

DATOS ABIERTOS PARA EL DESARROLLO SOCIO-ECONÓMICO

¿QUÉ SON LOS DATOS ABIERTOS?

Los Datos Abiertos, son datos disponibles en línea, sin procesar, en formato abierto, neutral e interoperable; que permite su uso y reuso, disponible para su descarga en forma completa, sin costo ni requisitos de registro y procesable en computadora. Deberá cumplir con las siguientes características:

- a. **Completo:** todo dato público debe estar disponible sin más excepciones que las establecidas por la ley.
- b. **Primario:** el dato público debe ser puesto a disposición desde la fuente, con el mayor nivel posible de desagregación y sin modificar.
- c. **Actualizado y oportuno:** el dato público debe ser puesto a disposición en forma pronta, según su periodo de generación para preservar su valor de reuso.
- d. **Accesible:** el dato público debe encontrarse fácilmente disponible para todas las personas y para cualquier uso no prohibido por ley.
- e. **Procesable automáticamente:** el dato público debe estar estructurado para permitir el procesamiento automático por computadoras.
- f. **No discriminatorio:** El dato debe estar disponible para cualquier persona sin requisitos de registro previo.
- g. **No propietario:** el dato público debe estar disponible en formatos neutrales e interoperables.
- h. **De libre uso y reuso (conocido como licencia de uso y reuso de datos):** El dato público debe estar libre de cualquier restricción de derechos de autor, propiedad intelectual o industrial, o de información comercial no divulgada. Puede tener restricciones razonables de privacidad o por niveles de usuario, acorde a la ley.

DATOS ABIERTOS COMO MEDIO EFICIENTE DE ACCESO A LA INFORMACIÓN PÚBLICA

- ! Cualquier persona podrá utilizar datos ya publicados o incluso solicitar a la institución respectiva la liberación de un conjunto de datos de su interés, en formato abierto, neutral e interoperable.

La idea es que se establezca un mecanismo digital, ágil y no discriminatorio para que las instituciones dispongan y publiquen datos.

PRINCIPIOS DE LOS DATOS ABIERTOS

1. **Abiertos por defecto:** Todos los datos públicos serán dispuestos en formatos abiertos estructurados, con la capacidad de ser procesados e interpretados por equipos electrónicos de manera automática. Deberán realizarse las acciones necesarias para que esta condición se de desde el momento en que los datos son generados o recolectados.
2. **Oportunos y Exhaustivos:** Los datos públicos abiertos se pondrán a disposición en forma oportuna, completa, precisa y exhaustiva. Además, se mantendrán versiones históricas de los datos para garantizar su permanencia.
3. **Accesibles y Utilizables:** Los datos públicos abiertos deben ser fácilmente visibles y accesibles, poniéndose a disposición de forma gratuita y sin barreras.
4. **Comparables, Neutrales e Interoperables:** los datos públicos abiertos deben estar disponibles en su fuente primaria, con el más alto grado posible de desagregación, o hacer referencia a la base principal cuando se disponga de forma agregada; deben ser fáciles de comparar dentro y entre sectores, a través de localizaciones geográficas y del tiempo. Además deben ser presentados en formatos estructurados y armonizados que faciliten su comparación, intercambio, trazabilidad y reutilización efectiva.
5. **Mecanismo para fortalecer la Gobernanza y la Participación Ciudadana:** Los datos públicos abiertos deberán fortalecer la confianza en las instituciones públicas y la gobernanza, entendida esta última como la gestión institucional y su relación con los actores sociales. De manera tal que provean un fundamento para la transparencia y la rendición de cuentas que mejore la toma de decisiones, incremente la demanda y prestación de servicios públicos; así como la formulación y evaluación de programas y políticas para satisfacer las necesidades de la sociedad.

6. Fomento del desarrollo incluyente y la innovación: La disponibilidad de datos abiertos debe estimular la creatividad y la innovación, generando procesos y espacios para identificar desafíos sociales y económicos, así como el monitoreo de programas de desarrollo sostenible con participación de todos los sectores involucrados.

