



El futuro digital
es de todos

MinTIC

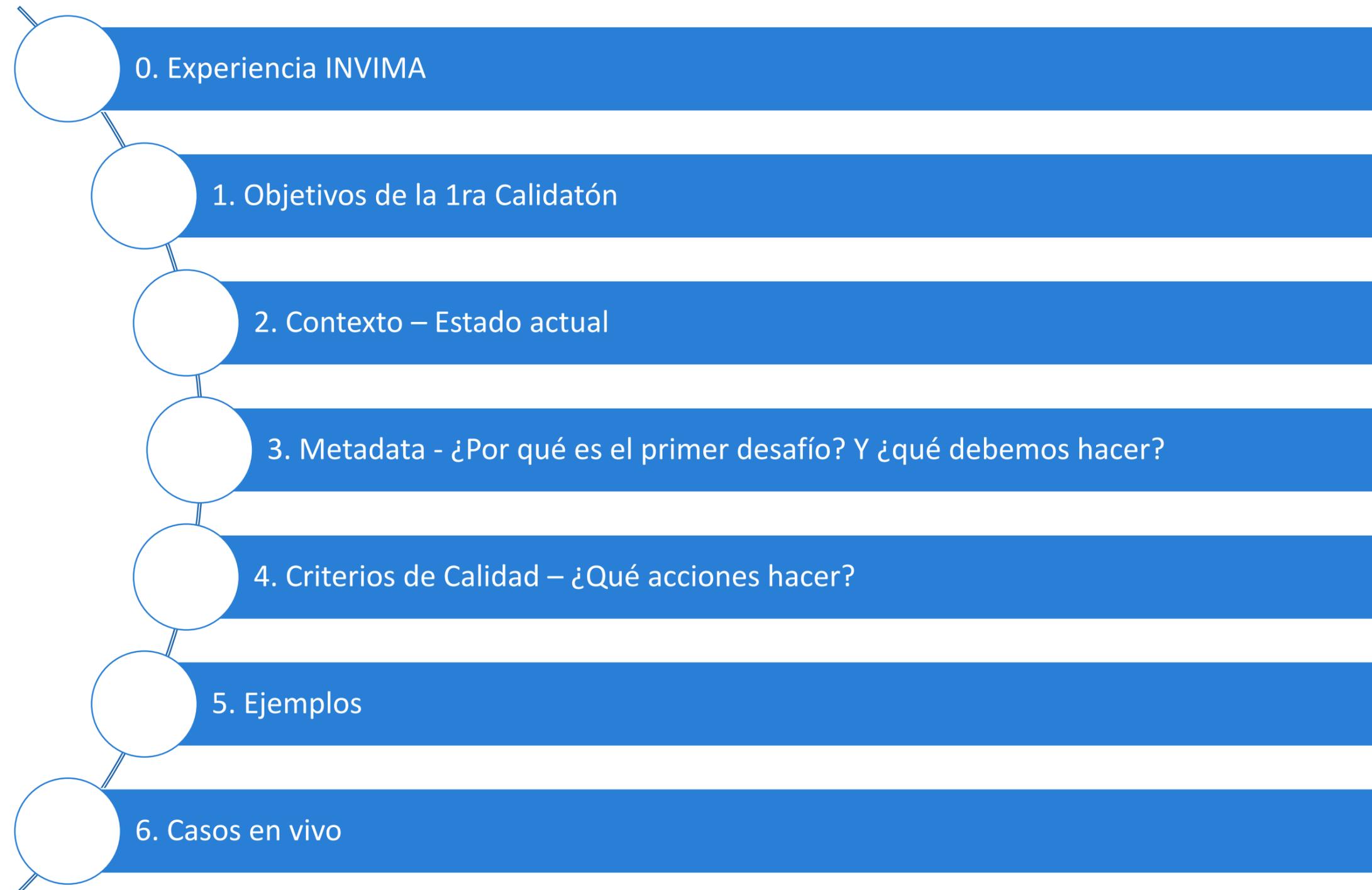


Primera Calidatón

Para que los datos sí den información

Marzo. 2020

Agenda



0. Experiencia INVIMA



1. Objetivos de la 1ra Calidatón



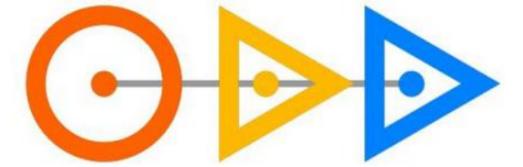
Objetivos de la Calidatón

- Compartir experiencias de calidad por parte de otras entidades
- Presentar cómo está a calidad de datos actualmente en datos.gov.co
- Explicar mejores prácticas para el uso de metadata
- Conocer los criterios y estándares de calidad de datos abiertos
- Tratar de forma práctica conjuntos de datos abiertos para que sí den información



