



# Lineamientos para la adecuación y publicación de **datos abiertos**

Versión 1.0



# **Lineamientos** para la adecuación y publicación de **datos abiertos**

# Lineamientos para la adecuación y publicación de datos abiertos

DAT-001

Este documento ha sido elaborado por los miembros del Consejo para las Tecnologías de Información y Comunicación del Estado Plurinacional de Bolivia (CTIC-EPB)

Coordinación Secretaría Técnica del CTIC-EPB: Cristina Loma y Carolina Ovale

Grupo de Trabajo de Datos Abiertos: Oscar Adriázola, Nila Asturizaga, Mario Silva, Ronny Espinoza, Rolando Aguilar, Elizabeth Quispe, Ivan Rojas, Fernando Fuentes, Stephanie Diaz, Adrian Maqueño, Vladimir Yugar, Ariel Zabala, Gary Ramos, José Ovando, Sylvain Lesage, Marcelo Pinto, Carlos Aguilar, Daniel Mendoza, Milenka Altamirano, Renan Yugar, Guillermo Choque, Nelson Huanquiri, Ximena Soto, Luis Rejas, Sergio Ardaya, Alejandro Salamanca, Javier Badani, Henry Taby, Ricardo Trujillo, Noemi Ticona, Judith Apaza, Miriam Jemio, Wilfredo Jordán, Eliana Quiróz, René Mendoza, Raúl Molina, Martin Meruvia, Mariela Padilla y Percy Zenteno.

Edición, diseño y diagramación: Natalia Antezana, Mariela Padilla y Orestes Sotomayor

Depósito Legal: 4-2-196-17P.O.

Se autoriza la reproducción total o parcial de este documento citando la fuente, así como el uso del mismo para obras derivadas que se distribuyan en las mismas condiciones.

La Paz, Bolivia 2017

## Contenido

<b>I. Introducción.....</b>	<b>6</b>
<b>II. Antecedentes.....</b>	<b>8</b>
<b>1. Marco normativo referencial.....</b>	<b>15</b>
<b>2. Objetivo.....</b>	<b>17</b>
<b>3. Alcance y ámbito de aplicación.....</b>	<b>17</b>
<b>4. Conceptualización de datos abiertos.....</b>	<b>17</b>
<b>5. Características.....</b>	<b>17</b>
5.1. Estructurados.....	17
5.2. Detallados.....	18
5.3. Con metadatos.....	18
5.4. Disponibles y permanentes.....	18
5.5. Oportunos.....	18
5.6. No discriminatorios.....	18
5.7. Veracidad de los datos.....	18
<b>6. Estándares de datos abiertos.....</b>	<b>18</b>
6.1. Formatos.....	18
6.1.1. Sobre el formato PDF.....	20
6.2. Estructura del metadato.....	20
<b>7. Protocolos de preprocesamiento.....</b>	<b>22</b>
7.1. Codificación estándar de caracteres.....	22
7.2. Encabezado de tablas y encabezados de los conjuntos de datos.....	22
7.3. Título de archivos generados.....	22
7.4. Delimitación numérica para millares y decimales.....	23
7.5. Formato de fechas y horas (24 horas).....	23
7.6. Compresión y empaquetamiento de archivos.....	23
<b>8. Licencia para la publicación y uso de datos abiertos.....</b>	<b>23</b>
8.1. Declaración de uso de Datos Abiertos.....	23
<b>Anexo 1.....</b>	<b>24</b>
<b>Anexo 2.....</b>	<b>25</b>

## I. Introducción

El Consejo para las Tecnologías de Información y Comunicación del Estado Plurinacional de Bolivia (CTIC-EPB) fue creado el año 2015 mediante D.S. 2514, con el fin de apoyar el desempeño de las entidades e instituciones de gobierno en el campo de la implementación de Tecnologías de Información y Comunicación y Gobierno Electrónico. Esto, mediante la elaboración consensuada de normas, estándares y guías orientadas al mejoramiento de la calidad de los servicios ofrecidos a la ciudadanía y al fortalecimiento de las capacidades institucionales, para así contribuir al desarrollo de acciones coordinadas.

La visión de trabajo del CTIC-EPB se tradujo en una amplia gama de actividades desplegadas a lo largo de su funcionamiento a partir de su lanzamiento

en el mes de mayo de 2016. Es destacable el compromiso e involucramiento de las entidades e instituciones nacionales, subnacionales, académicas y de la sociedad civil para el diseño de una estructura institucional orientada a la generación de normas técnicas a corto plazo.

Un reflejo de la tarea realizada hasta el momento, es la conformación de seis grupos temáticos de trabajo, en los que participaron más de 240 personas en alrededor de 60 reuniones presenciales. Las actividades de los equipos estuvieron respaldadas por un sostenido esfuerzo de reflexión, investigación y estudio de experiencias institucionales, plasmadas en textos, guías y normas técnicas. A la fecha se han revisado y aprobado dos documentos de trabajo elaborados por los grupos de Datos Abiertos y de Software Libre, de los cuales se pretende una amplia divulgación.

La propuesta del Grupo de Trabajo de Datos Abiertos establece los lineamientos específicos para la adecuación y publicación digital de datos por parte de las entidades del sector público u organismos del Estado, a fin de promover la transparencia, el acceso a información, la innovación y el control social. A través de las tareas realizadas por los participantes, con un alto porcentaje de personas provenientes de la sociedad civil, se logró conceptualizar, definir características y establecer una declaración que promueva el acceso, interpretación, análisis y comunicación de los datos abiertos e información pública.