NIVELES DE APERTURA

Existen diferentes escalas y parámetros para definir el nivel o grado de apertura de los datos. Tim Berners-Lee, creador de la Web y pionero de los Datos Enlazados (Linked Data), sugirió un esquema de clasificación a partir de 5 estrellas, cada una de las cuales representa un nivel o escalón de desarrollo en la apertura de los datos, del más simple al más complejo.

NIVEL	INFORMACIÓN	ACCESO	SERVICIOS
★ ★ ★ ★ ★	Datos disponibles a través de una licencia abierta e incluidos en documentos que sólo pueden ser leídos e interpretados gracias a una intervención humana significativa (por ejemplo, PDF).	Principalmente humano, también necesario para dar sentido a los datos incluidos en los documentos	Sólo las intervenciones humanas relevantes para la extracción y procesamiento de posibles datos permiten el desarrollo de servicios con la información disponible a este nivel.

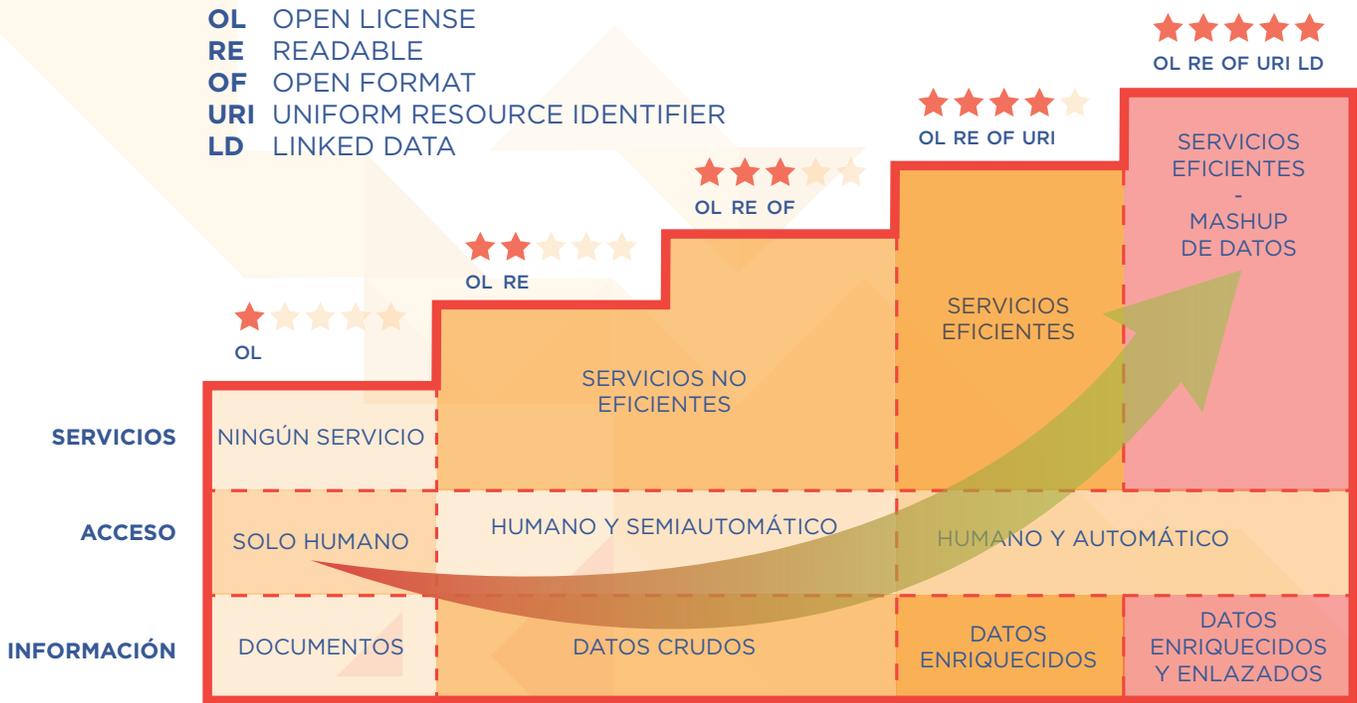
NIVEL	INFORMACIÓN	ACCESO	SERVICIOS
★ ★ ★ ★	Datos disponibles en forma estructurada y con licencia abierta. Sin embargo, los formatos son de propiedad exclusiva (por ejemplo, Excel) y la intervención humana es muy necesaria para el procesamiento de datos.	Los programas pueden procesar datos pero no son capaces de interpretarlos; por lo tanto, es necesaria la intervención humana para escribir programas ad-hoc para su uso.	Servicios ad-hoc que deben incorporar los datos para permitir el acceso directo a los mismos a través de la Web.