El futuro digital
es de todos

MinTIC



OPEN DATA DAY 2020



2. Contexto – Estado actual de los datos en el portal

Política de Gobierno digital



El futuro digital
es de todos

MinTIC



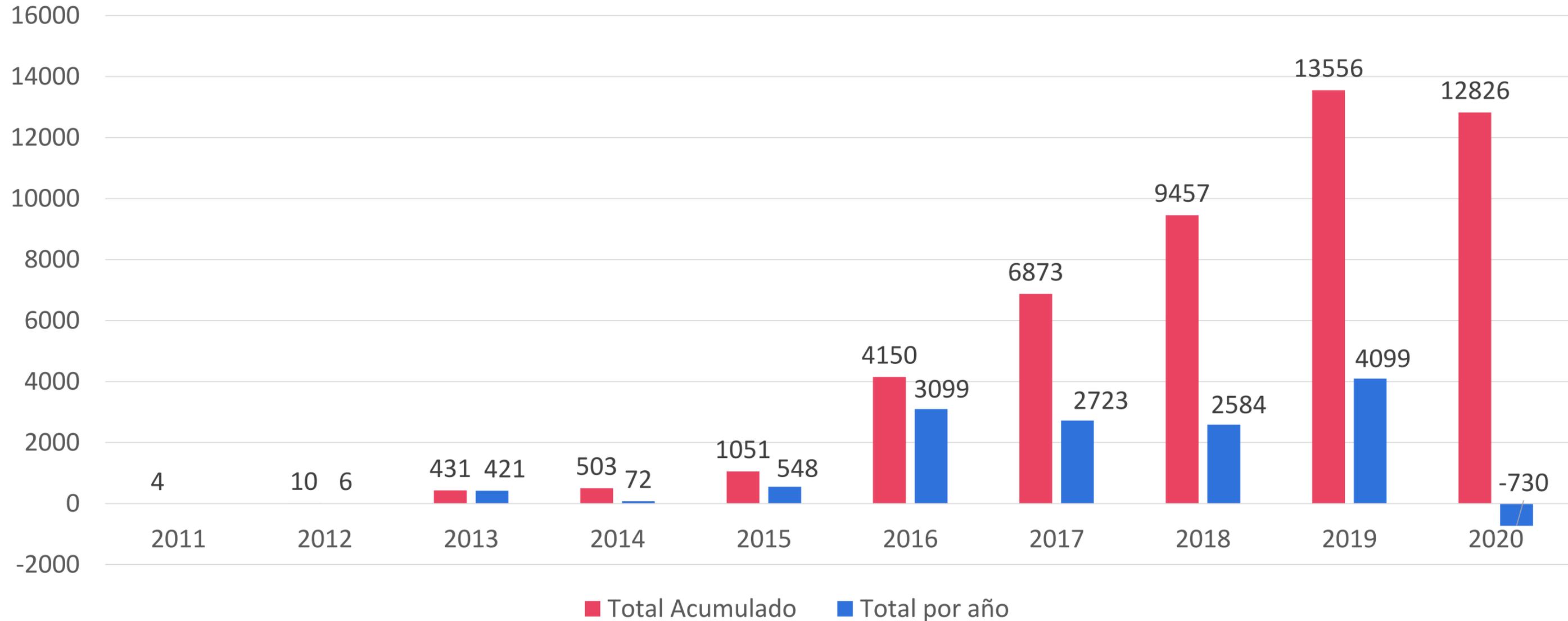
Portal Datos Abiertos Tasa de Crecimiento Anual



El futuro digital es de todos

MinTIC

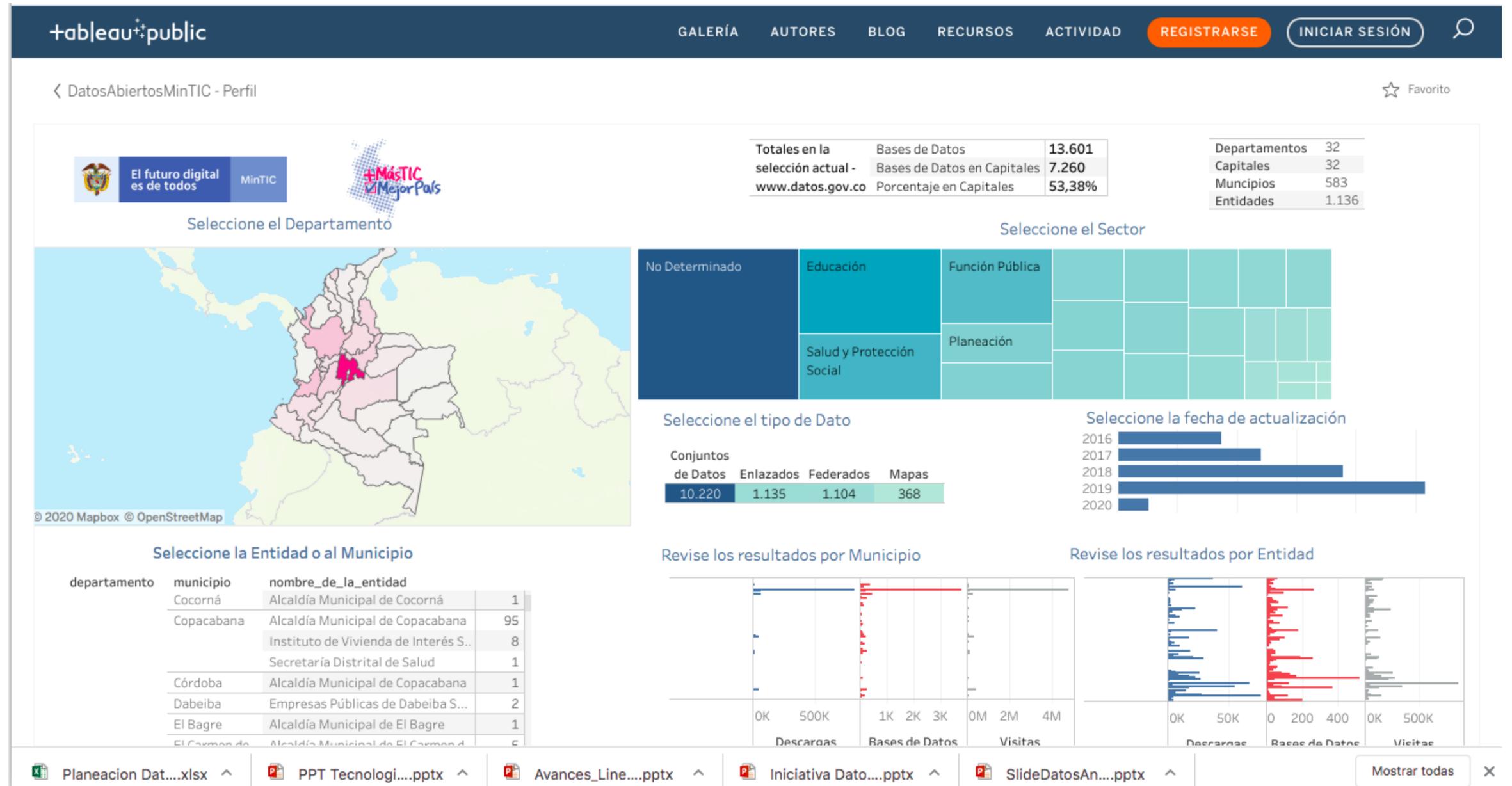
Crecimiento de Apertura de Datos por año



