaje y resolver problemas.³

Información pública: se considera información pública todo documento en cualquier formato, que se encuentre en poder de las instituciones públicas y de las personas jurídicas a las que se refiere la Ley Orgánica de Transparencia y Acceso a la Información Pública, contenidos, creados u obtenidos por ellas, que se encuentren bajo su responsabilidad o se hayan producido con recursos del Estado.

Metadatos: Se definen comúnmente como "datos acerca de los datos". Describen el contenido, la calidad, el formato y otras características que lleva asociadas un recurso, constituyendo un mecanismo para caracterizar datos y servicios de forma que usuarios y aplicaciones puedan localizarlos y acceder a ellos.

¹ <https://www.datasciencesociety.net>

² <https://dictionary.cambridge.org>

³ <http://www.gartner.com>

Ontología: En ciencias de la comunicación y en inteligencia artificial, la ontología es la representación formal del conocimiento donde los conceptos, las relaciones y las restricciones conceptuales son evidentes mediante formalismos en un determinado dominio del conocimiento.

Reutilizar: es la acción de volver a utilizar algo, bien con la función que desempeñaba anteriormente o con otros fines.

2. CICLO DE APERTURA DE LOS DATOS ABIERTOS

Se consideran pasos básicos para que las entidades abran y publiquen datos abiertos, buscando a través de este ciclo fomentar una cultura de apertura de datos abiertos en las entidades de la Administración Pública Central y en la ciudadanía, con la visión de llegar a la utilización, reutilización de los datos abiertos, y mejorar los procedimientos y contenidos de forma permanente.

Dentro de este ciclo se establecen los siguientes pasos básicos:

1. Conoce
2. Planea
3. Publica
4. Promueve
5. Evalúa y perfecciona.

2.1. CONOCE

Conocer, plantea el fundamento que posibilita la aplicación, partir del conocimiento es básico, esta sección responde preguntas básicas en torno a los datos abiertos, saber qué son, quiénes participan, por qué abrir datos, cuál es su potencialidad para impulsar procesos políticos, sociales y económicos.

Objetivo

Promover procesos y actividades para la comprensión de los principios de los datos abiertos, del marco normativo, política y guía de datos abiertos orientados al fomento de una cultura institucional de datos abiertos, a su promoción y utilización.

2.1.1. ¿Qué son los datos abiertos?

“Datos abiertos son datos digitales que son puestos a disposición con las características técnicas y jurídicas necesarias para que puedan ser usados, reutilizados y redistribuidos libremente por cualquier persona, en cualquier momento y en cualquier lugar”⁴.

⁴ <https://opendatacharter.net/principles-es/>

Esto significa que los datos abiertos pueden ser utilizados por todas las personas y con cualquier objetivo, ya sea para informarse, crear un emprendimiento, generar una investigación científica o periodística, crear nuevos servicios o mejorar servicios existentes, entre otros. Así también los datos abiertos constituyen una herramienta para una adecuada y oportuna transparencia en la asignación, distribución y manejo de los fondos públicos.

2.1.2. ¿Para qué sirven los datos abiertos?

Los datos abiertos tienen múltiples utilidades enfocadas de acuerdo con los usuarios de los mismos. Para la sociedad, los datos abiertos son importantes porque con ellos se puede conocer las problemáticas actuales y cuáles son las acciones que se están llevando a cabo para solucionarlas como por ejemplo: conocer las condiciones ambientales de determinadas zonas geográficas, o la localización de centros educativos y culturales; para el sector privado, los datos abiertos permiten realizar análisis que les ayuden a ver qué pasa en el mercado, dónde pueden encontrar oportunidades de negocios y de nuevos emprendimientos; para la Academia los datos abiertos son la fuente para realizar investigaciones, para el desarrollo de aplicaciones; para el sector público y sociedad civil son un instrumento para promover la transparencia, la rendición de cuentas, la participación ciudadana.

Los datos abiertos tienen grandes potenciales, el valor radica en su utilización y reutilización, en la medida en que existan más infomediarios que analicen, organicen, relacionen unos datos con otros y presenten los resultados de estos procesos como información que sea fácilmente entendible y útil para la ciudadanía.

2.1.3. ¿Quiénes participan en la demanda, apertura y uso de datos abiertos?

Todos los actores de la sociedad juegan un papel importante en el ecosistema de datos abiertos. Hay distintas acciones que se pueden tomar para promover la apertura y la utilización de estos datos, tomando ciertos roles, ya sea como productores, consumidores o infomediarios.

Productores: son aquellos actores que generan o producen datos; son grandes productores de datos las entidades públicas, la academia, la empresa privada y la sociedad civil.

Consumidores: en general son todos los actores que utilizan los datos abiertos. La Academia y la comunidad científica son demandantes de datos abiertos para realizar investigaciones, generar información, conocimiento y obtener nuevos datos. También son consumidores actores del sector privado, la sociedad civil, las entidades públicas, la ciudadanía.

Infomediarios: son personas o entidades que recopilan grandes cantidades de datos provenientes de diversas fuentes, los analizan y organizan de manera relevante para presentar de manera más comprensible o dar valor agregado, a usuarios que requieran esos datos.

2.1.4. ¿Quién está a cargo en el país del proceso de apertura de datos?

El MINTEL, en calidad de ente rector de Gobierno Electrónico, es responsable de coordinar con las demás entidades de la Administración Pública Central, con organizaciones de la Sociedad Civil y realiza el rol de Administrador de datos abiertos.

La Secretaría Técnica Planifica Ecuador gestiona la plataforma gubernamental de datos abiertos, facilita las herramientas necesarias para que las entidades de la Administración Pública Central publiquen los conjuntos de datos abiertos, considerando al portal www.datosabiertos.gob.ec, como punto de acceso a los datos abiertos gubernamentales.

Se considera la vinculación con los sectores público, privado, academia y sociedad civil, en todo el ciclo de vida los datos abiertos, y de manera permanente; considerando su involucramiento en los requerimientos como en la colaboración para la consecución de los fines que se busca en esta guía.

2.1.5. ¿Cómo organizarse dentro de la institución?

De acuerdo con lo señalado en la política de Datos Abiertos, las entidades identificarán al Responsable de datos abiertos y al Gestor institucional de datos abiertos que es quien preside el Comité de transparencia; su conformación respetará lo estipulado en la LOTAIP, así como en la Resolución 007 de la Defensoría del Pueblo.

El Gestor institucional de datos abiertos es quien organizará internamente un Equipo de Datos Abiertos que se encargará de cumplir el ciclo de los datos abiertos, según lo señalado en esta Guía.

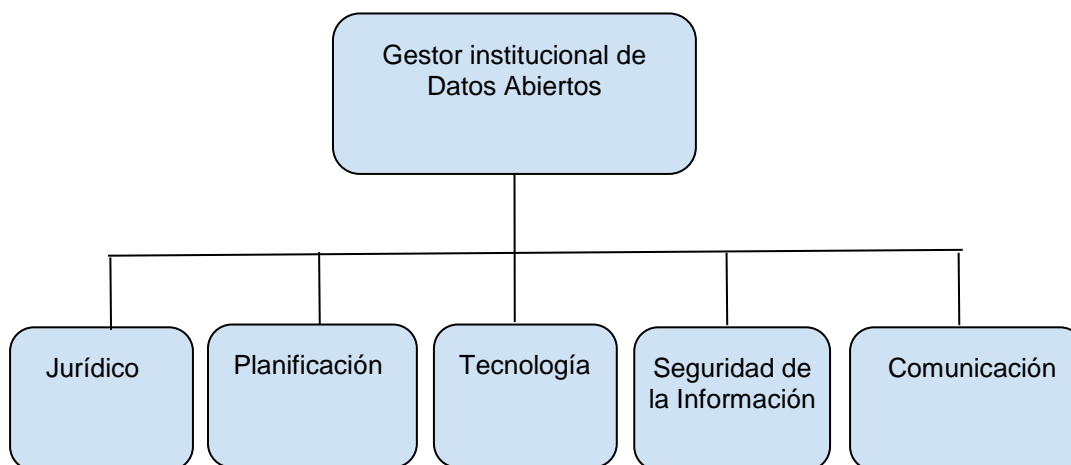


IMAGEN 1. EJEMPLO DE CONFORMACIÓN DEL EQUIPO DE DATOS ABIERTOS

El Equipo de Datos Abiertos incluirá al Comité de Transparencia que está integrado por los titulares de las Unidades Poseedoras de la Información; o como mínimo por delegados de las áreas clave de la organización: Jurídico, Planificación, Comunicación, Tecnología, Oficial de Seguridad de la Información, o quienes hagan sus veces, dependiendo de la realidad de cada institución.

El Equipo de Datos Abiertos debe asegurar el cumplimiento de las acciones planteadas en todo el ciclo de los datos abiertos que promueve esta Guía: conoce, planea, publica, promueve, evalúa y perfecciona.

El Equipo de Datos Abiertos deberá reunirse como mínimo tres veces en el año manteniendo actas de reuniones, que debe tener un formato libre, cuyo requisito mínimo es establecer las circunstancias de tiempo, modo, lugar, registro de asistencia, temas tratados (orden del día), los acuerdos alcanzados, y las tareas que se llevarán a cabo junto con los responsables de su ejecución y la fecha estimada de finalización.

2.2. PLANEA

Algunos de los inconvenientes que se pueden presentar al realizar una planificación de apertura de datos son, falta de estándares, diferentes niveles de accesos, datos que son resultados de aplicar procesos, datos incompletos; o como elementos asociados a su gestión como falta de colaboración entre las entidades, escasa participación e interacción entre actores del ecosistema de datos abiertos, estrategias y políticas sectoriales definidas, entre otras.

Objetivo

Planear consiste en establecer metas en cuanto a los datos que se van a abrir y publicar; comprende este paso el establecimiento de lo que se espera lograr dentro de un período determinado (un año como referencia), enfocándose en la transparencia y rendición de cuentas, la participación y colaboración ciudadana, la protección de los datos personales.

Son objetivos específicos:

- Realizar un proceso de planificación para establecer el Portafolio institucional de datos abiertos en el último trimestre del año, previo a la publicación.
- Establecer un plan de actividades, en el último trimestre de cada año (a excepción del año 2020), para la promoción y utilización de los datos abiertos institucionales.

