

# **ANÁLISIS** **DE TENDENCIA** **TIC**

## **DATOS OSCUROS**

**DEPARTAMENTO ADMINISTRATIVO DE TECNOLOGÍAS  
DE LA INFORMACIÓN Y LAS COMUNICACIONES**

**SUBDIRECCIÓN DE INNOVACIÓN DIGITAL**



[www.cali.gov.co](http://www.cali.gov.co)



ALCALDÍA DE  
SANTIAGO DE CALI  
DEPARTAMENTO ADMINISTRATIVO DE  
TECNOLOGÍAS DE LA INFORMACIÓN  
Y LAS COMUNICACIONES

# CONTENIDO

Introducción.....	2
Objetivo.....	2
Conceptos, siglas y abreviaturas.....	2
1. Análisis de antecedentes.....	3
2. Descripción de la tendencia.....	3
3. Organizaciones y personas que utilizan la tendencia.....	7
4. Beneficios (valor público).....	7
5. Causantes del cambio.....	9
6. Retos.....	9
7. Insights y/o Necesidades Básicas.....	9
8. Aplicaciones en la Alcaldía de Santiago de Cali.....	9
Conclusiones.....	11
Bibliografía.....	11
CRÉDITOS.....	12

# DATOS OSCUROS

ANÁLISIS DE LA TENDENCIA TIC

TIPO DE TENDENCIA: TECNOLOGÍA

Carlos Eduardo Jiménez Bacca. - carlos.jimenez.b@cali.gov.co

## INTRODUCCIÓN

El presente estudio de la tendencia digital Datos Oscuros o Dark Data por su nombre en inglés, se genera en cumplimiento al Decreto Extraordinario 0516 de 2016, específicamente en la segunda función de la Subdirección de Innovación Digital que reza "Realizar investigación para identificar y proponer soluciones tecnológicas orientadas a mejorar la calidad de vida de las personas."

En el estudio se analiza y se descomprime la tendencia en análisis de antecedentes, TIC o modelo de negocios que preceden, descripción de la tendencia, porque la tendencia es un punto de inflexión, aplicaciones, organizaciones y personas que utilizan la tendencia, beneficios (valor público) y amenazas para el usuario, causantes del cambio, retos, insights y/o necesidades básicas, aplicaciones en la Alcaldía de Santiago de Cali y conclusiones. Todos estos factores clave para comprenderla mejor y así generar soluciones innovadoras.

Esta investigación es fruto del procedimiento de Inteligencia y Análisis de las Tendencias TIC cuyo propósito es proveer información válida y relevante sobre tendencias e innovaciones que puedan tener un impacto positivo (oportunidades) o negativo (amenazas) sobre la Alcaldía de Santiago de Cali o sus organismos.

## OBJETIVO

Este estudio tiene como objetivo informar a la ciudadanía y organismos de la Alcaldía de Santiago de Cali de las tendencias digitales que puedan servir de insumo para la innovación digital de los procesos, productos o servicios que la Entidad presta la a sus ciudadanos, facilitando su análisis y por consiguiente, su aplicación en los ejercicios de ideación de soluciones innovadoras basadas en nuevas tecnologías.

## CONCEPTOS, SIGLAS Y ABREVIATURAS

**Big Data:** También llamados datos masivos, macrodatos, inteligencia de datos, datos a gran escala o big data (terminología en idioma inglés utilizada comúnmente) es un término que hace referencia al concepto relativo a conjuntos de datos tan grandes y complejos como para que hagan falta aplicaciones informáticas no tradicionales de procesamiento de datos para tratarlos adecuadamente.

**Databerg:** Iceberg de datos.

**Experiencia del Usuario:** Es el conjunto de factores y elementos relativos a la interacción del usuario, con un entorno o dispositivo concretos, cuyo resultado es la generación de una percepción positiva o negativa de dicho servicio, producto o dispositivo. Ésta depende no solo de los factores relativos al diseño (hardware, software, usabilidad, diseño de interacción, accesibilidad, diseño gráfico y visual, calidad de los contenidos, buscabilidad o encontrabilidad, utilidad, etc) sino además de aspectos relativos a las emociones, sentimientos, construcción y transmisión de la marca, confiabilidad del producto, etc.