Fuente: inventario de conjuntos de datos abiertos corte febrero 2020

Tablero de avance en línea del Portal de Datos Abiertos

- Consulta por departamento
- Consulta por Municipio
- Visualización en Mapa
- Consulta por Tipo de Dato
- Consulta por fecha de actualización



Fuente: inventario de conjuntos de datos abiertos corte febrero 2020

Ver <https://public.tableau.com/profile/datosabiertosmintic#!/vizhome/shared/Y84XXHW6M>

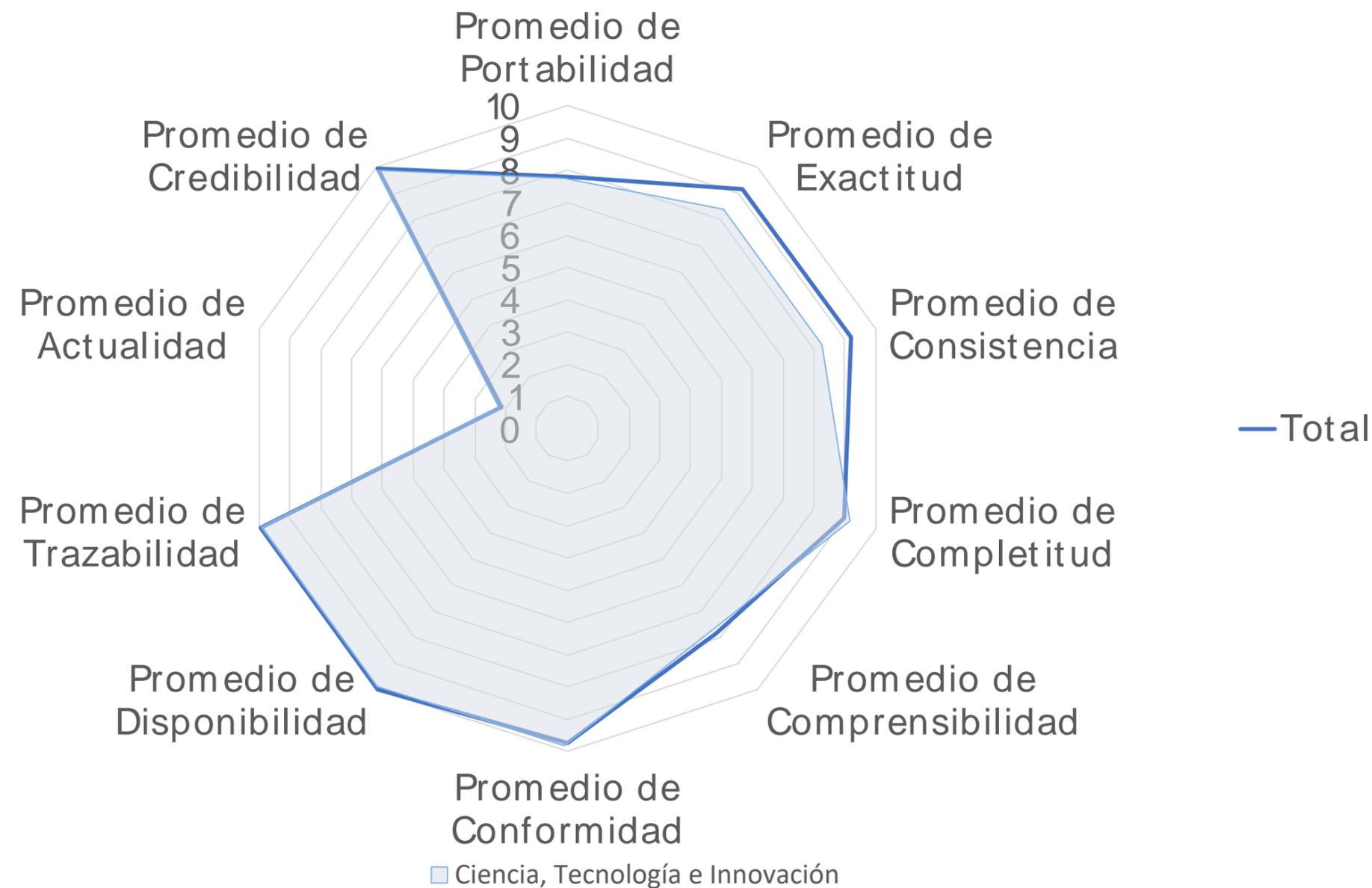
Herramienta de Calidad – Disponible próximamente