2.2.1. Construcción del Portafolio institucional de datos abiertos

El instrumento base de la planificación de la apertura de datos es el Portafolio institucional de datos abiertos, para cuya construcción incluirá los datos que les corresponden a los sujetos o representantes del Estado obligados a la transparencia, a la demanda ciudadana, a la oferta de datos estratégicos y sectoriales alineados con los planes de desarrollo y los planes estratégicos, al cumplimiento de los Objetivos de Desarrollo Sostenible (ODS), a las agendas e índices internacionales de datos abiertos, a la visión de su uso y reutilización de cualquier persona que así lo requiera.

El Portafolio institucional de datos abiertos tiene la finalidad de:

- Conformar un listado público de los datos generados o custodiados por las entidades que se van a abrir y publicar, los mismos que deben cumplir los principios de los datos abiertos.

- Indicar a las partes interesadas en un conjunto de datos publicado en el portal de datos abiertos de Ecuador, las fechas propuestas para la publicación de los datos que las entidades han determinado factibles de abrir y publicar.
- Llevar un control de la información y manejar de forma eficiente el flujo de información.
- Proporcionar enlaces de interés que faciliten el paso a otras plataformas con el detalle de la información requerida.
- Delimitar la estructura de la presentación y la publicación de los diferentes tipos de datos abiertos
- Los datos cumplen con el principio de oportunidad y utilidad para ciudadanos, academia y organizaciones sociales
- Contar con un catálogo de datos abiertos que permita contribuir a la usabilidad y reutilización de la información que se genera, produce o custodia en las entidades del Estado y las privadas que administran fondos públicos
- Ofrecer información clave a los usuarios para entender las características de los datos a publicarse.

El Equipo de Datos Abiertos trabajará en los Portafolios institucionales de datos abiertos, para lo cual debe realizar actividades de recopilación de la demanda de datos abiertos, selección y priorización de la demanda, y formalización del Portafolio institucional de datos abiertos.

Los Portafolios institucionales de datos abiertos, deben ser publicados periódicamente según se vayan generando, facilitando el acceso desde el portal gubernamental de datos abiertos.

El Equipo de Datos Abiertos debe asegurarse que las propuestas de datos a publicar cumplen con los requerimientos técnicos y jurídicos que promueve la Guía de datos abiertos.

Dentro de este proceso se deberá realizar un análisis entre los requerimientos de la ciudadanía y los conjuntos de datos disponibles para priorizar los datos que se van a abrir. La estructura del Portafolio institucional de datos abiertos se puede observar en los Anexos y debe incluir:

- Nombre de la entidad publicadora
- Código del documento
- Fecha de elaboración
- Período planificado
- Responsable
- Correo del responsable
- Código del conjunto de datos
- Nombre del conjunto de datos a publicar

- Descripción corta del conjunto de datos
- Entidades que generan la información de origen.
- Fecha prevista de publicación
- Formatos de publicación
- Licencia
- Solicitante / potenciales usuarios
- Categorías

Los Portafolios institucionales de datos abiertos deberán ser revisados al menos semestralmente, y ser actualizados de manera constante de acuerdo con la priorización, peticiones, uso y demanda de datos abiertos. Para su publicación se crearán versiones del mismo. Para nombrar al archivo se debe seguir la siguiente estructura:

SIGLASENTIDAD_PDA_Año_versión

Ejemplo: MINTEL_PDA_2021_v1

2.2.2. Recopilación de la demanda de datos abiertos

El Equipo de Datos Abiertos trabajará en los Portafolios institucionales de datos abiertos, para lo cual debe realizar actividades de identificar, analizar y priorizar datos que se encuentran en las unidades poseedoras de Información, y que responden a tres aspectos básicos:

2.2.2.1. Información pública que es solicitada frecuentemente por los ciudadanos

El Equipo de Datos Abiertos debe velar porque se mantenga actualizado el registro de la Información pública que es solicitada por los ciudadanos, de manera que pueda ser analizada y determinados cuáles datos cumplen las características y principios de los datos abiertos, y son factibles de incluir en el Portafolio institucional de datos abiertos.

Como parte fundamental para captar la demanda para la apertura de datos, las entidades deberán mantener habilitados, de manera permanente, canales de participación ciudadana; a través de la comunicación directa con los usuarios se recibirá propuestas y observaciones que deberán ser contestados de manera oportuna y tener un seguimiento (trazabilidad) ya sea que se incluyan o no, en el Portafolio institucional de datos abiertos.

Los canales de comunicación con la ciudadanía deberán contener ítems específicos para que la solicitud de información sea concisa y clara, de tal manera que se impulse una cultura de datos. Se sugiere las siguientes metodologías para recolectar información:

- Formulario de enlace en la sección de transparencia de la web de la institución.
- Sección o subsección específica para colocar solicitudes de conjuntos de datos.
- Registros administrativos.
- Procesos participativos que pueden incluir, entre otros: Rendición de cuentas participativas, mesas de diálogo para tratar la temática de datos abiertos y plantear nuevos criterios y elementos de interés de la ciudadanía.

- Observatorios regionales y globales, así como convenios con otras entidades para identificar metodologías específicas de recopilación de información para contrastar con estándares internacionales.
- Metodologías innovadoras y ejercicios virtuales que identifique la entidad como, por ejemplo, el uso de redes sociales, chatbots u otros elementos interactivos con los usuarios.

2.2.2.2. Datos que la institución considere oportuno publicar como datos abiertos

En los Portafolios institucionales de datos también se debe incluir los datos que la entidad considere contribuyen a un propósito institucional o pueden ser de beneficio para la sociedad, son criterios para considerar:

- Datos que, cumpliendo los principios de los datos abiertos, puedan resolver alguna problemática de la entidad.
- Proyectos de impacto país.
- Reportes frecuentes de información estratégica o complementaria.
- Potenciales soluciones a problemáticas públicas y/o ciudadanas.
- Peticiones canalizadas a través de grupos focales como infomediarios, periodistas e investigadores.
- Estándares internacionales en materia de transparencia, acceso a la información y lucha contra la corrupción.
- Investigaciones realizadas por la academia, los sujetos obligados a la rendición de cuentas.
- Herramientas para analizar datos: Inteligencia de Negocios (BI), Big Data u otras (Ej: Google Analytics).
- Sesiones con profesionales expertos por sectores para liberar datos.

2.2.3. La que disponga el ente rector de la política de datos abiertos

Como resultado del establecimiento anual de la demanda de datos abiertos, se incluirá también en el Portafolio institucional de datos abiertos los datos que sean el resultado del proceso realizado por el ente rector de la política de datos abiertos para establecer la demanda de datos abiertos, y que se han comunicado a la entidad, en concordancia con sus atribuciones.

2.2.4. Protección de datos personales

Los datos abiertos constituyen un recurso importante para aumentar el acceso de las personas a la información de interés generada por entidades del sector público, social y privado, y en consecuencia promover el desarrollo, la innovación, el crecimiento económico, entre otros aspectos de interés relevantes para la sociedad en general; sin embargo, no se puede dejar de lado, la preocupación de liberar datos abiertos que pueden afectar derechos como el de la protección de datos personales.

La Dirección Nacional de Registro de Datos Públicos (DINARDAP) hace énfasis en la Constitución de la República del Ecuador, que en su artículo 66, numeral 19 reconoce y garantiza el derecho a la protección de datos personales, en virtud de ello, debe

considerarse que uno de los principios que lo rigen es el de la confidencialidad, es por eso que, el acceso a los mismos solo puede darse por aquellos que están debidamente legitimados para hacerlo, es decir por una persona natural o jurídica que cuente con consentimiento del titular, mandato de ley u orden judicial.

Existen formas en que los datos personales pueden convertirse en datos abiertos, como por ejemplo ser sometidos a procesos de anonimización. Acogiendo las experiencias internacionales, las recomendaciones a seguir para garantizar el derecho a la protección de datos personales en relación a la materia de datos abiertos son:

1. Analizar los casos de uso y el valor de los datos personales.
2. Identificar los actores involucrados en el proceso.
3. Analizar e identificar los posibles riesgos ante la apertura de datos personales.
4. Es recomendable contar con el consentimiento de los titulares previo a la apertura de datos personales.
5. Reunir a todos los actores interesados en la construcción de guías, políticas y paquetes.
6. Contar y cumplir con la normativa en protección de datos personales.
7. Observar la garantía de derechos como la honra, la intimidad, la protección de datos personales, entre otros.
8. Generar confianza pública.
9. Únicamente se podrán liberar aquellos datos personales conforme a la normativa especializada en datos personales.
10. Anonimizar datos personales y monitorear permanente el anonimato de los datos.
11. Monitorear permanentemente la utilidad de los datos publicados y en caso de carecerla se sugiere eliminarlos.
12. Hacer seguimiento constante del contexto en que los datos abiertos están siendo utilizados.
13. Realizar un análisis de riesgo y de impacto que permita generar un plan de crisis a posibles casos en que se vulnere el derecho a la protección de datos personales.

Esta Guía promueve el cumplimiento de las disposiciones legales en lo relativo a la protección de datos personales. En caso de que la institución tenga solicitudes de información que incluyan datos personales, pero que a la vez se haya determinado, a través del Equipo de Datos Abiertos, que tiene gran potencial de que sean utilizados al ser publicados como datos abiertos, o que hagan eficiente la entrega de respuestas a pedidos ciudadanos que son reiterativos, o que hayan sido resultado del establecimiento de la demanda de datos abiertos realizada por el ente rector de la Política de Datos Abiertos, se debe proteger los datos personales.

2.2.4.1. Consideraciones para la anonimización

La anonimización es un proceso de disociación a través del cual los datos se desvinculan de la persona a la que pertenecen, de forma que dejan de ser “datos personales”. La anonimización consiste en un proceso que permite identificar y ocultar la información sensible, permitiendo su divulgación sin que ello implique vulnerar los derechos a la protección de datos de personas y organizaciones que se puedan referenciar en los mismos. Incluimos aspectos relevantes de la anonimización de datos producida por la CEPAL⁵.