NIVEL	INFORMACIÓN	ACCESO	SERVICIOS
	<p>Datos con características de la capa anterior pero en un formato no propietario (por ejemplo, CSV, JSON, geoJSON). Los datos son legibles por un programa, pero es necesaria la intervención humana para su procesamiento.</p>	<p>Los programas pueden procesar datos pero no son capaces de interpretarlos, por lo que es necesaria la intervención humana para escribir programas ad-hoc para su uso.</p>	<p>Servicios ad-hoc que deben incorporar los datos para permitir el acceso directo a los mismos a través de la Web.</p>

NIVEL	INFORMACIÓN	ACCESO	SERVICIOS
	<p>Datos con características de la capa anterior pero expuestos utilizando estándares W3C como RDF y SPARQL. Los datos se describen semánticamente a través de metadatos y ontologías.</p>	<p>Los programas son capaces de conocer la ontología de referencia y por lo tanto de procesar los datos casi sin intervención humana adicional.</p>	<p>Servicios, también para dispositivos móviles, que utilizan el acceso directo a la Web para encontrar los datos de interés.</p>

NIVEL	INFORMACIÓN	ACCESO	SERVICIOS
	<p>Datos con características del nivel anterior pero vinculados a los expuestos por otras personas y organizaciones (es decir, Linked Open Data 2). Los datos se describen semánticamente mediante metadatos y ontologías. Siguen el paradigma del RDF. En consecuencia, esta URI puede utilizarse para acceder directamente a la información sobre esa entidad. Los datos se denominan "enlazados" debido a la posibilidad de referirse (es decir, "conectarse") entre sí. En el caso de los datos abiertos enlazados, la intervención humana puede minimizarse y a veces incluso eliminarse.</p>	<p>Los programas son capaces de conocer la ontología de referencia y por lo tanto de procesar los datos casi sin intervención humana adicional.</p>	<p>Servicios, también para dispositivos móviles, que explotan tanto el acceso directo a la Web como la información adicional capturada a través de los enlaces de los datos de interés, facilitando la mezcla de datos.</p>

Este cuadro tiene la funcionalidad de diagnosticar el nivel de apertura de datos con los que cuenta cada institución y provee grados alcanzables para llegar a niveles cada vez más refinados de apertura de datos. Así mismo, esta guía recomienda un camino gradual hacia la producción nativa de Datos Abiertos Enlazados - LOD (nivel cinco estrellas), comenzando desde el nivel 3. (Ver figura)



¿POR QUÉ ABRIR DATOS PÚBLICOS Y CONVERTIRLOS EN DATOS ABIERTOS?

El sector público, los privados, las organizaciones de la sociedad civil, las universidades y/o cualquier ciudadano, puedan reutilizar los datos, agregarles valor, generar información y crear nuevos o mejores servicios y aplicaciones.

EL VALOR DE LOS DATOS ABIERTOS...

La generación de valor por medio de los datos requiere la participación y compromiso de todos los actores, pues es fundamental para entender las demandas de datos, guiar su entrega, conocer las principales fuentes y establecer el trabajo que debe realizarse para ponerlos a disposición del público.