El futuro digital
es de todos

MinTIC

Calificación Promedio por Criterios de Calidad



- Visualización promedio de cada categoría en comparación con el promedio global
- Calificación por criterio en cada conjuntos de datos
- Segmentación en el conjunto de datos por Categoría y por Entidad
- Documentación de modo de cálculo y modelo de uso

Herramienta de Calidad – Disponible próximamente



El futuro digital
es de todos

MinTIC

Calidad de Datos Abiertos

Basado en [Calidad de Datos Abiertos](#)

(No se ha proporcionado una descripción)

| ID DatSet | Nombre ... | Categoría | Nombre | Completo... | Credibil... | Actualid... |
|-----------|------------------|-------------------|------------------|-------------|-------------|-------------|
| 22wn-wxiu | Secretaria de... | Agricultura y ... | CULTIVOS CA... | 10 | 10 | 0 |
| 2339-t3uf | UNIDAD NAC... | Gastos Gube... | Ejecucion Pre... | 9 | 10 | 0 |
| 233f-b6ed | RED DE LA S... | Salud y Prote... | Esquema de ... | 8 | 10 | 0 |
| 238m-denz | Migracion col... | Cultura | HOSPEDAJES ... | 9 | 10 | 0 |
| 239z-ikj9 | Secretaria de... | Agricultura y ... | Rendimiento ... | 7 | 10 | 10 |

➤ Identificación de cada conjunto de datos, por ID y nombre.

➤ Categoría asociada a cada conjunto de datos

➤ Nombre de la entidad asociada a cada conjunto de datos

Diagnóstico Preliminar del Portal



El futuro digital
es de todos

MinTIC

1

- Desarrollo de la herramienta de calidad, la cual estará disponible para final del del trimestre.
- 100% de los conjuntos diagnosticados.

2

- Hay errores recurrentes en la metadata que dificultan el uso y re-uso de la información.
- Los errores recurrentes están en la metadata de entidad, municipio, departamento, sector, categoría y fecha de actualización.
- De los conjuntos de datos aproximadamente un 30% presenta más de un error en su metadata básica

3

- A nivel territorial se podría decir que un 46,3% de los conjuntos publicados no corresponden a la información de las capitales. De estos municipios podríamos priorizar los que más conjuntos están produciendo por departamento
- De estos municipios podríamos priorizar los que más conjuntos están produciendo por departamento

4

- Es requerido realizar un plan independiente de intervención para Bogotá
- de forma independiente produce 32% de los conjuntos de datos del portal
- Más del 90% de esta información está enlazada o federada

3. Metadata

¿Por qué es el primer desafío? Y ¿qué debemos hacer?

Herramienta de Calidad – Disponible próximamente



El futuro digital
es de todos

MinTIC

Calidad de Datos Abiertos

Basado en [Calidad de Datos Abiertos](#)

(No se ha proporcionado una descripción)

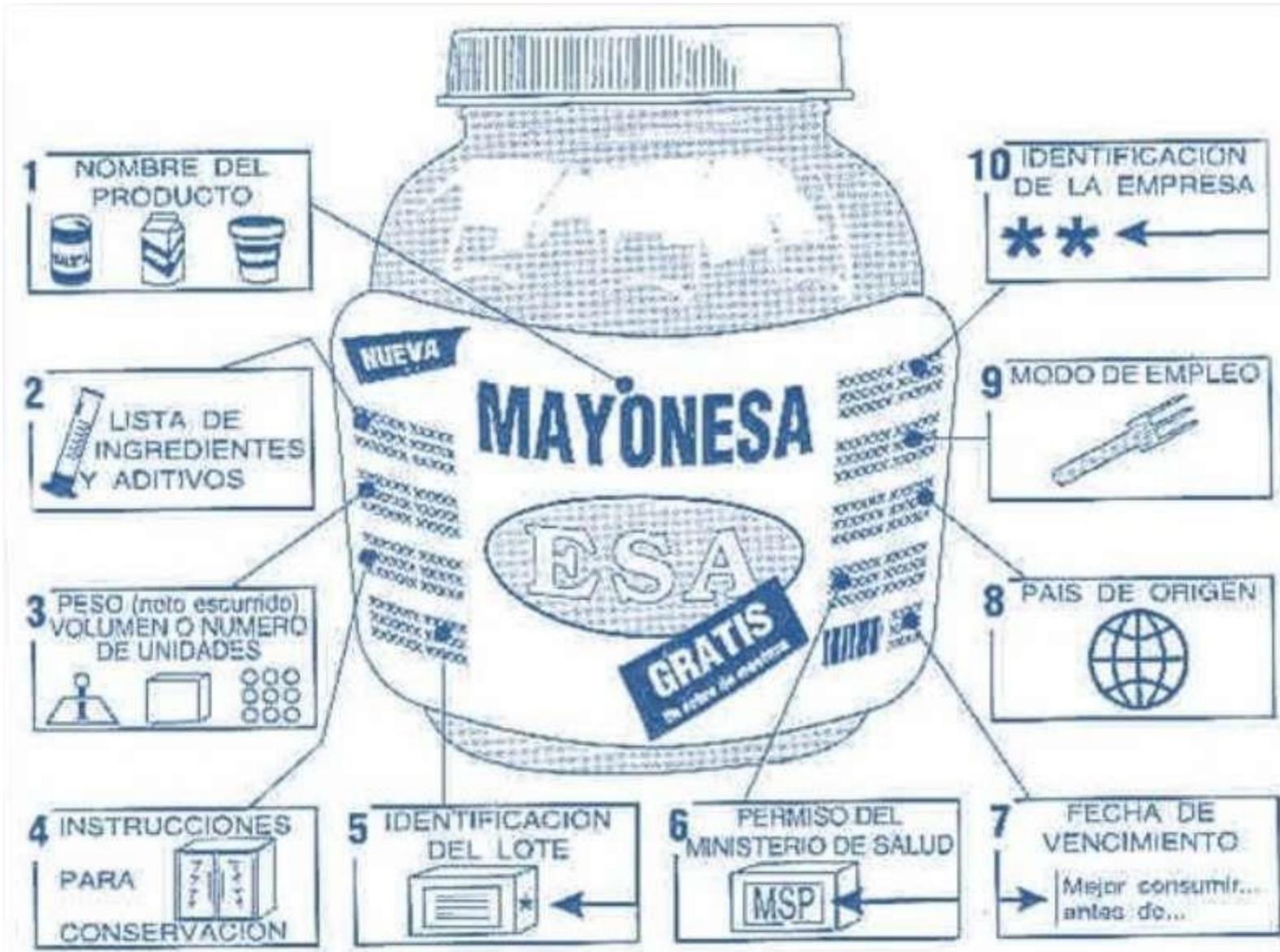
| ID DataSet | Nombre ... | Categoría | Nombre | Comple... | Credibil... | Actualid... |
|------------|------------------|-------------------|------------------|-----------|-------------|-------------|
| 22wn-wxiu | Secretaria de... | Agricultura y ... | CULTIVOS CA... | 10 | 10 | 0 |
| 2339-t3uf | UNIDAD NAC... | Gastos Gube... | Ejecucion Pre... | 9 | 10 | 0 |
| 233f-b6ed | RED DE LA S... | Salud y Prote... | Esquema de ... | 8 | 10 | 0 |
| 238m-denz | Migracion col... | Cultura | HOSPEDAJES ... | 9 | 10 | 0 |
| 239z-ikj9 | Secretaria de... | Agricultura y ... | Rendimiento ... | 7 | 10 | 10 |