La anonimización de datos efectuada de manera adecuada, preserva la utilidad de los datos. El análisis previo al proceso de disociación es esencial así también el personal implicado en el proceso de anonimización debe cumplir todos los requisitos de formación e información relativos al cumplimiento de la normativa de protección de datos personales.

Es aconsejable definir un protocolo de actuación. Algunas de las tareas o actividades que deben realizarse durante la fase de anonimización son:

- Determinar los recursos y equipo técnico necesarios para proceder a la anonimización de los datos.
- Determinar la técnica de anonimización que sea más apropiada en función de las variables que hubieran sido identificadas.
- Validar la técnica de anonimización por expertos.
- Aplicar la técnica seleccionada y ejecutar el proceso de anonimización.
- Incluir un sello de tiempo con el fin de garantizar la fecha y hora en la que la anonimización ha sido realizada, o la firma electrónica permiten garantizar la identidad electrónica de quien ha realizado la anonimización.
- Validar y aprobar los archivos anonimizados por el equipo técnico.
- Incluir en la tabla de metadatos la característica de la anonimización realizada, para la publicación del conjunto de datos.

Se debe tener en cuenta las garantías jurídicas necesarias para preservar los derechos de los titulares de los datos. Algunos de los aspectos a tomarse en cuenta son los acuerdos de confidencialidad que impliquen a los actores que participen en el proceso de anonimización.

Hay que tener en cuenta que la anonimización siempre generará, independientemente de las buenas prácticas empleadas, cierto grado de distorsión entre la información anonimizada y la información no anonimizada. Esta distorsión es conocida como “diferencial de privacidad”. Así también se debe considerar que el rápido avance tecnológico podría dejar obsoletas las técnicas de anonimización elegidas en un corto espacio temporal con la

⁵ <https://biblioguias.cepal.org/c.php?g=495473&p=4961125>

aparición de técnicas más seguras u otras que permitan vulnerarlas. La selección de técnicas de anonimización debería someterse también a análisis de riesgos que ayudarán al establecimiento de pautas que permitan optar o decidir por las técnicas más apropiadas en cada momento concreto.

2.2.4.2. Técnicas de anonimización

Existe una variedad de técnicas, así como de clasificaciones, como referencia tenemos algunos de los más conocidos y aplicados.

Algoritmos de Hash

Un algoritmo de hash es un mecanismo que, aplicado a un dato concreto, genera una clave que puede utilizarse para reemplazar el dato real. El algoritmo de Hash permite que, partiendo de un mismo dato podamos generar siempre la misma huella digital, pero partiendo de una determinada huella digital nunca podremos obtener el dato original, garantizando la confidencialidad al tratarse de una operación matemática de un solo sentido. Las claves resultantes de la aplicación de un algoritmo de hash son comúnmente conocidas como “huella digital”.

Algoritmos de cifrado

Un algoritmo de cifrado homomórfico permite realizar operaciones con datos cifrados de tal manera que el resultado de las operaciones es el mismo que si las operaciones se hubieran realizado con los datos sin cifrar. Los resultados de las operaciones con datos cifrados dan por resultado valores igualmente cifrados que pueden ser descifrados posteriormente por el usuario que disponga de la clave para descifrar. El esquema de cifrado homomórfico abre la posibilidad del tratamiento de datos personales anonimizados garantizando la privacidad del tratamiento y que los resultados de los tratamientos van a ser accesibles únicamente al poseedor de la clave de descifrado.

Perturbación de datos

Consiste en la variación y supresión sistemática de datos que evita que las cifras resultantes faciliten información sobre casos específicos. Puede aplicarse como:

- Microagregación: técnica utilizada para anonimizar datos numéricos consistente en la sustitución de valores numéricos concretos por el valor medio calculado para un determinado grupo de datos mediante la agrupación, segregación, supresión o sustitución de registros independientes.
- Intercambio aleatorio de datos: introduce una distorsión aleatoria en un conjunto de microdatos manteniendo el detalle y estructura de la información original.
- Datos sintéticos: como
 - Distorsión de datos: se generan datos aleatorios que mantienen los resultados del conjunto de datos originales; o

- Distorsión con microdatos híbridos: combinación de datos originales con datos sintéticos
- Permutación de registros: intercambio de valores de datos con valor clave que garantiza valores promedios y distribuciones estadísticas.
- Permutación temporal: movimiento aleatorio de rangos temporales que no genera distorsión sobre los resultados medios finales.
- Redondeo: sustitución de variables por valores redondeados de forma aleatoria.
- Reajuste de pesos: cuando se trabaja con muestras de datos conocidas, supone la distorsión de los valores de las muestras originales para evitar la reidentificación.
- Ruido aleatorio: inyecta ruido manteniendo la estructura de datos originales

Reducción de datos

Se reduce el número de datos originales sin alterar los mismos, disminuyen el nivel de detalle de los datos originales evitando la presencia de datos únicos o atípicos sin relevancia para el resultado final. Pueden aplicarse como:

- Eliminación de variables: eliminación de datos especialmente sensibles que pueden ser identificadores directos.
- Reducción de registros: cuando tras aplicar otra medida los sujetos sigan siendo identificables.
- Recodificación global: determinadas categorías de datos se agrupan en una nueva categoría reduciendo las posibilidades de reidentificación.
- Codificación superior o inferior: para casos en los que valores superiores o inferiores de un rango sean identificables, consiste en ampliar o reducir el rango mayor o menor.

2.2.5. Selección y priorización de la demanda de datos abiertos

Una vez establecida la demanda de datos abiertos, se debe aplicar criterios para la selección y priorización de lo requerido en función de los siguientes aspectos:

2.2.5.1. Factibilidad técnica y operativa

Los aspectos que se deben identificar en los datos a publicar son:

- Identificar si los datos requeridos no han sido ya publicados por otras entidades u otras áreas ya que, cuando hay similitud en la naturaleza de sus actividades, se pueden generar o custodiar datos iguales.

- Identificar la Unidad Poseedora de Información responsable de la generación, recopilación, custodia y actualización de los datos.
- Origen de la información (manual y/o automática).
- Periodicidad de la actualización de esta información.
- Si los datos están en diferentes repositorios, identificar el origen de los datos.
- Existencia de metadatos, descripción y contexto claro de los datos.
- Verificación de no exponer datos personales.

2.2.5.2. Pertinencia

- Permiten el cumplimiento a las metas institucionales de transparencia, participación, control social y colaboración.
- Contribuyen a un propósito de política pública específico, forman parte de mejores prácticas a nivel nacional e internacional en la materia.
- Están alineados con los objetivos estratégicos trazados en el Plan Nacional de Desarrollo y los planes institucionales.

2.2.5.3. Responden a la participación ciudadana

- Han sido solicitados por los ciudadanos en ejercicios virtuales o presenciales de consultas públicas.
- Han sido caracterizados como prioritarios por actores del ecosistema de datos abiertos

2.2.5.4. Aportan a estrategias e índices a escala internacional

- Los datos contribuyen al cumplimiento de derechos constitucionales e instrumentos internacionales.
- Están alineados a los Objetivos de Desarrollo Sostenible (ODS) y Agenda 2030.
- Datos requeridos para monitorización y observatorios que aportan con indicadores a fin de facilitar el análisis de la situación actual del país y contrastar con los de la región y el mundo.
- Aportan al cumplimiento de los Índices internacionales de datos abiertos, como Open Data Barometer, Open Data Index o Our Data Index.

2.2.5.5. Aportan a la eficiencia

- Los datos requeridos tienen alta concurrencia de usuarios y son de permanente actualización, esto con el fin de que se planee la construcción de procesos automáticos de extracción, transformación y carga de datos (ETL) desde los sistemas de información de la entidad al Portal gubernamental de datos abiertos.
- Al disponerlos como datos abiertos, se optimiza el costo (recursos, tecnología, tiempo) en preparar y entregar los datos.

2.2.5.6. Generan Impacto y articulación en el ecosistema de datos

- Articulan el sector público con la sociedad civil, el sector privado, la academia.
- Los datos pueden ser reutilizados para generar múltiples beneficios desde perspectivas económicas y sociales.
- Facilitan análisis que requieren datos históricos que sean relevantes y factibles de publicar.

2.3. PUBLICA

Es un proceso que va a permitir aplicar características técnicas de formato, estandarización y criterios de calidad para que los conjuntos de datos se encuentren disponibles abiertamente de modo que puedan ser utilizados por computadoras y personas.

Objetivo

Realizar las actividades necesarias para la publicación de los conjuntos de datos en formatos abiertos a través de las siguientes acciones, como parámetros mínimos:

- Estandarizar los formatos de publicación.
- Estructurar los metadatos.
- Estructurar los diccionarios de datos.
- Determinar las licencias de publicación.
- Verificar el cumplimiento de criterios como parte de evaluación de la calidad de los conjuntos de datos, previa a la publicación.
- Publicar los conjuntos de datos.

Para la publicación, esta sección se encuentra organizada en: fase preparatoria, niveles de apertura, fase de construcción y fase de publicación.

2.3.1. Fase preparatoria

2.3.1.1. Sistemas de origen

Los datos se recopilan de múltiples fuentes y se almacenan en distintas bases de datos, a partir de esos datos almacenados, se genera información útil para los usuarios. La información no es simplemente cadenas de números, listas de direcciones o resultados de pruebas almacenados en una computadora, la información es el producto de los procesos de negocio y continuamente se usa y reutiliza.

Una vez definido el conjunto de datos que se requiere publicar, se deberá analizar los sistemas de origen de los datos requeridos, por cuanto es importante tener el registro de la procedencia de la información que se va a publicar como datos abiertos, ya que muchas veces se encuentra dispersa, almacenada en fuentes (silos) diversas, en diferentes formatos, que no necesariamente están relacionados y al momento de publicarlos en formato abiertos,

ocasiona que sean difíciles de entender por no contar con información contextual adicional. En este aspecto la gestión se basa en utilizar una herramienta para unificar los datos en una sola fuente incorporando normas y reglas, como eliminación de registros duplicados, completar campos vacíos, entre otros, que permitan preparar el conjunto de datos previo a su publicación.

La herramienta que permite implementar el conjunto de datos que tiene diferentes fuentes es la Gestión de Datos Maestros o MDM (de sus siglas en inglés Master Data Management)⁶, que se basa en la creación del Archivo Maestro como estrategia de creación y apropiación de los datos, considerando sus diferentes orígenes⁷.