Esta propuesta es muy importante para el avance del Estado Plurinacional de Bolivia hacia una visión de gobierno electrónico con principios de transparencia activa. La información generada en el Estado debe estar al alcance de todas y todos sin ningún tipo de discriminación ni restricción. Esto permitirá iniciar procesos de co-creación de valor

e innovación a partir de los datos abiertos gubernamentales, generando valor público en el ámbito social, económico y político tanto para la sociedad en su conjunto como para la administración gubernamental.

Como ya se mencionó, el CTIC-EPB es un espacio propositivo en el que se coordina y delibera participativamente la implementación de las Tecnologías de Información y Comunicación y de Gobierno Electrónico en el Estado Plurinacional de Bolivia. Sin embargo, también es un espacio de diálogo y de encuentro entre el Estado y las y los ciudadanos, pues estos últimos son los beneficiarios finales de este profundo cambio en el modelo de gestión gubernamental. En este sentido, crear espacios para la elaboración de políticas de manera conjunta con la sociedad civil es vital, cuando el Estado pretende brindar servicios en función de las necesidades y demandas de las y los bolivianos.

## II. Antecedentes

La elaboración del presente documento se enmarca en el mandato institucional respaldado por:

- Decreto Supremo N° 2514 que en su Disposición Final Primera incorpora el inciso t) en el Artículo 22 del Decreto Supremo N° 29894, de 7 de febrero de 2009, de Organización del Órgano Ejecutivo, que establece que: "el Ministerio de la Presidencia es el ente rector de Gobierno Electrónico y de Tecnologías de Información y Comunicación para el sector público del Estado Plurinacional de Bolivia, siendo el encargado de establecer las políticas, lineamientos y normativa específica para su implementación, seguimiento y control".
- El Artículo 2 del Decreto Supremo N° 2514 sobre la creación de la Agencia de Gobierno Electrónico y Tecnologías de Información y Comunicación - AGETIC, como "una institución pública descentralizada de derecho público, con personalidad jurídica, autonomía de gestión administrativa, financiera, legal y técnica y patrimonio propio, bajo tuición del Ministerio de la Presidencia".
- Lo señalado en los artículos 9, 10 y 11 del Decreto Supremo N° 2514 sobre la creación del Consejo para las Tecnologías de Información y Comunicación para formular y presentar propuestas de políticas, normativa, programas y proyec-

tos de Gobierno Electrónico y Tecnologías de Información y Comunicación por parte de entidades del ámbito gubernamental.

El Consejo para las Tecnologías de Información y Comunicación del Estado Plurinacional de Bolivia (CTIC-EPB) se constituye en una instancia de coordinación técnica para la implementación de Gobierno Electrónico y para el uso y desarrollo de Tecnologías de Información y Comunicación en el país.

Entre las principales tareas asignadas al CTIC-EPB se encuentran:

- Formular propuestas de políticas y normativa relacionada con Gobierno Electrónico, a ser presentadas a la AGETIC;
- Presentar proyectos y programas de Gobierno Electrónico y Tecnologías de Información y Comunicación en el ámbito gubernamental a la AGETIC para su gestión;
- Generar mecanismos de participación de instituciones y organizaciones de la sociedad civil en la proposición y formulación de políticas y acciones relacionadas con Gobierno Electrónico y Tecnologías de Información y Comunicación en el ámbito gubernamental;
- Establecer espacios de coordinación entre las entidades del sector público para el desarrollo conjunto de programas, proyectos o acciones de Gobierno Electrónico y Tecnolo-



gías de Información y Comunicación en el ámbito gubernamental;

- Desarrollar y proponer estándares abiertos oficiales del Estado Plurinacional de Bolivia en materia de Gobierno Electrónico y Tecnologías de Información y Comunicación aplicables a las entidades del sector público;
- Establecer dentro del Estado, con la ciudadanía y a nivel internacional, espacios de coordinación de comunidades de desarrollo informático.

El 5 de mayo de 2016 se llevó a cabo la inauguración y la primera Reunión del Pleno del CTIC-EPB, en la que se conformaron seis grupos temáticos de trabajo: Interoperabilidad, Software Libre, Seguridad, Infraestructura, Desarrollo de Software y Datos Abiertos.

Cada Grupo de Trabajo estuvo conformado por servidores públicos de las entidades del nivel central del Estado: Órgano Ejecutivo, Legislativo, Judicial y Electoral, incluyendo sus instituciones descentralizadas, autárquicas, empresas públicas y autoridades de regulación sectorial; Ministerio Público y Procuraduría General del Estado. Adicionalmente, se invitó a participar, en calidad de miembros adjuntos, a representantes de entidades territoriales autónomas, universidades públicas e indígenas y sociedad civil, a fin de trabajar y elaborar propuestas a ser presentadas al Consejo para su posible implementación a nivel estatal.

Cabe mencionar que el desarrollo de los Grupos de Trabajo y del Consejo se enmarca en el Reglamento de Funcionamiento del CTIC-EPB, aprobado mediante la Resolución Administrativa N° 024/2016, del 31 de mayo de 2016.

El Grupo de Trabajo de Datos Abiertos se planteó como objetivo elaborar los lineamientos específicos para la implementación de iniciativas de datos abiertos del Estado Plurinacional de Bolivia. Sin embargo, al tratarse de datos generados y administrados por las entidades del sector público u organismos del Estado, también se consideró pertinente abordar la naturaleza y características de los mismos (por ejemplo, existen algunos datos geográficos sensibles), además de la legislación y normativa vigente.

El Grupo estuvo conformado por los representantes de las siguientes entidades y de sociedad civil:

- Vicepresidencia del Estado Plurinacional – Proyecto GeoBolivia.
- Agencia para el Desarrollo de la Sociedad de la Información en Bolivia (ADSIB).
- Ministerio de Gobierno.
- Ministerio de Economía y Finanzas Públicas (MEyFP).
- Autoridad de Supervisión del Sistema Financiero (ASFI).
- Servicio Nacional del Sistema de Reparto (SENASIR).