- Identificación de cada conjunto de datos, por ID y nombre.
- Categoría asociada a cada conjunto de datos
- Nombre de la entidad asociada a cada conjunto de datos

SON METADATOS QUE FUERON DILIGENCIADOS AL MOMENTO DE SUBIR CADA CONJUNTO DE DATOS – ¿QUÉ PASA SI ESTÁN MAL?

¿Qué son los metadatos?

En términos simples, es la información de la información



Fuente: <https://www.geoidep.gob.pe/conoce-las-ides/metadatos/que-son-los-metadatos>



Metadatos

Los metadatos hacen referencia a la información que contiene el conjunto de datos. De ahí la importancia de ser rigurosos y precisos en el momento de diligenciarlos

Guía del dominio de Información de TI

¿Cómo están los metadatos en el portal de datos abiertos?



El futuro digital es de todos

MinTIC

Sobre este conjunto de datos

Actualizado

26 de diciembre de 2019

| | |
|-----------------------------------|-----------------------------------|
| Datos actualizados por última vez | Última actualización de metadatos |
| 26 de diciembre de 2019 | 26 de diciembre de 2019 |

Fecha de creación
2 de diciembre de 2019

| | |
|------------|-----------|
| Vistas | Descargas |
| 132 | 1 |

[Ver todas las estadísticas para este conjunto de datos](#)

¿Qué hay en este Conjunto de Datos?

| | |
|------------|-----------|
| Filas | Columnas |
| 12K | 14 |

Calidad de Datos Abiertos

Ciencia, Tecnología E Innovación

Información de la Entidad

| | |
|----------------------|--|
| Área o dependencia | Gobierno de Datos |
| Nombre de la Entidad | Ministerio de Tecnologías de La Información y Las Comunicaciones |
| Departamento | Bogotá D.C. |
| Municipio | Bogotá D.C. |
| Orden | Nacional |
| Sector | Ciencia, Tecnología e innovación |

Información de Datos

| | |
|-----------------------------|------------|
| Idioma | Español |
| Cobertura Geográfica | Nacional |
| Frecuencia de Actualización | Mensual |
| Fecha Emisión (aaaa-mm-dd) | 2019-12-02 |

Temas

| | |
|-----------|---|
| Categoría | Ciencia, Tecnología e Innovación |
| Etiquetas | <i>Este conjunto de datos no tiene ninguna etiqueta</i> |

Columnas en este Conjunto de Datos

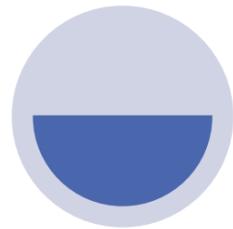
| Nombre de Columna | Descripción | Tipo |
|-------------------|-------------|----------------|
| ID DataSet | | Texto simple T |

Actualizado

26 de diciembre de 2019

Datos proporcionados por
Propios

¿Cómo están los metadatos en el portal de datos abiertos?



Información del conjunto de datos

Nombre

Fecha de Actualización

- Datos
- Metadatos

Número de Vistas

Número de Descargas

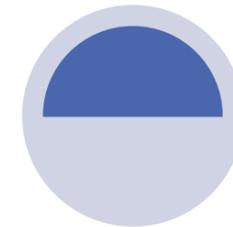
Columnas: Variables

- Nombre, descripción y tipo de dato

Filas: Registros (Observaciones)

Categoría

Palabras Clave



Información de la Entidad

Área / Dependencia

Nombre de la Entidad

Ubicación

- Departamento
- Municipio

Sector

Orden: Nacional, Municipal, territorial, etc.