**ROT data:** ROT (Redundant, Obsolete or Trivial). Datos redundantes, obsoletos o triviales.

**PAD:** Puntos de Apropiación Digital.

**TIC:** Tecnologías de la Información y las Comunicaciones.

**Inteligencia digital:** Es la capacidad para seleccionar los datos correctos y saber qué hacer con ellos. Análisis de datos para hacer previsiones, responder preguntas y exponer planes de acción relacionados

DATOS OSCUROS / TECNOLOGÍA

con cuestiones que ni siquiera sabía que debía hacerse, incorporar insights en los procesos de negocio, supervisar procesos con alertas automatizadas, implementar modelos predictivos para la optimización de toma de decisiones en el momento oportuno.

## 1. ANÁLISIS DE ANTECEDENTES

En esta era de iluminación impulsada por la tecnología, los datos son nuestra moneda competitiva. El problema es que no todos los datos son creados igual. A un nivel básico hay tres clases de datos: los Críticos para el negocio que deben ser protegidos, los Redundantes, Obsoletos y Triviales (ROT) que tienen poco o ningún valor, y los Datos Oscuros cuyo valor no ha sido determinado aún.

Los datos oscuros enterrados dentro de la información sin procesar generada en volúmenes alucinantes por sistemas transaccionales, redes sociales, motores de búsqueda e innumerables otras tecnologías, son puntos de vista críticos y estratégicos, de clientes y operacionales que, una vez iluminados por el análisis, pueden validar o aclarar suposiciones, informar la toma de decisiones, y ayudar a trazar nuevos caminos hacia el futuro.

En muchos casos, los miembros de una organización simplemente no son conscientes de la existencia de estos datos. En otros, hay desconexión entre los equipos y procesos, egoísmo al compartir información o limitaciones técnicas de software y hardware.

Los "datos oscuros" se están convirtiendo en un problema importante para las organizaciones, ya que la capacidad de recopilar datos supera con creces la capacidad de analizar datos.

### Línea de tiempo

Fecha	Hitos de la tendencia
1900s	Enfoque pasivo y retrospectivo.
1961	Ley de aumento exponencial.
1999	Big Data.
2014	Datos oscuros.
2016	Experiencia del usuario centrada en el ciudadano.
2018	Inteligencia digital.

Palabras clave: Datos oscuros. ROT data. Experiencia del ciudadano

## 2. DESCRIPCIÓN DE LA TENDENCIA

Los Datos Oscuros comprende un tipo de datos no estructurados, no etiquetados y desaprovechados, que se encuentran en los almacenes de datos y no se han analizado ni tratado. Son parecidos a los macrodatos (Big data) pero se diferencia en que en su mayoría, los administradores de negocio y de TI ignoran el valor que tienen.

“Cuando ese dato oscuro se ilumina, deja de serlo dark data y pasa a ser parte del proceso de negocio, por lo que debemos tener un proceso continuo de búsqueda de información hasta debajo de las piedras”, declara Jorge Fernández, mánager de Thinkalia.

Quizá la mejor forma de entender en qué consisten el Dark Data es su contraposición frente al Big data<sup>1</sup>, como se verá en un apartado más adelante.

El Databerg o el “iceberg de los datos” muestra que de todos los datos de una entidad, 12% son críticos para el negocio, 23% son ROT (redundantes, obsoletos y triviales) y el 65% son oscuros. Las razones por las que se presentan los Datos Oscuros son:

- 85% no hay una herramienta para desvelar o capturar los datos oscuros.
- 39% hay demasiados datos pero insuficiente analítica de datos.
- 25% solo hay acceso a datos estructurados.
- 66% hay datos perdidos o incompletos.

### **Clasificación de los Datos Oscuros**

1. Los datos que no se están recopilando actualmente.
2. Datos que se están recopilando, pero a los que resulta difícil acceder en el momento y lugar correctos.
3. Datos que se están recopilando y están disponibles, pero que aún no se han convertido en productos o no se han aplicado completamente.