- Yacimientos Petrolíferos Fiscales Bolivianos (YPFB).
- Ministerio de Desarrollo Productivo y Economía Plural (MDPyEP).
- Empresa Pública Quipus.
- Ministerio de Transparencia Institucional y Lucha Contra la Corrupción.
- Gobierno Autónomo Departamental de La Paz (GADLP).
- Universidad Mayor de San Andrés (UMSA).
- Administración de Aeropuertos y Servicios Auxiliares a la Navegación Aérea (AASANA).
- Banco Unión S.A.
- Agencia de Gobierno Electrónico y Tecnologías de Información y Comunicación (AGETIC).
- Alejandro Salamanca (Comunidad Software Libre Bolivia).
- Henry Taby (Datos.bo).
- Javier Badani (La Pública).
- Judith Apaza (Sociedad civil).
- Luis Rejas (Comunidad de Software Libre de Bolivia).
- Miriam Jemio (Sociedad civil).
- Noemí Ticona Angles (Comunidad de Software Libre de Bolivia).
- Ricardo Trujillo (Sociedad civil).

- Sergio Ardaya (Sociedad civil).

En el mes de agosto de 2016, se realizó una reunión virtual en la que participó la Comunidad de Software Libre de Bolivia (CSLB) apoyando en el proceso de construcción de los lineamientos.

Asimismo, es importante resaltar que otras entidades u órganos del Estado participaron a través de sugerencias y acotaciones presentadas al documento inicial elaborado por el grupo. Entre estas entidades se encuentran:

- Ministerio de Defensa.
- Aduana Nacional de Bolivia (ANB).
- Servicio de Impuestos Nacionales (SIN).
- Autoridad de Fiscalización y Control de Pensiones y Seguros (APS).
- Autoridad de Impugnación Tributaria (AIT).
- Autoridad de Fiscalización del Juego (AJ).
- Banco Central de Bolivia (BCB).
- Unidad de Análisis de Políticas Sociales y Económicas (UDAPE).
- Instituto Nacional de Estadística (INE).
- Fondo Nacional de Inversión Productiva Social (FPS).
- Fondo Nacional de Desarrollo Regional (FNDR).

- Agencia Nacional de Hidrocarburos (ANH).
- Servicio de Desarrollo de Empresas Públicas Productivas (SEDEM).
- Empresa de Apoyo a la Producción de Alimentos (EMAPA).
- Fondo de Desarrollo del Sistema Financiero y Apoyo al Sector Productivo (FONDESIF).
- Administradora Boliviana de Carreteras (ABC).
- Empresa Estatal de Transporte por cable "Mi Teleférico".
- Empresa de Correos de Bolivia (ECOBOL).
- Autoridad de Regulación y Fiscalización de Telecomunicaciones y Transporte (ATT).
- Autoridad de Fiscalización y Control Social de Electricidad (AE).
- Autoridad Jurisdiccional Administrativa Minera (AJAM).
- Escuela de Gestión Pública Plurinacional (EGPP).
- Observatorio Plurinacional de Bolivia.
- Ministerio de Medio Ambiente y Agua (MMAyA)
- Observatorio Agroambiental y Productivo (OAP).
- Autoridad de Fiscalización y Control Social de Agua Potable y Saneamiento Básico (AAPS).
- Ministerio de Trabajo, Empleo y Previsión Social (MTEyPS).
- Servicio Estatal de Autonomías (SEA).
- Instituto Geográfico Militar (IGM).
- Tribunal Agroambiental.
- Gobierno Autónomo Departamental de Chuquisaca.
- Gobierno Autónomo Departamental de Santa Cruz.
- Gobierno Autónomo Departamental de Pando.
- Universidad Mayor, Real y Pontificia de San Francisco Xavier de Chuquisaca (USFX).
- Universidad Técnica de Oruro (UTO).
- Universidad Autónoma Mayor de San Simón de Cochabamba (UMSS).

Como resultado de las reuniones de trabajo del Grupo de Datos Abiertos, sus respectivas discusiones, exposiciones y el análisis de las observaciones contenidas en los informes y notas remitidas a la propuesta inicial, se elaboró el documento "Lineamientos para la adecuación y publicación de Datos Abiertos" para aquellas iniciativas que emprendan el proceso de apertura de datos.



# **Lineamientos** para la adecuación y publicación de **datos abiertos**



## 1. Marco normativo referencial

Conforme al Artículo 21 de la Constitución Política del Estado, Capítulo Tercero (Derechos Civiles y Políticos), Punto 6: las bolivianas y los bolivianos tienen el derecho a acceder a la información, interpretarla, analizarla y comunicarla libremente, de manera individual o colectiva. En este marco, uno de los mecanismos y procedimientos para garantizar el ejercicio de este derecho se da a través de la publicación y difusión de datos abiertos de las entidades u organismos del Estado. Además, como también lo establece el Artículo 232, Capítulo Cuarto (Servidoras Públicas y Servidores Públicos): la Administración Pública se rige por los principios de legitimidad, legalidad, imparcialidad, publicidad, compromiso e interés social, ética, transparencia, igualdad, competencia, eficiencia, calidad, calidez, honestidad, responsabilidad y resultados.

Asimismo, el Parágrafo I del Artículo 109 de la Constitución Política del Estado establece que todos los derechos reconocidos en la Constitución son directamente aplicables y gozan de iguales garantías para su protección.