La estructura del Archivo maestro deberá contemplar las características planteadas en la norma ISO 8000, que son: origen de los datos almacenados, identificación de estado, así como las relaciones o formas de conectarse con otros datos, por cuanto se trata de la recopilación de datos de diferentes orígenes.

El Equipo de Datos Abiertos debe tener identificado este Archivo maestro para que las siguientes publicaciones se tomen los mismos orígenes de datos y tengan coherencia con las actualizaciones que se realicen.

2.3.1.2. Aspectos de interoperabilidad

La interoperabilidad se basa en la estandarización, comprende un conjunto de criterios que deben ser aplicados con el fin de facilitar el intercambio seguro y eficiente de información.

La interoperabilidad en el **dominio semántico** permite garantizar que, en el momento de intercambiar datos, el significado de los datos sea el mismo para todos los actores involucrados.

La interoperabilidad en el **dominio técnico** hace referencia a las condiciones que se deben cumplir para conectar los sistemas de información con el propósito de intercambiar datos.

2.3.1.3. Estándares para los datos abiertos

Se priorizará el uso de estándares internacionales para los datos, tomando como referencia las publicaciones: Datos enlazados de información geográfica del Ecuador⁸ y Datos ecuatorianos geoespaciales enlazados⁹, como se detalla a continuación:

Dato	Estándar a aplicar	Detalle
------	--------------------	---------

⁶ Norma ISO 8000

⁷ https://www.researchgate.net/publication/266476656_Chapter_1_-_Master_Data_and_Master_Data_Management

⁸ Datos enlazados de información geográfica del Ecuador (julio-diciembre 2018), <https://pdfs.semanticscholar.org/36d2/5f8f743aabf2eb80719bb9adc0787418c85a.pdf>

⁹ Ecuadorian Geospatial Linked Data, https://www.researchgate.net/publication/260983879_Ecuadorian_Geospatial_Linked_Data

Dato	Estándar a aplicar	Detalle
Unidades de medidas	Ontologías de cantidades, Unidades, Dimensiones y Tipos de Datos QUDT http://www.qudt.org/qudt/owl/1.0.0/unit/Instances.html	Unidades de medida
Moneda	ISO 4217-2015 https://www.iso.org/iso-4217-currency-codes.html	La moneda de cada monto debe especificarse usando el código de moneda de 3 letras mayúsculas del ISO 4217-2015
Tipo de archivos	IANA media types https://www.iana.org/assignments/media-types/media-types.xhtml	Considerar los estándares para publicar datos abiertos
Idioma	ISO 639-1 https://docs.oracle.com/cd/E13214_01/wli/docs92/xref/xqisocodes.html	Se usa el código de 2 dígitos estandarizado en la ISO 639-1
Países	ISO 3166 https://docs.oracle.com/cd/E13214_01/wli/docs92/xref/xqisocodes.html	Se usa el código de 2 dígitos estandarizado en la ISO 3166
Fecha corta	ISO 8601 https://www.w3.org/TR/NOTE-datetime	Se usa el estándar ISODATE con año, mes, y día YYYY-MM-DD
Fecha larga	ISO 8601 https://www.w3.org/TR/NOTE-datetime	Se usa el estándar ISODATE con año, mes, día, horas, minutos, segundos y zona horaria. YYYY-MM-DDThh:mm:ssTZD
Codificación	ISO 10646 https://www.iso.org/standard/73773.html	El formato de codificación de caracteres es el UTF-8. Define todos los caracteres necesarios en la mayoría de los idiomas.
Catalogar los datos	DCAT-2 https://www.w3.org/TR/vocab-dcat-2	Para definir campos de los metadatos

TABLA 1. ESTÁNDARES INTERNACIONALES PARA LOS DATOS

Para los casos que no están detallados, por ser aplicados en el ámbito ecuatoriano, se considerarán los siguientes estándares locales:

Dato	Estándar a aplicar	Detalle
Provincia, cantón, parroquia	INEN http://apps.normalizacion.gob.ec/descarga/ https://www.ecuadorencifras.gob.ec/documentos/web-inec/Poblacion_y_Demografia/CV_aplicativos/modulo_cpv/dpa_.pdf	Se usan los códigos normalizados por el INEN

TABLA 2. ESTÁNDARES NACIONALES PARA LOS DATOS

2.3.1.4. Vocabulario estandarizado para nombres de archivos

- Formato para la denominación del nombre del conjunto de datos:

SIGLAS DE LA ENTIDAD_NombreConjuntoDatos_AñoMes

Ejemplos:

MINTEL_RemuneracionPuesto_2020 Mayo

MERNNR_ServicioOtorgamientoPersonalidadJuridica_2020 Enero

- Formato para la denominación del nombre del diccionario de datos:

SIGLAS DE LA ENTIDAD_NombreConjuntoDatos_DD_AñoMes

Ejemplos:

MINTEL_RemuneracionPuesto_DD_2020Mayo

MERNNR_ServicioOtorgamientoPersonalidadJuridica_DD_2020 Enero

- Formato para la denominación del nombre del archivo de metadatos:

SIGLAS DE LA ENTIDAD_NombreConjuntoDatos_PM_AñoMes

Ejemplos:

MINTEL_RemuneracionPuesto_PM_2020Mayo

MERNNR_ServicioOtorgamientoPersonalidadJuridica_PM_2020Enero

- No usar artículos ni preposiciones.
- Usar únicamente letras y números ASCII, comprendidos en el rango "a-z" y "0-9".
- Separar las palabras con guión bajo "_".
- No usar tildes ni caracteres especiales.
- Usar palabras en singular.
- Usar palabras en español.
- Incluir año (4 dígitos) y mes (en letras con mayúscula inicial).

2.3.1.5. Consideraciones al generar los archivos a ser publicados

Los archivos a ser publicados deben poseer las siguientes características¹⁰:

- Los datos deben ser claros para su reutilización, las imágenes deben ser legibles, y en caso de ser audios deben ser entendibles.
- Los archivos deben garantizar una cómoda descarga de acuerdo con su volumen.
- Cuando se trata de mostrar gráficos estadísticos o geográficos, se debe incluir los datos que originaron las imágenes, caso contrario no se consideran datos abiertos.

Se promueve la publicación de datos abiertos en formatos estructurados, siendo los más utilizados:

Formato	Funcionalidades
---------	-----------------

¹⁰ <http://www.iim.csic.es/wp-content/uploads/2017/09/Datos-abiertos-2020.pdf>

Formato	Funcionalidades
CSV (Comma Separated Values)	Es un tipo de documento en formato abierto sencillo para representar datos en forma de tabla donde las columnas son separadas por comas (o punto y coma, donde la coma es el separador decimal), y las filas por saltos de línea.
ODS / XLS (Open Documento Spreadsheet / Archivos de Excel)	Documento abierto en el caso de ODS, y propietario en el caso de Microsoft XLS; muestra la información en celdas organizadas en filas y columnas.
XML (Extensible Markup Language)	Es un lenguaje general basado en etiquetas utilizado para representar información estructurada, define un conjunto de reglas de codificación a través de modelos específicos legibles y procesables. Es uno de los formatos más utilizados para la compartición de información entre programas, entre personas, entre programas y personas, tanto a nivel local como en red.
RDF (Resource Description Framework)	Es un modelo de representación para la descripción de recursos siguiendo los fundamentos de la web semántica. Se representa la información utilizando expresiones con la forma sujeto-predicado-objeto; el sujeto es el recurso que se describe, el predicado es la propiedad sobre la cual se quiere establecer el recurso y el objeto es el valor de la propiedad con el cual se establece la relación. A fin de que sea interpretable, se representa en formato XML.
RSS (Really Simple Syndication)	Es un formato de datos basado en XML utilizado para redifundir contenidos a suscriptores de un sitio web. Se suelen utilizar para describir brevemente contenidos cuya actualización es bastante frecuente.
JSON (JavaScript Object Notation)	Es un formato ligero para el intercambio de datos basado en una simplificación de la notación JavaScript.
SHP (ESRI Shapefile)	Formato propietario de datos espaciales, sirve para almacenar tanto la localización de los elementos geográficos como los atributos asociados a ellos.
VMS (Web Map Service)	Formato de servicio web para producir mapas de datos referenciados espacialmente, de forma dinámica a partir de información geográfica.
GML (Geography Markup Language)	Vocabulario XML para el modelaje, transporte y almacenamiento de información geográfica.
KML (Keyhole Markup Language)	Vocabulario XML para presentar datos geográficos en tres dimensiones. Suelen distribuirse comprimidos en formato KMZ.

TABLA 3. FORMATOS DE DATOS ABIERTOS

2.3.2. Niveles de apertura de los datos abiertos

A continuación, se expone el modelo de cinco estrellas propuesto por Tim Berners-Lee¹¹, con las características a imprimir a los conjuntos de datos, cuya aplicación responderá a las capacidades técnicas de la institución en el manejo de la información.

★	Publicar datos en la web con cualquier formato y bajo una licencia abierta
★★	Publicar como datos estructurados
★★★	Usar formatos no propietarios

¹¹ <https://5stardata.info/es/>

★★★★	Uso de estándares abiertos de W3C (Consortio World Wide Web) de manera que los datos puedan ser identificados mediante una URI (Identificador uniforme de recursos) y que así estén integrados en la web. Para este nivel una forma de representación de los datos es RDF.
★★★★★	Vinculando además los datos con los que otras personas o instituciones publican, de modo que se proporcione un contexto para ellos. El formato para este nivel es el linked RDF

TABLA 4. ESQUEMA DE DESARROLLO DE 5 ESTRELLAS PARA DATOS ABIERTOS, ADAPTADO DE [HTTPS://5STARDATA.INFO/EN/](https://5stardata.info/en/)
Características de metadatos

Los metadatos deben estar presentes en todos los niveles de apertura y se estructurarán considerando el estándar DCAT, priorizando los datos en origen previo a la publicación. En la sección Estructura de Metadatos se describe los campos a implementarse.

2.3.2.1. Primera Estrella

Exige que los datos estén disponibles en la web, sea en el formato que sea, con licencia abierta. Entre los tipos de archivos para cumplir este nivel se encuentra el PDF.