¿Para qué necesitamos metadatos?

Nos permiten identificar la información de interés, buscar y encontrar la información, compartirla y recuperarla entre otros.



Lenguaje Claro

¿quién buscará la información?
Debe ser un lenguaje apropiado, natural y lógico. Evitar tecnicismos y siglas.

Ej: MEN



Consolidado

La segmentación de la información no es conveniente, en general se presenta esto por territorio y fechas

Ej: Desplazados



Palabras Clave

Son la llave principal de búsqueda que empleamos para referirnos a la información. Suele ser la respuesta de ¿qué información necesito?

Ej: Barrios Popayan



Preciso

Que la información corresponda tanto con la entidad como con el conjunto de datos. Entre más detallada la precisión, más adecuada la búsqueda

Ej: Cultivos

¿Cómo diligenciarlo en el portal?

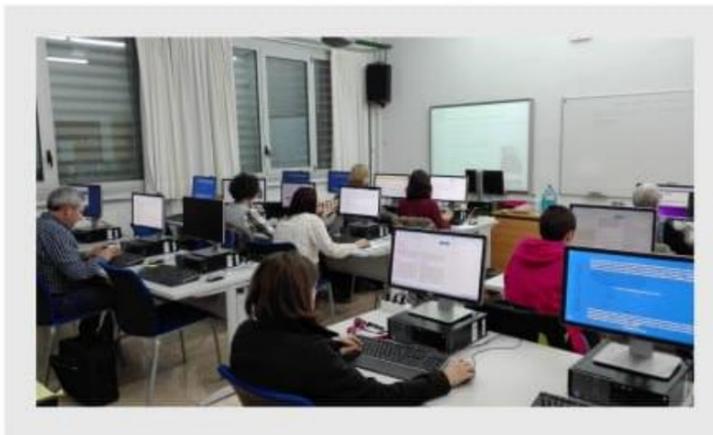


1 Tenga claro su usuario

Actualiza tus usuarios registrados en el Portal de Datos Abiertos

Publicado por Site Admin, Febrero 13 de 2020

Categoría: Eventos



El portal nacional de datos abiertos www.datos.gov.co está dispuesto para que todas las entidades de orden nacional y territorial publiquen datos abiertos de calidad para consulta, uso y aprovechamiento de los ciudadanos.

Estamos en un proceso de actualización de usuarios y de mejorar la calidad de los datos abiertos disponibles en nuestro portal y queremos invitarte a verificar:

1. Si tu entidad dispone datos abiertos en www.datos.gov.co
2. Si tienes datos disponibles, verifica que el o los usuario(s) que los administra se encuentren vigentes en tu Entidad. Consulta los siguientes conjuntos de datos:

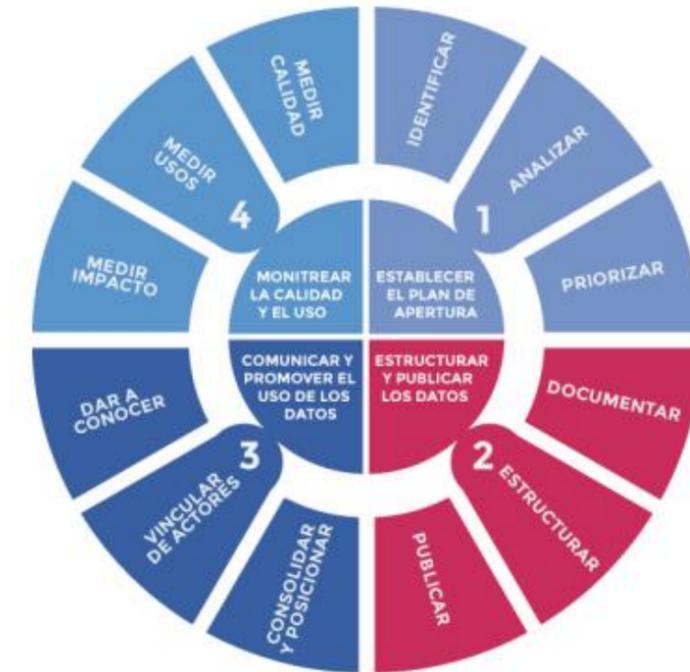
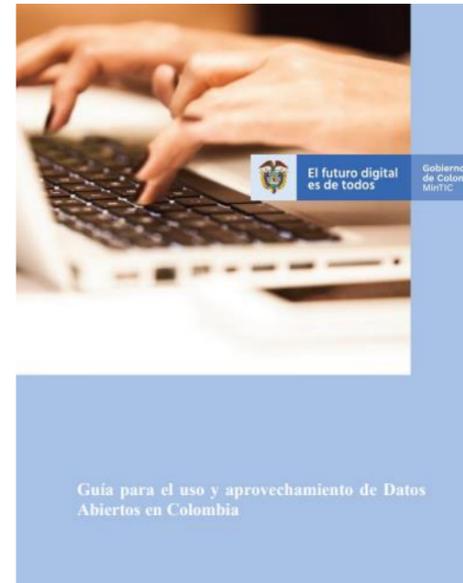
- Listado de usuarios publicadores del portal
- Listado de conjuntos de datos del portal datos.gov.co y sus usuarios publicadores

3. Si los datos publicados se acogen a los términos y condiciones del portal

Para dudas sobre el estado del usuario de tu Entidad, contacta a nuestra mesa de ayuda a la línea gratuita: 01-8000-910742, Bogotá (571) 390 79 51 de Lunes a Viernes de 7 a.m. a 6 p.m. y Sábados de 8 a.m. a 1 p.m. o al Correo electrónico: sopORTECCC@mintic.gov.co

Te invitamos a consultar nuestro tutorial sobre cómo crear, actualizar y eliminar un usuario

2 Use la guía de datos abiertos



3 Tenga claros los metadatos

Pueden asociarse a los archivos originales

Evite que dependa del usuario editor

El productor de datos debe proveer esa información

¿Cómo diligenciarlo en el portal?