El Artículo 1 del Decreto Supremo N.º 28168, de 17 de mayo de 2005, señala como objeto: "garantizar el acceso a la información como derecho fundamental de toda persona y la transparencia en la gestión del Poder Ejecutivo". Y el Artículo 2 se refiere al ámbito de su aplicación: "Poder Ejecutivo tanto a

nivel central como descentralizado, autárquico y desconcentrado; empresas y sociedades del Estado y sociedades con participación estatal mayoritaria".

El Punto 5 (Marco conceptual) de la Política Nacional de Transparencia y Lucha contra la Corrupción, aprobada mediante Decreto Supremo N.º 214, de 22 de julio de 2009, define el acceso a la información como "un derecho fundamental de las personas a conocer el manejo de la cosa pública", que "permite a los ciudadanos saber acerca del destino y uso de los recursos públicos, constituyéndose en un instrumento de participación ciudadana". Dicha Política establece en su Punto 10 que: "Resulta prioritario profundizar los mecanismos de transparencia en la gestión pública y fortalecer el derecho de todos los actores y organizaciones sociales de acceder a la información en poder del Estado, consagrando el principio general de publicidad de los actos del Órgano Ejecutivo, para lo cual es indispensable 'que las autoridades estatales se rijan por el principio de máxima divulgación, el cual establece la presunción de que toda la información es accesible, sujeto a un sistema restringido de excepciones' (Corte Interamericana de Derechos Humanos, sentencia de 19.09.2006, caso 12.108)".

El Numeral 1 del Parágrafo II del Artículo 4 de la Ley N.º 341 de 5 de febrero de 2013, Ley de Participación y Control Social, establece como principios esenciales a la Transparencia, entendida como el manejo honesto y adecuado de los

recursos públicos, así como la facilitación de información pública desde los órganos del Estado y las entidades privadas que administran recursos fiscales y/o recursos naturales, de forma veraz, oportuna, comprensible y confiable.

El Artículo N° 3 del Decreto Supremo 1793, de 13 de noviembre de 2013, Reglamento a la Ley N° 164, de 8 de agosto de 2011, para el Desarrollo de Tecnologías de Información y Comunicación, define Estándar Abierto "como una especificación técnica o protocolo normalizado":

- Cuyas especificaciones técnicas, completas y coherentes, están sujetas a una evaluación pública completa, se puede usar sin restricciones y está disponible por igual para todos los usuarios y/o partes, sin costo alguno para su uso;
- Que no necesita ningún componente o extensión adicional que tenga dependencias con formatos o protocolos que no cumplan la definición de Estándar Abierto;
- Que está libre de cláusulas legales o técnicas que limiten o restrinjan su utilización por cualquier usuario y/o parte o en cualquier modelo de negocio;
- Que es gestionado y puede ser desarrollado independientemente por cualquier organización en un proceso abierto a la participación equitativa e inclusiva de competidores,

usuarios, especialistas del área de aplicación y terceras partes;

- Que esté disponible en al menos una implementación completa, cuya documentación y especificación técnica está disponible para todas las partes con grado de detalles suficientes para un desarrollo correcto y de calidad.

En base al anterior contexto normativo referencial, es imperativo materializar el derecho fundamental que reconoce la Constitución respecto al acceso, interpretación, análisis y comunicación de la información. En este sentido, los datos públicos, con las excepciones que establece la legislación vigente, en resguardo de los derechos fundamentales y la seguridad del Estado, se constituyen en un recurso y una oportunidad para cumplir este derecho.

También se debe aclarar que habitualmente, en el derecho comparado, estos derechos relacionados a la información se hacen aplicables mediante el uso de licencias de obras culturales libres como son las licencias Creative Commons<sup>[1]</sup> y Open Data Commons<sup>[2]</sup>. Las licencias de obras culturales libres se compatibilizan con nuestra Constitución Política del Estado porque permiten acceder, utilizar, interpretar, crear derivados y distribuir copias del documento original solo con ciertas limitaciones. En este contexto, se sugiere una declaración respecto a las

[1] Ver: <https://creativecommons.org/>

[2] Ver <http://opendatacommons.org/>



libertades y condiciones para el uso de datos abiertos.

## 2. Objetivo

Este documento tiene el objetivo de establecer los lineamientos específicos para la adecuación y publicación digital de datos por las entidades del sector público u organismos del Estado, a fin de promover la transparencia, el acceso a información, la innovación y el control social.

## 3. Alcance y ámbito de aplicación

El presente documento comprende la conceptualización de datos abiertos, características, estándares de datos abiertos, protocolos de preprocesamiento y licencia para publicación y uso de datos abiertos, para posibilitar la adecuación y publicación digital en formato de datos abiertos de los conjuntos de datos que las entidades del sector público u organismos del Estado hayan publicado o que vayan a poner a disposición de la ciudadanía<sup>[3]</sup>, conforme a legislación y normativa vigente.

Los lineamientos contenidos en este documento deberán ser admitidos por todas las entidades del sector público u organismos del Estado, sin perjuicio del trabajo desarrollado por aquellas que ya hayan asumido esta tarea en el marco de

los parámetros rectores, normas y estándares nacionales e internacionales vigentes o diferentes.

## 4. Conceptualización de datos abiertos

Datos abiertos son los datos accesibles, estandarizados y reutilizables que las entidades del sector público u organismos del Estado generan, administran o custodian y que pueden ser obtenidos de forma libre y sin restricciones, con el fin de que los distintos actores puedan utilizarlos.

## 5. Características

Se deben considerar las siguientes características cuando se desarrollen iniciativas de apertura de datos:

### 5.1. Estructurados

Se debe garantizar la lectura, reutilización y automatización de los conjuntos de datos a través del uso de formatos legibles por computadores (estructurados), lo que facilitará la agregación de otros conjuntos de datos, permitirá una mejor búsqueda y tratamiento y potenciará el grado de acceso y uso de los conjuntos de datos. En todos los casos, se deben utilizar formatos abiertos<sup>[4]</sup> así como la codificación estándar de caracteres.