La analítica de datos oscuros se enfoca principalmente en datos brutos basados en texto que no se han analizado, con énfasis en datos no estructurados, que pueden incluir cosas como mensajes de texto, documentos, correo electrónico, archivos de video y audio e imágenes fijas

### **Dark Data frente a Big Data**

El problema inherente al Big Data radica en las llamadas tres V: Volumen, Velocidad y Variedad. Es decir, el gran volumen de datos que nos llega, la velocidad y frecuencia constante a la que lo hace, y la variedad en el formato y contenido de estos datos.

La dificultad principal del Big Data es el análisis de un volumen grande de datos, la rapidez en el análisis y la selección de los datos, mientras que en el Dark Data la complejidad radica en la identificación y recolección de esos datos.

El Big Data aporta esencialmente información de clientes, proveedores, competencia y productos sustitutos. Su análisis permite identificar cómo una organización se relaciona con la cadena de valor externa. Sin embargo, en Dark Data los objetos de análisis son sus procesos de negocio. Se trata de hacer accesible y analizable la información generada en los procesos de la cadena de valor interna.

Otra diferencia sustancial es la envergadura de los proyectos y los costos de implementación de una u otra tecnología. El costo de abordar un proyecto de Big Data es elevado. Es como intentar construir una nave espacial. Te enfrentas a problemas desconocidos con tecnología insuficiente. Y la identificación y generación de dicha información puede tener un costo considerable. En cambio, el costo de un proyecto de Dark Data es relativamente bajo, pues la dificultad radica en encontrar los datos. Equivale a cavar un gran hoyo con un pico y una pala. “Va a costar, pero sabemos que el dato está allí. Sólo hay que encontrarlo”, afirma el responsable de Thinkalia.

### **Dos maneras de aproximarse**

<sup>1</sup> <https://es.wikipedia.org/wiki/Macrodatos>

Hay dos formas de acercarse a los Datos oscuros. Por un lado, se puede utilizar aquellos que ya están registrados en los sistemas de información. Puede que el 60%-70% de los datos que se pueden utilizar ya estén recolectados de una u otra manera y se estén desperdiciando. Se trata de cambiar los procedimientos en los sistemas de información para activar esos datos, capturarlos y darlos a conocer dentro de la entidad. En este caso, puede que se necesite un servicio de consultoría para identificar los datos susceptibles de ser rescatados y utilizados para mejorar su eficiencia. Yendo un paso más allá, quizá sea preciso instalar algún software adicional, esencialmente para traducir los datos a un formato adecuado para su visualización.

La otra manera de acceder es mediante la instalación de un hardware específico que recopila estos datos de difícil recolección, como un dispositivo para monitorizar contadores de electricidad, lectores de códigos de barras, etc. Y después habrá que tratar y analizar esos datos. Como es lógico, los servicios contratados – consultoría, software, hardware...– determinarán el costo del proyecto.

### **¿Por qué esta tendencia es un punto de inflexión?**

IBM estima que para 2020, el volumen de Datos oscuros ascenderá a 93%, poniendo como ejemplo que los carros generarán 350 MB de datos cada segundo.

Con la creciente acumulación de datos estructurados, no estructurados y semiestructurados en las organizaciones –cada vez más a través de la adopción de aplicaciones de macrodatos– los datos oscuros llegan para denotar datos operativos que se dejan sin analizar. Tales datos se ven como una oportunidad económica para las empresas sí pueden aprovecharlos para obtener nuevos ingresos o reducir costos internos.

Los ejemplos más comunes de datos oscuros son:

- Hojas de cálculo (en un estudio, una empresa con 1.500 empleados tenía 2,5 millones de hojas de cálculo, lo que equivale a miles de millones de celdas de datos).
- Múltiples versiones anteriores de documentos.
- Archivos adjuntos de correo electrónico y .zip archivos que se descargan y luego se ignoran.
- Bases de datos inactivas e información de clientes no utilizada.
- Archivos de empleados anteriores y contenido (por ejemplo, notas del proyecto).
- Informes de análisis y datos de encuestas.
- Archivos de registro, información de cuenta e historial de transacciones.

En resumen, los datos oscuros que son descartados por un proceso u organismo, pueden ser muy valiosos para otro.

Adicionalmente hoy en día la innovación digital está siendo promovida por una cultura impulsada por los datos.