Dominio semántico (vocabulario común)

Las características de este nivel permiten lo siguiente:

- Acceso libre al archivo
- Posibilidad de modificación
- Capacidad de compartir
- Se requiere de la creación de software la extracción de los datos del documento o su copiado manual
- Facilita la publicación de los datos
- Dificulta la reutilización por máquina
- Los tipos de archivos en esta estrella, incluyen:

Tipo de archivo	Cerrado	Abierto	Libre
Documentos solo de lectura		PDF	EPUB, MOBI, DJVU, PS
Imágenes			PNG, JPEG-2000, TIFF
Archivos de texto		RTF	ODT
CAD		DXF, SAT	IGS, STP
Audios			OGA, FLAC
Videos	MPEG-4		OGV, WEBM
Data Geo		KML, KMZ	

TABLA 5. TIPOS DE ARCHIVOS ★ ESTRELLA

Al preparar conjuntos de datos en este nivel, se debe tener las siguientes consideraciones:

- Verificar que las imágenes sean claras, si se trata de letras que sean legibles.
- Definir claramente los metadatos que facilite la utilización del recurso.
- No utilizar varios formatos en mismo documento (A3, A4), lo cual dificulta la lectura.
- Verificar que los datos representen a los de origen.

2.3.2.2. Segunda Estrella

Los datos deben estar disponibles de manera estructurada, para que sean legible por máquinas. Los formatos que se manejan en esta estrella pueden ser: XLS, DOC, MDB.

Dominio semántico (vocabulario común)

Las características de este nivel permiten lo siguiente:

- Exportar los datos, porque se encuentran en formato tabular.
- Se requiere un software propietario para extraer datos.
- Facilitan su procesado directo para realizar modificaciones, cálculos o visualizaciones.
- Los datos están accesibles en la web en forma estructurada
- El conjunto de datos contiene datos numéricos y alfanuméricos dispuestos en tablas. Cada tipo de unidad observacional es una entidad, donde cada variable es una columna y cada observación es una fila.

Al preparar conjuntos de datos en este nivel, se debe tener las siguientes consideraciones:

- No hacer uso de colores u otras señales estilísticas para codificar la información.
- El archivo no debe tener macros ni fórmulas que puedan alterar el contenido al momento de copiarla a otro archivo.
- Incluir dominios de datos estandarizados (por ejemplo, distribución política, tipo de institución, nombre de instituciones), emitido por INEC.

2.3.2.3. Tercera estrella

Al igual que la segunda estrella los datos deben estar disponibles en forma estructurada, pero en este caso los formatos deben ser no propietarios.

Dominio semántico (vocabulario común)

Las características de este nivel permiten lo siguiente:

- Se podrá hacer todas las cosas que puede hacer con el nivel dos estrellas, pero sin la limitación que impone el uso de un software en concreto.
- La preparación es sencilla previa a la publicación de datos, pero podrían surgir necesidades como la conversión de formatos o la necesidad del uso de plugins de exportación desde un formato propietario
- Datos disponibles en la web y dispuestos de manera que puede utilizarse sin restricciones.
- Formatos generales, los más habituales son CSV, XML y RDF, JSON, KML.

2.3.2.4. Cuarta estrella

Este nivel requiere el uso de estándares abiertos de W3C (Consortio World Wide Web) de manera que los datos puedan ser identificados mediante una URI (Identificador uniforme de

recursos) y que así estén integrados en la web. Para este nivel una forma de representación de los datos es RDF.

Dominio semántico (vocabulario común)

Las características de este nivel permiten lo siguiente:

- Enlazar los datos desde otro sitio web, marcarlos como favoritos, reutilizar partes de los mismos, usar herramientas y librerías existentes incluso en el caso en que estas solo puedan manejar parte de los patrones de los datos, o incluso combinarlos con otros.
- Ofrece distintos formatos de serialización entre los más conocidos están: RDF/XML, RDF/JSON.
- Facilita la interpretabilidad de los contenidos con respecto al uso de XML o RDF por parte de aplicaciones o máquinas si se utiliza el Lenguaje de Ontologías Web (OWL).
- Las URI seguirán una estructura uniforme con la finalidad de facilitar a los agentes reutilizadores su lectura e interpretación por ejemplo para reconocer la temática del conjunto de datos, a qué organismo pertenece, etc.

Al preparar conjuntos de datos en este nivel, se debe tener las siguientes consideraciones:

- Se requiere el análisis de los datos, la preparación para su representación, la asignación de las URI y la búsqueda y creación de patrones para aplicarlos a la información.
- Declarar recursos en forma de expresiones denominadas triplets que tienen la forma sujeto-predicado-objeto. El sujeto se encarga de indicar el recurso, el predicado expresa la relación entre el sujeto y el objeto apuntando sus rasgos o aspectos, y el objeto sería el valor de la relación.
- Utilizar estándares abiertos de la Web semántica para identificar y consultar recursos: RDF, SPARQL.
- Describir la semántica de los recursos siguiendo las recomendaciones definidas por la W3C, y utilizar en medida de lo posible los vocabularios definidos en Linked Open Vocabularies (ontologías).
- El dataset debe ser descrito a través de la recomendación DCAT definida por la W3C.
- Definir la procedencia de datos (confianza), por ejemplo, con la ontología PROV-O¹² que permite expresar la información del origen de los conjuntos de datos.
<https://www.w3.org/TR/dwbp/#provenance>

¹² <https://www.w3.org/TR/prov-o>

2.3.2.5. Quinta Estrella

Este nivel requiere haber logrado los niveles anteriores y vinculando además los datos con los que otras personas o instituciones publican, de modo que se proporcione un contexto para ellos. El formato para este nivel es el linked RDF. (Berners-Lee, 2009)

Dominio semántico (vocabulario común)

Las características de este nivel permiten lo siguiente:

- Descubrir otros datos relacionados mientras se utilizan, y aprender directamente acerca de su esquema.
- Aumentar el valor de los datos y se logra que sea posible descubrirlos de manera más sencilla.
- Utilización de estándares abiertos de la Web semántica para identificar y consultar recursos: RDF, SPARQL.
- Los enlaces permitirán realizar integración de datos basado en los recursos y descubrir relaciones entre fuentes de datos diferentes.

Al preparar conjuntos de datos en este nivel, se debe tener las siguientes consideraciones:

- Se debe incluir URI para los recursos.
- Se requiere realizar los enlaces con otros en la web con el fin de proveer contexto, así como realizar un mantenimiento óptimo de los datos, sus URI identificativas y sus enlaces.
- Describir la semántica de los recursos siguiendo las recomendaciones definidas por la W3C, y extender en medida de lo posible los vocabularios definidos en Linked Open Vocabularies (ontologías).
- La licencia puede ser definida de las siguientes maneras:
 - Genérica para todos los recursos que se publiquen en el portal.
 - Licencias independientes por Catálogos de datos, Conjuntos de datos (datasets), Recursos, Servicio de datos, vocabularios semánticos, ontologías, y otros elementos que forman parte del dataset abierto.
 - Algunos ejemplos de licencias se pueden encontrar en: <https://www.w3.org/TR/dwbp/#licenses>
- Definir la procedencia de datos (confianza), por ejemplo, con la ontología PROV¹³ que permite expresar la información del origen de los conjuntos de datos.

¹³ <https://www.w3.org/TR/prov-o>

2.3.3. Fase de construcción

2.3.3.1. Licencias de datos abiertos

Al establecer la licencia bajo la cual se van a publicar los datos, se debe considerar que, cuando el origen del conjunto de datos es de varias instituciones, la responsable de la información de origen definirá el tipo de licencia con la que se abrirá el conjunto de datos. Para efectos de la publicación de los conjuntos de datos, en los metadatos, en el campo Licencias, se debe señalar el nivel de utilización que puede darse a los conjuntos de datos. Para esto cada entidad deberá definir el nivel de licenciamiento de cada conjunto de datos el cual será aprobado por el Comité de Transparencia en conjunto con el área Jurídica, para esto se deberá establecer los siguientes acuerdos:

- Determinar qué derechos de propiedad intelectual existen en el conjunto de datos.
- Aplicar una licencia adecuada considerando las alternativas que ofrece Creative Commons¹⁴ (definidas internacionalmente), que identifique todos esos derechos y soporte la definición de dato abierto.

2.3.3.2. Estructura de los metadatos

La importancia que tienen los metadatos en los procesos de apertura de datos está estrechamente ligada a la reutilización, ya que ésta podría mejorarse proporcionando metadatos estructurados considerables que permitan la comprensibilidad de los datos a reutilizar¹⁵. El uso adecuado de los conjuntos de datos abiertos muchas veces se ve obstaculizado porque los metadatos están mal documentados, es decir son incompletos e imprecisos, y por lo tanto la semántica de los datos puede ser ambigua¹⁶.

Los metadatos son el puente entre el conjunto de datos y el usuario (potencial reutilizador), el mismo permitirá entender la estructura del conjunto de datos, así como otros datos relacionados al manejo de la información publicada.

Basados en el Vocabulario de Catálogos de Datos (DCAT)¹⁷ y considerando importante incluir el nivel de anonimización, licencia y periodicidad del dato, por cuanto estos datos son requeridos para entender la semántica del conjunto de datos publicado, se utilizará la siguiente estructura de metadatos:

Campo	Descripción	Nombre del campo	Obligatorio
Identificador	Identificador único de un conjunto de datos de la organización.	met:identificador	SI
Nombre	Título del conjunto de datos.	met:nombre	SI

¹⁴ <https://creativecommons.org>

¹⁵ <https://www.jedem.org/index.php/jedem/article/view/138>

¹⁶ <https://dl.acm.org/doi/abs/10.1145/2463728.2463731>

¹⁷ <https://www.w3.org/TR/vocab-dcat-2/>

Descripción	Descripción textual completa del conjunto de datos.	met:descripcion	SI
Código institución	Sitio web de la organización.	met:institucion	SI
Licencia	Licencia aplicable al conjunto de datos (es decir, condiciones de reutilización).	met:licencia	SI
Fecha de creación	Fecha de creación del conjunto de datos.	met:creacion	SI
Frecuencia de actualización	Frecuencia con la que se actualizan los datos del conjunto de datos: Tiempo Real Diaria Semanal Mensual Anual	met:frecuencia	SI
Archivo maestro	La actualización se realiza a través de un archivo maestro de datos abiertos.	met:archivomaestro	OPCIONAL
Anonimización	Estado de anonimización del conjunto de datos.	met:anonimizacion	SI
Responsable	Correo de contacto del gestor institucional	met:responsable	SI
Lenguaje	Idioma de publicación.	met:lenguaje	SI
Etiquetas	Palabras clave de búsqueda para el conjunto de datos.	met:etiquetas	SI
URI	Dirección permanente del conjunto de datos	met:uri	SI

TABLA 6. ESTRUCTURA DE METADATOS

2.3.3.3. Estructura del Diccionario de datos

El diccionario es una herramienta que contribuye al fortalecimiento de la gestión del conjunto de datos por parte de los usuarios, en la cual se encuentran las descripciones de cada uno de los campos del conjunto de datos. De la misma manera que se publica el conjunto de datos, la especificación del diccionario de datos debe encontrarse en formato abierto.