[3] En este contexto, no corresponde a este documento establecer qué datos son publicables o no, sino el cómo se debe proceder para su adecuación y/o publicación.

[4] En el Capítulo 6 de este documento se especifican los formatos que se deben utilizar para el procesamiento y publicación de datos abiertos, siguiendo la definición de Estándar Abierto del Decreto Supremo 1793, de 13 de noviembre de 2013, Reglamento para el Desarrollo de Tecnologías de Información y Comunicación.

## 5.2. Detallados

Los conjuntos de datos deben ofrecer el mayor nivel de detalle posible a fin de posibilitar un análisis profundo y desagregado por parte de los usuarios (mayor granularidad). Asimismo, se debe procurar que el conjunto de datos contenga la mayor cantidad de atributos disponibles.

## 5.3. Con metadatos<sup>[5]</sup>

Los conjuntos de datos a ser publicados deben poseer información y datos que contextualicen al usuario sobre su origen y contenido.

## 5.4. Disponibles y permanentes

A fin de garantizar la disponibilidad de los datos, se debe prever que tras la publicación de un conjunto de datos, este permanezca accesible en el tiempo.

## 5.5. Oportunos

Se deben evitar periodos considerables de tiempo entre la producción de un conjunto de datos y su publicación. De esta manera, se garantiza su valor y relevancia.

## 5.6. No discriminatorios

Se debe otorgar al usuario tanto el derecho a reutilizar el conjunto de datos<sup>[6]</sup> así como su libre acceso. De esta manera,

[5] Los metadatos son elementos descriptivos que otorgan al usuario información necesaria para entender y procesar un conjunto de datos. Así, los metadatos describen el contenido, la calidad, la condición y otras características de los datos, dependiendo del modelo que se decida utilizar.

[6] El apartado 8 de este documento otorga una Declaración de Uso de Datos Abiertos equivalente a las licencias de obras culturales libres (copyleft) que debe ser incorporada a las iniciativas de datos abiertos.

también se asegura la gratuidad a la hora de acceder a estos datos.

## 5.7. Veracidad de los datos

Todo dato abierto debe provenir de una fuente veraz y verificable a través de su publicador.

## 6. Estándares de datos abiertos

### 6.1. Formatos

Con respecto a los formatos de los conjuntos de datos, estos se pueden agrupar en cuatro grandes categorías: uno genérico y tres especializados, cuyas características técnicas exigen un tratamiento particular que se describe a continuación. Estos formatos han sido seleccionados en aplicación a la definición de Estándar Abierto del Decreto Supremo 1793: Reglamento para el Desarrollo de Tecnologías de Información y Comunicación.

#### a) Datos genéricos

Son aquellos datos que no tienen aplicaciones especializadas y que normalmente corresponden a datos provenientes de bases de datos, reportes, como planillas de cálculo o archivos de texto. Las características de los formatos a usar son:

Para datos de planillas y bases de datos tradicionales se debe considerar el uso de los formatos JSON, SQL, CSV, XML (XSD asociado).

Para conjuntos de datos científicos (grandes volúmenes), se debe considerar el uso de los formatos HDF o NetCDF.

Es posible considerar otros formatos similares siempre y cuando cumplan con la definición de Estándar Abierto establecida en Decreto Supremo N° 1793.

## b) Datos estadísticos

Son datos usados para fines de análisis estadístico. Como formato de datos, se deben publicar los conjuntos de datos en formato CSV<sup>[7]</sup>.

## c) Imágenes

Corresponden a datos gráficos cuyos formatos recomendados son los siguientes:

Se deben usar tipos de archivos de imágenes basados en el método JPEG de compresión con pérdida y los archivos basados en el método PNG (Portable Network Graphics, por sus siglas en inglés) de compresión sin pérdida, usualmente con extensión.

Para formatos vectoriales y de visualización se recomienda el formato SVG (Scalable Vector Graphics, por sus siglas en inglés).

## d) Datos geográficos

Son aquellos que representan la superficie terrestre o la geometría de objetos en el espacio geográfico. Los formatos de publicación de coberturas y datos georeferenciados deberán estar en concordancia con los formatos definidos

por la Norma Nacional de Generación, Intercambio y Difusión de Información Geográfica del Estado Plurinacional de Bolivia (NIG)<sup>[8]</sup>.

Como recomendación de base, los formatos de publicación pueden variar dependiendo del tipo de uso que tengan los datos, distinguiendo entre formatos orientados a los objetos (vectoriales) u orientados a la superficie (raster o malla).

Para visualización de imágenes, los formatos recomendados son JPEG y PNG.

Para uso en servidores de mapas y en análisis se recomienda el uso de archivos World File con sus seis líneas de metadato y la convención de usar el mismo nombre del archivo raster con el sufijo "w".

Para formatos vectoriales y de visualización en la Web se recomienda utilizar el formato PNG.

Para geoservicios de transferencia de datos del tipo Web Feature Service (WFS) y Web Coverage Service (WCS) se recomiendan los formatos GML<sup>[9]</sup>, GPX, GeoJSON y KML 2.2 (OGC)<sup>[10]</sup>.

El siguiente cuadro resume los tipos de datos y formatos recomendados.

[8] Esta norma ha sido trabajada y consensuada por representantes de entidades públicas en el comité: Infraestructura de Datos Espaciales del Estado Plurinacional de Bolivia (IDE-EPB), que se viene reuniendo desde 2012 (<https://ideepb.geo.gob.bo/>).

[9] Basado en el ISO 19136.