Hasta la fecha, las empresas han explorado sólo una pequeña fracción del universo digital para obtener un valor analítico. IDC estima que para 2020, hasta el 37 por ciento del universo digital contendrá información que podría ser valiosa si se analiza.

¿Pero exactamente qué tan valioso? IDC también proyecta que las organizaciones que analicen todos los datos relevantes y entreguen información procesable lograrán \$430 mil millones adicionales en ganancias de productividad sobre sus pares menos orientados analíticamente para 2020<sup>2</sup>.

### **Aplicaciones**

<sup>2</sup> <https://www.idc.com/getdoc.jsp?containerId=259835>

1	Asesoramiento e individualización de la oferta	Los contadores, que eran totalmente oscuros, han dejado de serlo. Ahora, Gas Natural tiene todos los datos cada segundo, pudiendo contactar con sus clientes para explicarles por qué consumen y de qué manera, ofrecer un servicio de valor añadido de eficiencia para las industrias, etc.
2	Optimización de consumo y detección de fallas	La utilización de los datos oscuros permite reunir información acerca del funcionamiento de los equipos con el fin de reducir consumos innecesarios o detectar fallas.
3	Mejora de la venta online	Todos los procesos a través de Internet dejan un rastro de datos. Sin embargo, dicha información no siempre se utiliza. Se analizan cuáles son los productos más buscados, los más vendidos, etc., pero hay mucha información que se pierde sin ser usada, ya que no es registrada.
4	Optimización de compras y stocks	Todo el producto que se ha solicitado por parte de los clientes pero que no se ha llegado a vender porque no estaba en stock. Revisando esos datos, los analistas son capaces de establecer una política de compras que les va a reportar mucho más beneficio, porque van a vender más o con mejor margen. Adicionalmente, el aprovechamiento de estos datos puede ayudar a mejorar la rotación de almacén, evitar roturas de stock o la acumulación de producto no vendido.
5	Registro del funcionamiento de plataformas	Captura de todos los datos de los sistemas en una plataforma Big Data para seguimiento de lo que está sucediendo. Así la entidad se centra en su negocio.
6	Estudios de mercado	<p>Uno de los grandes problemas de los fabricantes es que no pueden controlar cómo, cuándo y dónde se venden sus productos. Se tienen que fiar de lo que les dicen las empresas de investigación de mercado. La tecnología de Datos oscuros permite saber cómo se venden sus productos.</p> <p>De esta manera, el fabricante puede saber en qué tiendas se vende cada uno de sus artículos, qué días y a qué horas se registra mayor venta, si estas referencias se venden solas o junto a otros productos propios o de la competencia. Y esto también permite conocer mejor el efecto real de las campañas de marketing.</p>
7	Mejora del proceso comercial	<p>Los datos oscuros permiten conocer mejor el proceso de negociación de un proveedor con las compañías que adquieren sus bienes y servicios. Normalmente, el sistema de información de una empresa sólo registra el estado final de los pedidos, pero no los estados intermedios, como variaciones en los precios o en el volumen, la interacción con el cliente. Y el business intelligence visualiza la información en 'fotos' –ventas de un mes, evolución del total de ventas en el último año, pero no ofrece una visión de la evolución de la relación con el cliente desde el inicio hasta la fecha.</p> <p>Este proceso incluye Datos oscuros, que se puede emplear para mejorar el proceso comercial desarrollando el registro de todo aquello que envuelve al proceso de creación de la oferta, como cambios de productos, modificaciones de precio, incrementos de descuentos, entre otros.</p>

8	Medición de tráfico en tienda	Con la ayuda de un pequeño dispositivo, se puede hacer un seguimiento del flujo de clientes, midiendo cuántas personas entran o salen, cuánto tiempo pasan en la tienda, con qué frecuencia vuelven, etc. Esos datos se pueden extraer sirviéndose de diferentes tecnologías.
9	Seguimiento de empleados	Hacer tracking de los empleados sólo con analizar qué máquinas están encendidas o apagadas, permitiendo conocer cuándo vienen y se van, si están presentes o no, tiempo de permanencia, etc.
10	Carreras profesionales	Capturar la información de la formación que se va recibiendo para conocer qué motivaciones tiene cada empleado. Quizá una persona que trabaja en logística esté realizando cursos orientados al marketing, por lo que tal vez sería preciso que dicho departamento evaluará una trayectoria profesional horizontal de este empleado.