Para la creación del diccionario de datos se debe considerar lo siguiente:

- No contener tildes en el nombre de los campos.
- Ser lo más claro y detallado posible, logrando que pueda ser interpretado y usado fácilmente.
- Utilizar la codificación estándar definida por INEC (Instituto Nacional de Estadísticas y Censos) para país/provincia/cantón/parroquia, y para la identificación de los ministerios.
- Para los datos utilizar codificación UTF-8.

2.3.3.4. Aplicación de criterios de calidad de datos abiertos

En la actualidad la cantidad de datos generados y custodiados por las entidades ha crecido exponencialmente, esta tendencia responde a la constante evolución de la tecnología marcada fuertemente por la expansión de Internet, el crecimiento en las capacidades de los sistemas de información y la consecuente demanda de acceso a la información.

En el contexto de uso y reutilización de los datos abiertos surge la importancia sobre la calidad de los mismos. Si bien no abarcamos todos los aspectos, proponemos criterios mínimos para realizar una evaluación previa a la publicación, que responden a normas internacionales; en base a buenas prácticas de otros países sugerimos aplicar el estándar ISO/IEC 25012 - Data Quality Model que define un modelo general aplicable a los datos que se encuentran almacenados de manera estructurada y forman parte de un sistema de información.

Los criterios a evaluar son:

Métrica	Definición	Cómo calcular
Proporción de errores	¿Cuántos errores tiene en relación con el tamaño de su conjunto de datos?	Divida el número total de datos erróneos (campo) por el número total de elementos.
Cantidad de valores vacíos	Los valores vacíos indican que falta información en un conjunto de datos.	Cuente el número de campos que están vacíos dentro de un conjunto de datos.
Tasas de error de transformación de datos	¿Cuántos errores surgen al convertir la información a un formato diferente?	Cuente el número de campos que se alteraron.
Tasa de campos con codificaciones definidas	¿Qué porcentaje de registros que están dentro de los rangos de las codificaciones establecidas?	Cuente el número de campos que no cumplen con las codificaciones que utiliza la organización o estándares nacionales e internacionales.
Tasa de campos con simbología inapropiada	¿Qué porcentaje de registros tienen caracteres especiales?	Relación del número de registros con caracteres especiales en relación al total de registros

TABLA 7. MÉTRICAS PARA LA CALIFICACIÓN DE CRITERIOS DE CALIDAD DE LOS DATOS ABIERTOS

Si se detectan errores en base a las mediciones propuestas, debe realizar la inspección y analizar la causa raíz de fallas, aplicar las medidas correctivas que sean necesarias para reducir el número de errores.

2.3.4. Fase de publicación

La información que será puesta a disposición de los usuarios a través de los portales web, son:

- Portafolio institucional de datos abiertos: la planificación de los conjuntos de datos que la institución va a publicar
- Conjuntos de datos: archivos o enlaces que contienen datos sin restricciones de acceso y utilización para el usuario.
- Metadatos: información complementaria que facilita el entendimiento de lo contenido en los conjuntos de datos.
- Diccionario de datos: información que describe el contenido de los conjuntos de datos.

El detalle de estos instrumentos se encuentra en las secciones: 2.2.1. Construcción del Portafolio institucional de datos abiertos; 2.3.1. Fase preparatoria; 2.3.3.2. Estructura de los metadatos; 2.3.3.2. Estructura del Diccionario de datos.

2.3.4.1. Plataforma gubernamental de publicación de datos

La publicación se realizará en una plataforma de código abierto que facilita la publicación de catálogos de datos, y es el lugar donde se puede encontrar los datos generados por las entidades dependientes de la Administración Pública Central.

En el catálogo se puede realizar la búsqueda de datos por palabras clave, nombre de entidad, categoría, etiqueta o formatos. Cada conjunto de datos publicado tiene asociados una ficha de metadatos y el diccionario de datos que proporcionan el detalle de su contenido.

2.3.4.2. Interfaz de Programación de Aplicaciones (API)

Es el punto de entrada donde un sistema recibe peticiones para ejecutar una tarea previamente programada, facilitando el trabajo de los reutilizadores de datos que buscan crear nuevos productos o servicios basados en los datos abiertos.

Es recomendable el uso de las API abiertas para complementar y mejorar el acceso a los datos.

2.3.4.3. Conjuntos de datos

Los datos son publicados en unidades llamadas “datasets” o conjuntos de datos. Cuando un usuario realiza búsquedas, los resultados serán datasets individuales.

Los conjuntos de datos se construyen a partir de datos digitalizados y que se encuentran disponibles en una URI.

2.3.4.4. Recursos

Posterior a la creación del conjunto de datos con sus correspondientes metadatos, se procede a la inclusión de uno o más recursos, es decir la incorporación de los archivos que cuentan con los registros que alimentarán el conjunto de datos, mismo que puede ser un

fichero de cualquier tipo, un enlace a un fichero en otro lugar de la red o un enlace a una API.

2.3.4.5. Usuarios de la plataforma

La información disponible en la plataforma de publicación de datos es de libre acceso y utilización, y no es necesario registrarse para buscar y acceder a los datos; no obstante, para crear o editar conjuntos de datos es necesario disponer de una cuenta para la gestión del contenido en la plataforma.

La administración de la plataforma gubernamental de datos abiertos es de responsabilidad del ente rector de la Planificación, quien creará usuarios individuales por cada entidad de la Administración Pública Central.

Las entidades de la Administración Pública Central, tienen a su cargo los conjuntos de datos, y para gestionar su ingreso es necesario que se registre el “login” de un único responsable, mediante credenciales; corresponde a los Gestores institucionales manejar las claves de ingreso que se refieren a las credenciales de acceso.

Dentro de cada entidad se identifican los siguientes usuarios con sus respectivas atribuciones:

Administrador: su rol permite añadir, eliminar y cambiar roles para los miembros de la institución.

Editor: su rol permite coordinar internamente la creación, edición y publicación de los datos abiertos de la entidad. El editor inicialmente creará el conjunto de datos en privado y visible; y posteriormente podrá ser publicado para su visualización y acceso.

Participante: su rol le permite ver los conjuntos de datos que la entidad ha publicado.

2.4. PROMUEVE

El rol de las entidades del Estado en el uso de datos abiertos no termina con su publicación. También deben desempeñar un papel activo para apoyar una reutilización eficaz e innovadora de los mismos.

Objetivo

Determinar los medios, recursos, herramientas, esfuerzos a realizar, tanto al interior de la entidad así como hacia fuera de ella, para promover el conocimiento de los datos abiertos y de los conjuntos de datos publicados, e impulsar su utilización.

2.4.1. Difusión para la promoción de datos abiertos

Los datos abiertos cobran vida gracias a los usuarios, es necesario que los datos puedan ser encontrados con facilidad de manera que se promueva la generación de comunidades tales como desarrolladores de software, visualizadores, periodistas de datos, investigadores, colectivos sociales, entre otros, para que puedan reutilizarse y generar valor a partir de ellos.

Generar comunidad o una audiencia definida será una tarea central que las entidades deberán considerar en sus planes institucionales, para poder promover el uso de los datos.

Es necesario que la promoción y difusión de los datos involucre al menos:

- Identificar sus usuarios y potenciales usuarios, para ello será importante identificarlos aliados de otros sectores (sociedad civil, academia, sector privado, agencias de cooperación, etc.) y generar vínculos con los mismo
- Incorporar en las estrategias comunicacionales la promoción de los datos institucionales.
 - Generar contenidos tales como infografías, videos, cápsulas, artículos, sobre los datos disponibles de la entidad.
 - Vincular a distintos actores a través de acciones que permitan el encuentro e interacción, que promuevan la reutilización de los datos.
 - Promover el uso de tecnologías abiertas emergentes para generar productos o servicios a partir de los datos abiertos, en conjunto con los distintos actores.

Para lograrlo, se deben desarrollar al menos los siguientes pasos por parte de la entidad:

- Segmentar y caracterizar a los potenciales usuarios.
- Desarrollar un plan de comunicaciones interno que promueva la oferta, así como la demanda, facilitando la reutilización y consumo de datos; y sensibilizando sobre los beneficios.
- Identificar los canales de difusión disponibles, de acuerdo con la realidad de la institución, así como los medios de participación y colaboración para la gestión de ideas y/o propuestas, con el fin de acercar a la sociedad civil y a la ciudadanía en general. Son alternativas a considerar:
 - Hacer uso de redes sociales, en función de la población objetivo, para difundir la información disponible. Usar el perfil institucional de la entidad para la publicación.