[10] Las especificaciones KML 2.2 fueron presentadas ante el Open Geospatial Consortium (OGC) con el fin de alcanzar un estatus de estándar abierto como formato de intercambio de información geográfica. En noviembre de 2007, la OGC creó un nuevo grupo de trabajo sobre KML 2.2 donde se solicitaron alegaciones hasta el 4 de enero de 2008. El 14 de abril de 2008, el formato KML 2.2 se convirtió definitivamente en un estándar abierto.

[7] Específicamente se recomienda el CSV: RFC 4180.

TIPO DE DATOS	FORMATO REQUERIDO
Genéricos	JSON, SQL, CSV, XML, HDF, NetCDF
Estadísticos	CSV
Imágenes	JPEG, PNG, SVG
Geográficos	World File, GML, GPX, GeoJSON, KML 2.2

El conjunto de datos debe ser publicado siempre en estándar abierto. En caso de que la fuente de datos no sea de formato abierto, se deberá publicar dicha fuente.

### 6.1.1. Sobre el formato PDF

En el país se ha extendido el uso del formato PDF para la publicación de varios documentos y datos públicos. Si bien se trata de un formato de estándar abierto, la publicación en este formato impide la reutilización y procesamiento de sus datos desde ordenadores, por lo que se recomienda, para la publicación en versiones posteriores de conjuntos de datos, utilizar los formatos recomendados en el punto 6.1 de este documento.

### 6.2. Estructura del metadato

De acuerdo a la DCMI (*Dublin Core Metadata Initiative, por sus siglas en inglés*), un metadato debe contener la siguiente información:

**Título:** El nombre dado a un recurso, habitualmente por el autor. **(OBLIGATORIO)**

Etiqueta: *DC. Title.*

**Claves:** Los temas del recurso. Típicamente *Subject* expresará las claves o frases que describen el título o el contenido del recurso. Se fomentará el uso de vocabularios controlados y de sistemas de clasificación formales.

Etiqueta: *DC. Subject.*

**Descripción:** Es una descripción textual del recurso. Puede ser un resumen, en el caso de un documento; o una descripción del contenido, en el caso de un conjunto de datos o base de datos. En este espacio también se debe especificar el carácter temporal o permanente del conjunto de datos, además, si corresponde, las fórmulas a utilizar o utilizadas en la obtención de los datos o su tratamiento. En el caso de que los datos sean preliminares, se debe aclarar que estos son de carácter referencial y no pueden ser utilizados como información oficial. En caso de que los datos se modifiquen, se deberá especificar la modificación y origen de los datos sujetos a modificación. **(OBLIGATORIO)**

Etiqueta: *DC. Description.*

**Fuente:** Secuencia de caracteres usados para identificar unívocamente un trabajo a partir del cual proviene el recurso actual. **(OBLIGATORIO)**

Etiqueta: *DC. Source.*

**Tipo del Recurso:** Categoría del recurso. Por ejemplo, página institucional.

Etiqueta: *DC.Type.*

**Relación:** Identificador de un segundo recurso y su relación con el recurso actual. Este elemento permite enlazar los recursos relacionados y las descripciones de los recursos.

Etiqueta: *DC. Relation.*

**Cobertura:** Característica de cobertura espacial o temporal del contenido intelectual del recurso. La cobertura espacial se refiere a una región física, utilizando por ejemplo coordenadas. La cobertura temporal se refiere al contenido del recurso, no a cuándo fue creado (*que ya lo encontramos en el elemento Date*). **(OBLIGATORIO)**

Etiqueta: *DC. Coverage.*

**Autor o creador:** La persona u organización responsable de la creación del contenido intelectual del recurso. **(OBLIGATORIO)**

Etiqueta: *DC. Creator.*

**Editor:** Entidad responsable de hacer que el recurso se encuentre disponible en la red en su formato actual. **(OBLIGATORIO)**

Etiqueta: *DC. Publisher.*

**Otros colaboradores:** Persona u organización que haya tenido una contribución intelectual significativa, pero que esta sea secundaria en comparación con las de las personas u organizaciones especificadas en el elemento *Creator*. (por ejemplo: *Editor*).

Etiqueta: *DC. Contributor.*

**Derechos:** Referencia (por ejemplo una URL) de una nota sobre la licen-

cia para un servicio de gestión de derechos o para un servicio que dará información sobre libertades y condiciones de acceso y uso de un recurso. En este campo se incluirá la Declaración de Publicación de Datos Abiertos especificada en el punto 8 de este documento. **(OBLIGATORIO)**

Etiqueta: *DC. Rights.*

**Fecha:** Fecha en la cual el recurso se puso a disposición del usuario en su forma actual. Esta fecha no debe confundirse con la que pertenece al elemento *Coverage*, que está asociada con el recurso en la medida en que el contenido intelectual está de alguna manera relacionado con aquella fecha. Asimismo, cuando corresponda se deben especificar la fecha de modificación y origen de los datos. **(OBLIGATORIO)**

Etiqueta: *DC. Date.*

**Formato:** Formato de datos de un recurso usado para identificar el software y, posiblemente, el hardware que se necesitaría para mostrar el recurso. **(OBLIGATORIO)**

Etiqueta: *DC. Format.*

**Identificador (único) del recurso:** Secuencia de caracteres utilizados para identificar unívocamente un recurso. Ejemplos para recursos en línea pueden ser las URL y URN. **(OBLIGATORIO)**

Etiqueta: *DC. Identifier.*

**Lengua:** Lengua/s del contenido intelectual del recurso. **(OBLIGATORIO)**

Etiqueta: *OD. Language.*

**Diccionario de Datos:** Listado organizado de nombres y definiciones que contiene un conjunto de datos para que el usuario tenga entendimiento común del contenido del conjunto de datos. **(OBLIGATORIO)**

Etiqueta: *OD. Dictionary.*

Si bien se recomienda esta estructura para documentar un conjunto de datos, se puede utilizar una versión mínima de metadato aplicando aquellas etiquetas marcadas con el rótulo OBLIGATORIO. Asimismo, si algún proyecto o iniciativa de datos abiertos lo desea, puede utilizar versiones más amplias a la estructura de metadato recomendada en este documento.