Aquí deben diligenciarse los metadatos antes de publicar

Acciones de datos

Para agregar el conjunto de datos

Añadir datos
Añada datos tabulares, geoespaciales, no analizables o vinculados externamente
Añadir datos

Revisar y configurar datos
Revisar datos

Editar los metadatos del conjunto de datos
Añada un contexto para ayudar a los usuarios a comprender sus datos
Editar metadatos

Editar columna de metadatos
Editar metadatos

Gestionar metadatos

Metadatos del conjunto de datos
No hay cambios

Descripciones de las columnas

Título y descripción
Haga su título y descripción tan claros y simples como pueda.

Título de Conjunto de Datos *

Prueba

Breve Descripción

Introduzca una descripción

Corresponden justamente con cada uno de los campos ya descritos

Ingresando como usuario

4. Criterios de Calidad – ¿Qué acciones hacer?

¿Qué es Calidad de Datos?



La capacidad del dato de prestar adecuadamente una función o servicio, es decir el dato debe permitir....

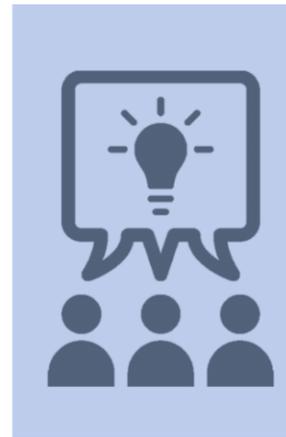


En términos simples es lo que permite que haya
“Aptitud para el uso”

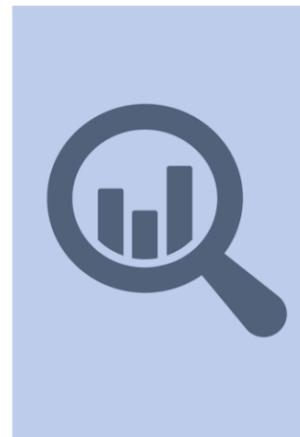
¿Qué podemos hacer para que nuestros datos sí tengan calidad?



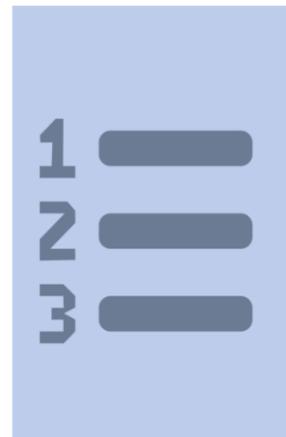
Conozca las guías disponibles tanto el [Manual de Gobierno Digital](#), como la [Guía de estándares de calidad e interoperabilidad](#).



Tenga claro cuál es la información que produce por su gestión y función y realice un **plan de datos** para disponer de ella.



Tenga clara cuál es la información, metadatos, y su contenido, datos.



Haga uso de **herramientas para la depuración y limpieza de datos**, muchas son abiertas y libres como [Open Refine](#).

¿ Qué podemos hacer para que nuestros sí tengan calidad?



Familiarizarse con los términos y condiciones del portal

TÉRMINOS Y CONDICIONES DE USO PORTAL DE
DATOS ABIERTOS

TÉRMINOS Y CONDICIONES DEL PORTAL DE DATOS

Recuerde que no se trata de formación procesada

Ejemplo

Emplee herramientas de depuración y análisis de datos



Open Refine

Desde el grupo de datos
preparamos video tutoriales para
iniciar el uso de estas herramientas
libres



Lenguaje R

3. ¿ Cuales son los 16 principios de calidad?



1. Confidencialidad

2. Relevancia

3. Actualidad

4. Trazabilidad

5. Conformidad

6. Exactitud

7. Completitud

8. Consistencia

9. Precisión

10. Portabilidad

11. Credibilidad

12. Comprensibilidad

13. Accesibilidad

14. Eficiencia

15. Recuperabilidad

16. Disponibilidad

3. ¿ Cuales son los 16 principios de calidad?