- Coordinar acciones interinstitucionales, de ser posible, a fin de impulsar la difusión de las publicaciones que se realicen desde las entidades para promover el uso de datos.
- Hacer uso de las bases internas de usuarios y enviar e-mailing con publicidad sobre la información que va a ser publicada por la entidad.
- Establecer, dentro de las prioridades, procesos de capacitación dentro de la institución y hacia los infomediarios, academia, sector privado, sociedad civil, etc., sobre los datos que se están publicando.
- Generar campañas informativas con medios de comunicación sobre los datos publicados.
- Realizar intercambios de experiencias para retroalimentar conocimientos.
- Determinar y habilitar espacios de participación, ya sean físicos o virtuales donde se realicen iniciativas encaminadas a la utilización efectiva de los datos buscando la participación e integración entre la institución, sociedad civil, gobierno, academia, sector privado. Se deberá realizar cada año, al menos una de las siguientes opciones:
 - Concursos colegiales, universitarios de uso y aprovechamiento de la información, que motive a los estudiantes a hacer trabajos con los datos de la institución.
 - Webinars que hablen sobre el uso de datos, y visibilice estudios de casos que han utilizado los datos de la institución.
 - Gamificación que motive la utilización y reutilización de los datos de la institución a través de incentivos y/o reconocimientos.
 - Congresos o seminarios que se centren en el uso de datos, y visibilice estudios de casos que han utilizado los datos de la institución.
 - Socialización de estudios internos de la entidad con la utilización de sus datos.

2.4.2. Promover la innovación con datos de la institución

La promoción de los datos de la institución como vía para promover la innovación es uno de los usos específicos que puede darse.

La innovación es la introducción de un nuevo o significativamente mejorado producto, es decir bienes y/o servicios con cambios en su calidad, así como la introducción de nuevos o significativamente mejorados procesos de producción, es decir la forma de generar y proveer los bienes y/o servicios; métodos organizativos, con cambios estructurales y de gestión a nivel micro y macro; a través de la conexión de ideas y conocimientos

multidisciplinarios que permitirán impactos positivos y darán como resultado valor social. Innovación es una idea implementada efectivamente para resolver un problema¹⁸.

El uso de los datos de la institución se convierte en un insumo para la generación de valor por parte de los gobiernos, sociedad civil, academia, sector privado y ciudadanía en general. El desarrollo de la innovación basada en datos es fundamental para una transformación digital y que permita a toda la sociedad, beneficiarse de la cuarta revolución industrial. Las actividades de innovación incluyen todas las actuaciones científicas, tecnológicas, organizativas, financieras y comerciales que conducen a la innovación.

De este modo, para promover procesos de innovación en el país a través de la utilización de los datos abiertos gubernamentales se debe considerar lo siguiente:

- Difundir y promocionar los datos institucionales (referirse a la sección anterior).
- Realizar un mapeo de actores institucionales y potenciales usuarios de los datos.
 - Nivel interno: Actores que potencialmente usen/brinden los datos para la generación de herramientas, productos y servicios a nivel del sector público, y que orienten la transformación digital.
 - Nivel externo: Actores que potencialmente usen/brinden los datos para la construcción de innovación hacia afuera del sector público, mediante procesos de datos e innovación abierta.
- Generar y promover el desarrollo de habilidades de procesamiento y manejo de datos para que los actores tengan la oportunidad y los incentivos de utilizar la información de la institución.
 - Se sugiere generar alianzas con actores claves para brindar capacitaciones o webinars sobre procesos de innovación a partir de datos abiertos.
- Asegurar espacios de intercambio de ideas colectivas y mecanismos de participación de actores multisector en base al uso de datos de la institución en proyectos innovadores. Algunas vías son:
 - Realizar webinars que se centren en casos de éxito de proyectos de innovación con el uso de datos abiertos, y en el caso de haber, estudios que han utilizado los datos de la institución.
 - Organizar/participar de congresos o seminarios que tengan como objetivo presentar casos de éxito de proyectos de innovación con el uso de datos abiertos, y en el caso de haber, estudios que han utilizado los datos de la institución.
 - Compartir proyectos innovadores desarrollados internamente de la entidad con la utilización de sus datos.

¹⁸ PNUD, 2019

- Promover ejercicios de uso de datos tales como retos públicos de innovación, datatones, hackatones, concursos de visualización, entre otros similares.
- Implementar/crear programas que apoyen a proyectos de innovación con el uso de los datos abiertos de la institución.
- Difusión de la Innovación. Este proceso podría estar inmerso en las rendiciones de cuentas de cada entidad.
 - Se deberán difundir las innovaciones a fin de aprovechar el conocimiento desde los distintos sectores de la sociedad.
 - Las instituciones deberán valorar y reconocer la gestión y toma de decisiones en base a datos abiertos en sus diferentes áreas y deberán promover procesos de gestión del conocimiento.
 - Documentación: Se deberá guardar registro, de preferencia virtual, de las acciones (seguimiento, control, monitoreo, evaluación, resultados, retroalimentaciones) que se generen dentro de la gestión de innovación.

2.4.3. Datos abiertos para promover la investigación científica

Promover la investigación científica es uno de los usos específicos que puede darse a los datos abiertos.

La Investigación es la búsqueda sistemática y lógica de nueva información sobre un tópico particular. Cuando se aplican los pasos de método científico se tiene investigación científica, en este campo, los datos son la base sobre la cual se construye la ciencia. Un intercambio de datos más eficiente y un acceso más abierto a la información y los recursos facilitarán la confirmación de las observaciones, la replicación de los experimentos, el respaldo, el rechazo o el refinamiento de las hipótesis y, en última instancia, la respuesta a los desafíos de la sociedad.

Para promover procesos de investigación científica a través de la utilización de los datos institucionales, se debe identificar a los actores del proceso de intercambio de datos:

Investigador: Solicita datos para sus tareas de investigación. La investigación la realiza: el personal académico de las instituciones de educación superior públicas y privadas como parte de su carga académica; los investigadores nacionales o extranjeros; y, el personal de investigación de los institutos nacionales de investigación.

Institución fuente: Es la entidad que genera, custodia y almacena en cualquier medio la información solicitada.

Custodio de datos: Funcionario responsable de la Unidad Poseedora de Información donde se encuentran los datos que son requeridos para la investigación.

Con el objetivo de lograr un verdadero y efectivo fomento a la investigación, se proponen los siguientes criterios para el accionar de los interesados.

Comunicación: La entrega de datos debe ser directa entre el custodio de los datos y el investigador.

Agilidad: El proceso de solicitud, autorización e intercambio de datos debe realizarse por vía electrónica utilizando las credenciales de las instituciones donde laboren.

Trazabilidad: Los custodios e investigadores deben registrar el intercambio de datos y la evidencia de los productos generados con ellos a fin de diseminar el conocimiento generado.

Documentación: Se deberá guardar registro, de preferencia virtual, de las acciones (seguimiento, control, monitoreo, evaluación, resultados, retroalimentaciones) que se generen dentro de la gestión de investigación.

2.5. EVALÚA Y PERFECCIONA

Se establece evaluar: la percepción de los usuarios si los datos publicados son adecuados para su uso, la publicación versus la planificación del Portafolio institucional de datos abiertos, el nivel de descargas de los conjuntos de datos.

Objetivo

Medir cualitativa y cuantitativamente los resultados de las acciones emprendidas y el cumplimiento versus lo planificado.

2.5.1. Resultados de la implementación de la Política y la Guía de Datos Abiertos

El Gestor institucional de datos abiertos deberá ingresar, anualmente, en el portal gubernamental de datos abiertos, los primeros 15 días del año información que permita realizar la evaluación de la implementación de la Política y de la Guía de Datos Abiertos.

La información será suministrada a través de un formulario en línea que será dispuesto en el portal gubernamental de datos abiertos., para el efecto, el Equipo de Datos Abiertos contribuirá con el establecimiento de la siguiente información:

1. Equipo responsable.
2. Portafolio institucional anual de datos abiertos, y sus actualizaciones.
3. Conjuntos de datos publicados vs. los planificados (ver sección 2.5.2).
4. Resumen de acciones para la difusión y utilización de los datos publicados.
5. Resumen de la evaluación de la percepción de los usuarios sobre los datos publicados.
6. Indicadores de uso y descarga de datos (ver sección 2.5.3).
7. Indicadores e percepción de los usuarios sobre los datos publicados (ver sección 2.5.4).
8. Conclusiones y recomendaciones para la institución, para el ente rector y para el gestor del portal gubernamental de datos abiertos. Aportan en este aspecto, establecer el ranking de conjuntos de datos más descargados, identificar oportunidades de mejora tanto en la disponibilidad como en la calidad de los datos.

En base a los datos ingresados, el gestor del portal gubernamental de datos abiertos publicará Información pertinente, relacionada con el cumplimiento de las entidades, de manera abierta en el portal gubernamental de datos abiertos.

2.5.2. La publicación versus la planificación del Portafolio institucional de datos abiertos

El Equipo de Datos Abiertos deberá evaluar el cumplimiento del Portafolio comparando lo publicado versus lo planificado, tanto en conjuntos de datos como en tiempos.

Se debe considerar las solicitudes de publicaciones de datos abiertos, y el estado en que se encuentran, incluyendo aquellas que van a formar parte del nuevo Portafolio especificando el estado en el que se encuentran: en proceso, incluida, no considerada.

2.5.3. Nivel de descarga de conjuntos de datos

Contribuyen a tener una percepción del nivel de utilización:

- Nivel de descarga de conjuntos de datos, incluir indicadores numéricos de descargas realizadas en los medios donde se encuentren disponibles, como portales institucionales y portal gubernamental.
- Solicitudes de apertura de conjunto de datos recibidas y atendidas.
- Casos de utilización y reutilización de datos que ha registrado la institución
- Otros estudios o reportes de terceros respecto al uso de datos abiertos de la institución

2.5.4. Percepción de los usuarios de los datos publicados

La institución deberá realizar la evaluación por parte de los usuarios con el fin de validar que los datos publicados cumplen ciertas características, se propone los siguientes indicadores¹⁹ como mínimo; la entidad puede añadir otros.

Indicador	Descripción
Comprensión semántica: Porcentaje del conjunto de datos que se entiende el contenido.	Relación entre el vocabulario común reconocido y los términos o definiciones utilizados en el diccionario de datos.
Comprensión del diccionario de datos: Porcentaje de comprensión del diccionario de datos.	Relación de comprensión de los términos utilizados en el diccionario de datos, con el uso de un vocabulario común.

¹⁹ https://link.springer.com/chapter/10.1007/978-3-030-40907-4_17

Comprensión de los metadatos: Porcentaje de comprensión de los metadatos.	Relación de comprensión entre el número de metadatos entendidos frente a los metadatos de interés para los reutilizadores potenciales.
Comprensión de la licencia de uso: Porcentaje de comprensión de las licencias de los conjuntos de datos	Relación entre el conocimiento de la utilización que se le puede dar a los conjuntos de datos, entendido por la licencia que tengan.