## 7. Protocolos de preprocesamiento

Los protocolos de preprocesamiento son necesarios para facilitar el intercambio, la organización y el procesamiento de los datos. Para llevar a cabo este trabajo se recomienda tener en cuenta los siguientes criterios:

### 7.1. Codificación estándar de caracteres

A efectos de optimizar el trabajo con datos y evitar los caracteres incompatibles en documentos de texto, conjuntos de datos, tablas y otros formatos de datos donde se manejan caracteres especiales que son propios de los idiomas oficiales de nuestro país, se debe utilizar la codificación universal UTF-8 sin BOM.

### 7.2. Encabezado de tablas y encabezados de los conjuntos de datos

Para garantizar la reutilización de datos en formato de tabla o base de datos, se debe considerar el tipo de caracteres y la forma de escritura de los encabezados o nombre de los atributos (columnas), para lo cual se sugiere usar solo minúsculas de la a...z, números de 0...9 y guión bajo (" \_ "); se excluyen la ñ y las vocales con tilde o diéresis.

### 7.3. Título de archivos generados

Considerando que los generadores de datos almacenan sus productos en archivos digitales para su reutilización, se debe tener en cuenta que el nombre del archivo digital debe describir, de manera muy resumida, la fuente y el contenido del mismo. Para tal efecto, se sugiere denominar el nombre del archivo de la siguiente manera:

Nombre o sigla de la entidad que genera el dato.

El contenido del archivo (éste debe ser descriptivo y pensarse intuitivamente para el público que accederá al archivo).

Fecha o gestión en la que se generó el dato.

El título del archivo debe ser separado por guión bajo " \_ " o guión medio " - "

Ejemplo:

a) agetic\_localizacion\_entidades\_estado\_2015.csv

## 7.4. Delimitación numérica para millares y decimales

Considerando que gran parte de los datos abiertos contienen valores numéricos, ya sea en formato de tablas o bases de datos, se debe tener en cuenta que en nuestro país la separación de decimales y los millares no están estandarizadas ni en su uso ni en su generación. Por lo tanto, para garantizar la interrelación e interoperabilidad de datos numéricos, se establece que la separación de decimales se efectúe mediante punto (.), tal como se presenta en el siguiente ejemplo:

8425366.88

## 7.5. Formato de fechas y horas (24 horas)

Para garantizar la interrelación e interoperabilidad de datos con respecto a las fechas, se establece utilizar el formato: AAAA-MM-DD, donde AAAA es año de cuatro dígitos; MM es mes de dos dígitos; y DD es día de dos dígitos.

## 7.6 Compresión y empaquetamiento de archivos

Se recomienda usar algoritmos de compresión y empaquetamiento sin pérdidas y abiertos (no propietarios), como por ejemplo: tar, bzip2, zip, etc.

## 8. Licencia para la publicación y uso de datos abiertos

Actualmente, en Bolivia no se cuenta con una licencia compatible con las

licencias de obras culturales libres. No obstante, no se puede limitar la aplicación de un derecho Constitucional a la falta de un desarrollo normativo. En ese sentido, se realiza la presente declaración que propone otorgar permisos para el acceso, interpretación, análisis y comunicación de los datos abiertos e información pública bajo los siguientes requisitos:

### 8.1. Declaración de uso de Datos Abiertos

La presente declaración promueve el uso y reutilización de los conjuntos de datos abiertos bajo las siguientes libertades y condiciones:

I) Usted es libre y responsable de:

- a) Hacer, publicar y distribuir copias del conjunto de datos;
- b) Adaptar, remezclar y/o transformar el conjunto de datos;
- c) Extraer total o parcialmente el contenido del conjunto de datos;
- d) Crear conjuntos de datos derivados del conjunto de datos o su contenido;

II) Siempre y cuando cumpla con las siguientes condiciones:

- a) Citar la fuente de origen de donde obtuvo el conjunto de datos<sup>[11]</sup>,

[11] Los conjuntos de datos deben ser citados de la siguiente manera: Sigla o nombre completo de la entidad generadora del conjunto de datos (año de publicación de la base de datos). Nombre del conjunto de datos. Base de datos disponible en: enlace de Internet de los datos descargados (fecha de consulta: día, mes y año).  
Ejemplo de cita: Ministerio de Educación (2014). Estudiantes Matriculados 2014. Base de datos disponible en <https://datos.gob.bo/dataset/estudiantes-matriculados-2014> (fecha de consulta: 30 de julio de 2016).

**b)** En el caso de copiar, distribuir, re-mezclar, transformar y crear a partir de un conjunto de datos y sus resultados, se mantengan la mismas libertades y condiciones de la presente Declaración.

**III)** Los portales e iniciativas de datos abiertos deben hacer uso de la presente Declaración de uso de datos abiertos que garantiza la atribución de la autoría de los conjuntos de datos y vela por el carácter abierto de los mismos.

**IV)** Toda interpretación no es responsabilidad de la entidad que publica los datos.

## Anexo 1

### Glosario

**Apertura de datos:** acción y resultado de hacer accesibles públicamente datos ocultos o restringidos en una organización presentándolos de forma que puedan ser aprovechados por cualquier agente reutilizador.

**Conjunto de datos:** Datos relacionados, convenientemente estructurados y organizados, de forma que puedan ser tratados (procesados) apropiadamente para obtener información.