Sin embargo, lograr que la información quede a disposición del público no es suficiente para crear el valor mencionado. La creación de valores económicos, sociales o de buena gobernanza requiere, en efecto, de la activa participación pública.

La OCDE identifica tres amplios conjuntos de valores que los gobiernos focalizan mediante iniciativas de datos abiertos: valor económico, valor social y valor de buena gobernanza. Los beneficios de los datos abiertos incluyen, entre otros, los siguientes:

- ▲ Una mejor toma de decisiones para los sectores público, privado y social.
- ▲ Una prestación de servicios públicos más rápida, innovadora y eficiente.
- ▲ Una mayor rendición de cuentas del gobierno y una mayor confianza pública.
- ▲ Empoderamiento de la población y la sociedad civil.
- ▲ Una mayor participación y compromiso del público.
- ▲ Espíritu emprendedor e innovación social.

GLOSARIO

- ▲ **Dato:** representación simbólica (numérica, alfabética, algorítmica, espacial, etc.) de un atributo o variable cuantitativa o cualitativa.
- ▲ **Dato público:** cualquier tipo de dato de interés público que sea generado o resguardado por quien ejerza una función o potestad pública y que no tenga su acceso restringido por ley.
- ▲ **Datos personales:** cualquier dato relativo a una persona física identificada o identificable.
- ▲ **Datos personales de acceso irrestricto:** los contenidos en bases de datos públicas de acceso general, según dispongan leyes especiales y de conformidad con la finalidad para la cual estos datos fueron recabados.
- ▲ **Datos personales de acceso restringido:** los que, aun formando parte de registros de acceso al público, no son de acceso irrestricto por ser de interés solo para su titular o para la Administración Pública.
- ▲ **Datos sensibles:** información relativa al fuero íntimo de la persona, como por ejemplo los que revelen origen racial, opiniones políticas, convicciones religiosas o espirituales, condición socioeconómica, información biomédica o genética, vida y orientación sexual, entre otros.
- ▲ **Conjunto de Datos:** la serie de datos estructurados, vinculados entre sí y agrupados dentro de una misma unidad temática, de forma que puedan ser procesados apropiadamente para obtener información.
- ▲ **Catálogo de Datos Abiertos:** el inventario de los conjuntos de datos puestos a disposición en un portal de datos abiertos disponible en la web.
- ▲ **Formato Abierto, Neutral e Interoperable:** Es el formato de archivo que no supone una dificultad de acceso, que está libre de cualquier restricción legal o económica para su utilización y que tiene la capacidad de integrarse y entrelazarse con otros datos.
- ▲ **Ecosistema:** Conjunto de actores institucionales, del sector empresarial, la academia y la ciudadanía que interactúan en los procesos de apertura y uso de los datos abiertos.

- ▲ **Armonización:** aplicación de estructuras comunes, previamente definidas y acordadas para facilitar la integración de conjuntos de datos.
- ▲ **Metadatos:** son los datos estructurados y actualizados que describen el contexto y las características de atribución, contenido, captura, procesamiento, calidad, condición, acceso, distribución u otras de un conjunto de datos; que faciliten su búsqueda, identificación y uso.
- ▲ **Portal de Datos Abiertos:** Plataforma digital, neutral e interoperable que sirven para almacenar, compartir y conectar los catálogos de datos y todos los recursos destinados a facilitar su uso, tales como: aplicaciones, herramientas de visualización, demos, normativa o guías.
- ▲ **Licencia para uso y reuso datos abiertos:** documento que especifica el derecho de uso de los datos públicos abiertos. Establece y garantiza condiciones para el acceso, la redistribución, el reuso, la ausencia de restricciones tecnológicas, la atribución de origen, la no discriminación de grupos o personas y la integridad de los datos.
- ▲ **Bases de Datos:** conjunto de datos pertenecientes a un mismo contexto y almacenados sistemáticamente para su posterior uso
- ▲ **Interoperabilidad tecnológica:** entendida esta como la capacidad de generar la interconexión e interacción funcional o la capacidad para intercambiar información con otros sistemas a través de sus interfaces