TABLA 8. INDICADORES PARA EVALUAR LA PERCEPCIÓN DE CALIDAD DE LOS DATOS PUBLICADOS

ANEXOS DE LA GUIA:

1. HERRAMIENTA DE PLANIFICACIÓN

1.1. Portafolio institucional de datos abiertos (ejemplo sujeto a ajustes)

ENTIDAD		Portafolio de datos abiertos				Código Documento: PDA-2021-v01		
Fecha de Elaboración		01/12/2020						
Período planificado		2021						
Responsable		Juan Pérez						
E-mail Responsable		juan.perez@xx.gob.ec						
Código	Nombre	Descripción	Entidades que generan la información de origen	Fecha prevista de publicación	Formatos de publicación	Licencia	Solicitante	Categorías
A-IN-001	Censo de Información Ambiental Económica en Gobiernos Autónomos Descentralizados Municipales 2016	Contiene información de la gestión ambiental, riego y drenaje, fomento y desarrollo productivo, y gastos e inversión en gestión ambiental, de todos los Gobiernos Autónomos Descentralizados Provinciales del Ecuador	GADX, INEC	Mayo 2021	csv	CC-BY-4.0		Ambiente y vivienda
A-IN-002	Censo de Población y Vivienda Galápagos 2018	Contiene información de todas las viviendas existentes y de la población que reside habitualmente en la provincia	INEC	Septiembre 2021	csv	CC-BY-4.0		geográficos

1.2. Anexo sobre metodologías de priorización de datos

Otras metodologías y sugerencias que se pueden usar son:

- <https://www.ibge.gov.br/estatisticas/sociais/saude/9160-pesquisa-nacional-de-saude.html?=&t=o-que-e>
- <https://drive.google.com/file/d/1hYmoTZTDgFe9E8CtxAW6qfbj5W93DL/view>,
- Guías que se construyen para resolver problemáticas comunes, ejemplos de apertura de datos con propósito: <https://docs.google.com/document/d/182USj4La896XgMZ-LTw7A3HYbA5i7aP0p2JLBL8a-24/edit#heading=h.9vgybo4y3gv2>
- Consulta a las personas consumidoras de datos e información
- Buzón de sugerencias de los usuarios
- Talleres con expertos y líderes de las organizaciones de la sociedad civil, que tengan proyectos o procesos relacionados a la apertura de datos
- Open Data Guide, <https://opendata.guide/chapter2.html>
- Rally de Datos en la Calle, experiencia mexicana que consiste en organizar a la ciudadanía en grupo de 3 personas aproximadamente para que puedan salir a las calles a verificar y constatar la ejecución de políticas públicas de manera directa.
- Determinar criterios de priorización de los datos (impacto y dificultad) y ponderación de estos criterios con lo cual se obtendrá los plazos de publicación de los datos si es a corto, mediano o largo plazo.
- Construir una matriz de interesados, donde se incluya los interesados, la información que requieren, la urgencia, la frecuencia con que requieren los datos, para que serán usados los datos, y la unidad responsable de entregar la información especificando la importancia y celeridad con que atiende cada solicitud (bajo, medio alto).

2. HERRAMIENTA DE PREPARACIÓN

2.1. Referencia de licencias para publicación de datos abiertos

<https://creativecommons.org/licenses/by-sa/4.0/legalcode.es>

2.2. Ficha de metadatos (ejemplo sujeto a ajustes)

Institución	Ministerio de Telecomunicaciones y de la Sociedad de la Información
Identificador (Código de este Documento)	TM-01-2020-V01
Documento Referencia (nombre del archivo de datos)	MINTEL_Directorio_2020Mayo
Nombre (Nombre o título del Conjunto de Datos)	Directorio Institucional
Descripción (Descripción del Conjunto de Datos)	Contiene información para ubicación de los funcionarios que laboran en la institución.
Código institución	www.telecomunicaciones.gob.ec
Licencia	CC-BY-4.0
Fecha de creación	15/6/2020
Frecuencia de actualización	mensual
Archivo maestro (Fuente)	Base de Datos Sistema de Recursos Humanos
Anonimización	No
Responsable (correo del funcionario dentro de la institución)	juan.perez@mintel.gob.ec
Lenguaje	español
Etiquetas	recursos humanos, personal
URI	www.telecomunicaciones.gob.ec/transparencia/

2.3. Diccionario de datos (ejemplo sujeto a ajustes)

Institución	Ministerio de Telecomunicaciones y de la Sociedad de la Información
Identificador (Código de este Documento)	DD-01-2020-V01
Documento Referencia (Nombre del archivo de datos)	MINTEL_Directorio_2020Mayo
Nombre (Nombre o título del Conjunto de Datos)	Directorio Institucional
Descripción (Descripción del Conjunto de Datos)	Contiene información para ubicación de los funcionarios que laboran en la institución.
Nombre del campo	Descripción del campo
No.	Número de ítem o funcionario
Apellidos y Nombres de los servidores y servidoras	Nombre del archivo del conjunto de datos al que está asociado este diccionario de datos+B7:D12
Unidad a la que pertenece	Unidad departamental a la que pertenece el funcionario dentro de la institución
Dirección institucional	Dirección física de la institución
Ciudad en la que labora	Ciudad en la que desempeña sus labores
Teléfono institucional	Número de teléfono o PBX de la institución
Extensión telefónica	Número de extensión telefónica del funcionario dentro de la institución
Correo Electrónico institucional	Dirección de correo electrónico (email) del funcionario dentro de la institución

3. HERRAMIENTA PARA LA PUBLICACIÓN

3.1. Categorías

Las categorías involucran a los sectores donde los datos puedan causar impacto o responden a intereses comunes. Los conjuntos de datos pueden asociar una o más categorías, entre las siguientes:

- 1 Administración Pública
- 2 Ambiente y agua
- 3 Sociedad y Bienestar
- 4 Cultura y Deporte
- 5 Comercio e Industria
- 6 Conocimiento e Investigación
- 7 Estadísticas demográficas
- 8 Energía
- 9 Economía y Finanzas
- 10 Educación
- 11 Geográficos
- 12 Salud
- 13 Seguridad y Defensa
- 14 Transporte
- 15 Vivienda
- 16 Tecnología y telecomunicaciones

17 Turismo

Otras categorías que pueden usarse para proyectos concretos, obligación legal o un tema en particular. Pueden ser:

- Los Objetivos de Desarrollo Sostenible
- Objetivos del Plan Nacional de Desarrollo
- Datos para la Transparencia y ética en la administración pública
- Sectores
- Datos para la Contratación Pública
- Datos de investigación científica

PARTICIPANTES DE LA CO-CREACIÓN

Equipo coordinador del proceso:

Nombre	Sector
Andrea Llerena	Sociedad Civil
Francisco Silva	Gobierno
Johanna Pazmiño	Gobierno
Juan Carlos Castillo	Gobierno
Julio López	Sociedad Civil
Margarita Yépez	Sociedad Civil
Robert Enríquez Reyes	Academia
Susana Cadena Vela	Academia

Equipo de expertos que participaron en la construcción de la propuesta de Guía de Datos Abiertos:

Nombre	Sector	País
Alexandra Suasnavas	Ciudadanía	Ecuador
Ana Grijalva	ONG	Ecuador
Anderson Castro	Gobierno	Ecuador
Andrés Vivanco	Ciudadanía	Ecuador
Aníbal Garrido	Privado	Ecuador
Augusto Polivio Gómez Bautista	Gobierno	Ecuador
Carlos Julio León Caicedo	Gobierno	Colombia
Carlos Ruz	Ciudadanía	Ecuador
Claudia Olguín	Ciudadanía	Ecuador
Claudio Alejandro Valdivia Martínez	Privado	Perú

Nombre	Sector	País
Daniel Ochoa	Academia	Ecuador
Efrén Díaz	Privado	España
Emilefran Pazmiño	Transparencia y Control Social	Ecuador
Fabio A. Bazurto Blacio	Privado	Ecuador
Freddy Tapia León	Academia	Ecuador
Gabriela Revelo	Gobierno	Ecuador
Genny Vélez	Transparencia y Control Social	Ecuador
Gissela Carballo	Ciudadanía	Ecuador
Guido Moncayo	Transparencia y Control Social	Ecuador
Héctor Revelo Herrera	Academia	Ecuador
Irma Elizabeth Cadme Samaniego	Academia	Ecuador
Isabel Irene Montero	Gobierno	Ecuador
Ivanova Teresa Procel Vega	Ciudadanía	Ecuador
Jessica Vásquez	Ciudadanía	Ecuador
Jorge Andrey Beita Jiménez	Sociedad Civil	Costa Rica
Jorge Chicaiza	GAD	Ecuador
Jorge López	Academia	Ecuador
José Orellana	Privado	Ecuador
Juan Pablo Marín	Privado	Colombia
Julián Salazar	Privado	Ecuador
Liliana Elizabeth Gutiérrez Torres	Gobierno	Ecuador
Luis Alfredo Guambo Alcocer	Ciudadanía	Ecuador
María del Carmen Cabrera Loayza	Academia	Ecuador

Nombre	Sector	País
María José Bucheli	Gobierno	Ecuador
Mayra Montalván	Ciudadanía	Ecuador
Nelson Piedra	Academia	Ecuador
Nina León Álvarez	Ciudadanía	Ecuador
Oscar Patricio Alvear	Academia	Ecuador
Paola Espinoza Arias	Ciudadanía	Ecuador
Rene Rolando Elizalde Solano	Academia	Ecuador
Richard Urrego	Gobierno	Ecuador
Sofía Cañadas	Academia	Ecuador
Stalyn Darío Castro Tapia	Ciudadanía	Ecuador
Stephanie Naranjo	Ciudadanía	Ecuador
Susana Soto González	Ciudadanía	Costa Rica
Tania Gualli	Academia	Ecuador
Víctor Hugo Saquicela Galarza	Academia	Ecuador
Wilson Rodrigo Pérez Rocano	Academia	Ecuador