**CSV:** *Comma separated values.*

**CTIC-EPB:** Consejo para las Tecnologías de Información y Comunicación del Estado Plurinacional de Bolivia.

**Datos:** Caracteres, números o símbolos recogidos para su tratamiento informático,

análisis estadístico o referencia. Información no procesada.

**Diccionario de datos:** Es un listado organizado de nombres y definiciones que contiene un conjunto de datos, para que el usuario tenga entendimiento común del contenido del conjunto de datos.

**Dublin Core:** Modelo de metadatos elaborado y auspiciado por la DCMI (*Dublin Core Metadata Initiative*), una organización dedicada a fomentar la adopción extensa de los estándares interoperables de los metadatos.

**Formato:** Conjunto de características técnicas y de presentación.

**GML:** *Geography Markup Language.*

**GPX:** *GPS Exchange Format.*

**HDF:** *Hierarchical Data Format.*

**ISO:** *International Organization for Standardization.*

**JPEG:** *Joint Photographic Experts Group.*

**JSON:** *JavaScript Object Notation.*

**KML:** *Keyhole Markup Language.*

**Metadato:** Datos que describen otros datos, es decir, son el conjunto de características que todo conjunto de datos lleva asociado.

**NetCDF:** *Network Common Data Form.*

**PDF:** *Portable Document Format;* formato portable de documentos estándar



para la publicación de documentos en la web.

**PNG:** *Portable Network Graphics.*

**Reutilizar:** Utilizar algo con la función que desempeñaba anteriormente o con otros fines.

**SVG:** *Scalable Vector Graphics.*

**URL:** *Uniform Resource Locator.*

**UTF:** *Unicode Transformation Format.*

**UTF-8:** UTF-8 (8-bit Unicode Transformation Format). Es un formato de codificación de caracteres.

**XML:** *Extensible Markup Language.*

## Anexo 2

### Ejemplo de metadato

#### Contenido:

**Título:** Proyectos del programa MI AGUA I, 2011-2015 (**OBLIGATORIO**)

Etiqueta: *DC. Title.*

**Claves:** agua, agua potable, municipio, riego

Etiqueta: *DC. Subject.*

**Descripción:** Tabla de registro de proyectos de agua potable y riego del Programa MI AGUA I a nivel nacional; desglosado por departamentos y municipios, 2011-2015. El objetivo del Programa MI AGUA es apoyar las estrategias de reducción de la pobreza y el logro de la soberanía alimentaria del país a través

del incremento de la capacidad productiva de pequeños y medianos productores agrícolas y el de contribuir a mejorar las condiciones de vida y salud de la población a través del incremento del acceso y de la calidad en la provisión del servicio de agua potable y saneamiento básico en el territorio nacional. La población objetivo de los proyectos son familias de comunidades campesinas, agricultores, población rural dispersa y residentes más pobres de las áreas urbanas que no disponen de agua potable y saneamiento básico de manera formal y continua. Los responsables del Programa MI AGUA I como co-ejecutores son el Ministerio de Medio Ambiente y Agua (MMAyA) y el Fondo Nacional de Inversión Productiva y Social (FPS). Con relación al presupuesto, éste está compuesto por la inversión del Organismo Financiador - Corporación Andina de Fomento (CAF) y la contraparte del Municipio (aporte local); 82% y 18%, respectivamente.

Etiqueta: *DC. Description.*

**Fuente:** Fondo Nacional de Inversión Productiva y Social (FPS) (**OBLIGATORIO**)

Etiqueta: *DC. Source.*

**Tipo del Recurso:** <http://www.fps.gob.bo/>

Etiqueta: *DC. Type.*

**Relación:** Industria y Servicios.

Etiqueta: *DC. Relation.*

**Cobertura:** Gestión de inicio 2011 y gestión de conclusión 2015. (**OBLIGATORIO**)

Etiqueta: *DC. Coverage.*

**Autor o Creador:** Fondo Nacional de Inversión Productiva y Social (Lic. Mónica Vargas) (OBLIGATORIO)

Etiqueta: *DC. Creator.*

**Editor:** Fondo Nacional de Inversión Productivo y Social (OBLIGATORIO)

Etiqueta: *DC. Publisher.*

**Otros Colaboradores:** Ministerio de Medio Ambiente y Agua – Ministerio de Planificación

Etiqueta: *DC. Contributor.*

**Derechos:** Declaración de Publicación de Datos Abiertos respecto al punto 8 de este documento. (OBLIGATORIO)

Etiqueta: *DC. Rights.*

**Fecha:** 24 de septiembre de 2016 (OBLIGATORIO)

Etiqueta: *DC. Date.*

**Formato:** CSV. (OBLIGATORIO)

Etiqueta: *DC. Format.*

**Identificador (único) del recurso:** <https://datos.gob.bo/dataset/673806b3-e5d4-4f9d-b85e-e773a1e448e9> (OBLIGATORIO).

Etiqueta: *DC. Identifier.*

**Lengua:** Castellano. (OBLIGATORIO)

Etiqueta: *OD. Language.*

**Diccionario de Datos:** Lid=identificador; cod\_mun=código municipal; tipo\_proy=tipo de proyecto; nomb\_proy=nombre del proyecto; monto\_

fps\_en\_bs=monto en bolivianos del FPS; monto\_contrap\_en\_bs=monto en bolivianos de la contraparte (aporte local); monto\_total\_en\_bs=monto total en bolivianos; estado\_proy=estado del proyecto; conex\_dom=conexiones domiciliarias de agua potable; pil\_publ=piletas públicas; hect\_increm=hectáreas incrementales por riego; hog\_benef=hogares beneficiados (OBLIGATORIO).

Etiqueta: *OD. Dictionary.*