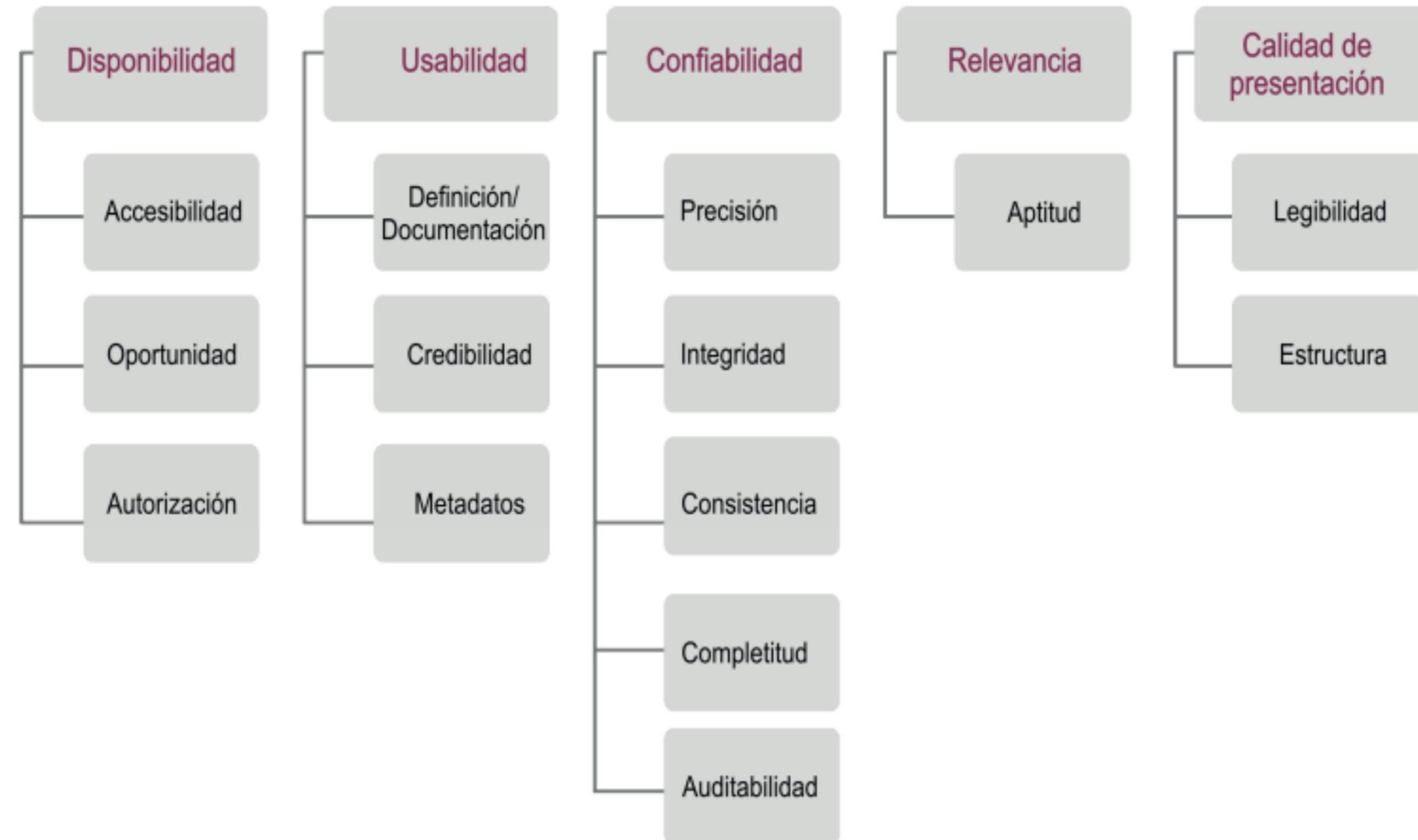


GUÍA DE ESTÁNDARES DE CALIDAD E INTEROPERABILIDAD DE LOS DATOS ABIERTOS DEL GOBIERNO DE COLOMBIA

- Datos abiertos para la transparencia y el control social
- Datos abiertos para empoderar al ciudadano
- Datos abiertos para mejorar o crear productos y modelos de negocio
- Datos abiertos como una puerta abierta a la innovación
- Datos abiertos para mejorar la eficiencia y eficacia del Estado
- Datos abiertos para medir el impacto de las políticas

Estándar universal de calidad de datos de dos capas



4.

¿ Cómo mejorar?



Para aumentar la cantidad de filas se recomienda evitar las agregaciones de registros a través de operaciones como promedios, mínimos, máximos, etc.



Para aumentar la cantidad de filas se recomienda incluir variables de hora y/o fecha, conocido como formato panel, obteniendo registros tipo transacción en lugar de consolidados.



Para mejorar los metadatos se deben diligenciar todas las casillas del formulario en el portal. Especialmente descripción y palabras clave.



Hagan uso de las herramientas que están disponibles, tanto la de evaluación de calidad de datos como los videotutoriales.

5. Ejemplos

Ejemplo guiado

Centros de Transformación digital

Ciencia, Tecnología E Innovación

Ver datos

Visualizar y explorar ▾

Exportar

API

...

Centros de Transformación digital

Actualizado

21 de febrero de 2020

Datos proporcionados por

Ministerio de Tecnologías de la
Información y las Comunicaciones

Sobre este conjunto de datos

Actualizado

21 de febrero de 2020

Datos actualizados
por última vez

21 de febrero de 2020

Última actualización
de metadatos

27 de enero de 2020

Información de la Entidad

Área o dependencia

Dirección de transformación digital

Nombre de la Entidad

Ministerio de Tecnologías de La Información y Las
Comunicaciones

Departamento

Bogotá D.C.

- Empecemos con Open Refine
- También trabajemos con Excel

<https://www.datos.gov.co/Ciencia-Tecnolog-a-e-Innovaci-n/Centros-de-Transformaci-n-digital/b84s-hkiv>

6. Casos en vivo



El futuro
es de todos

Alta Consejería
para la Innovación
y la Transformación
Digital

CAROLINA MATAMOROS
Líder de analítica y calidad de datos
abiertos
datosabiertos@mintic.gov.co

Ministerio de Tecnologías de la
Información y las Comunicaciones
Tel:+57(1) 344 34 60
Edif. Murillo Toro Cra. 8a entre calles 12 y
13, Bogotá, Colombia - Código Postal 111711
www.mintic.gov.co