### 3. ORGANIZACIONES Y PERSONAS QUE UTILIZAN LA TENDENCIA

- Datumize<sup>3</sup>: Proveedor de Dark Data que busca que su negocio “aumente las ventas y mejore la eficiencia al integrar y transformar sus datos oscuros en nuevos y poderosos conocimientos sobre su negocio, cliente y inteligencia operativa” con clientes en el sector turismo, sector hotelero, negocios de almacenaje.
- USA.gov o Data.gov<sup>4</sup>: Es la guía en línea de información y servicios gubernamentales del Gobierno de Estados Unidos. Cuenta con más de 285,614 conjuntos de datos. Sirve para buscar información oficial.
- IBM<sup>5</sup>: de los datos a la disrupción: innovación a través de la inteligencia digital.
- IBM Watson<sup>6</sup> con la nube de IBM se puede ayudar a las empresas a ver los datos de muchas formas: video, audio, blogs, tweets entre otros. La nube de IBM es capaz de analizar lo que antes parecía imposible.
- Apple<sup>7</sup>: Informes recientes sugieren que Apple compró la empresa de inteligencia artificial (IA) Lattice Data por \$ 200 millones. Lattice Data aplicó un motor de inferencia habilitado para AI para extraer datos oscuros.
- Chan Zuckerberg Initiative<sup>8</sup> (CZI) compró Meta por un monto no revelado. Meta es una empresa de motores de búsqueda de investigación impulsada por inteligencia artificial que CZI planea poner a disposición libremente. Por lo tanto, la tecnología de extracción de datos oscuros y la propiedad intelectual desarrollada en la empresa también son potencialmente valiosos de manera independiente. GovLoop se asoció con DLT Solutions y Veritas para comprender mejor el costo de los datos oscuros en el gobierno y cómo las agencias pueden superarlo<sup>9</sup>.

### 4. BENEFICIOS (VALOR PÚBLICO)

- Diseño de nuevas experiencias del ciudadano a partir de los datos.
- De los datos a la disrupción: innovación a través de la inteligencia digital.
- Los datos oscuros son valiosos porque a menudo contienen información que no está disponible en ningún otro formato. Por lo tanto, las organizaciones continúan pagando el costo de recopilar y almacenar datos oscuros para fines de cumplimiento y con la esperanza de explotar los datos (para obtener información valiosa) en el futuro.
- Análisis de mejor calidad. Con acceso a mejores fuentes de datos y más información, la calidad de los análisis mejora dramáticamente. El análisis no solo se basa en un conjunto más grande de datos de

<sup>3</sup> <https://www.datumize.com/>

<sup>4</sup> <https://www.data.gov/>

<sup>5</sup> <https://www.ibm.com/analytics/>

<sup>6</sup> <https://www.youtube.com/watch?v=PBpCR7inFhQ>

<sup>7</sup> <https://www.applesfera.com/general/apple-adquiere-la-empresa-de-inteligencia-artificial-lattice-data-por-200-millones-de-dolares>

<sup>8</sup> <https://chanzuckerberg.com>

<sup>9</sup> <http://dlt.com/>

alta calidad, sino que los datos están disponibles para su análisis con prontitud. El resultado es una toma de decisiones más rápida y mejor basada en datos, que a su vez conduce al éxito comercial y operacional.

- Costos y riesgos reducidos. La extracción de datos oscuros deja a las organizaciones menos expuestas a los riesgos y la responsabilidad en la obtención de información confidencial. Las organizaciones también pueden purgar de forma segura los datos innecesarios, lo que reduce los costos recurrentes de almacenamiento y conservación. El cumplimiento regulatorio también se vuelve más fácil.
- La tecnología oscura de extracción de datos es valiosa. Además de los datos oscuros en sí, las tecnologías de extracción de datos oscuros son extremadamente valiosas.

### El valor de Dark Data

El primer paso para entender su valor es identificar qué información contienen, dónde se encuentra y el estado actual respecto a su exactitud, antigüedad, etc. Llegar a este estado exige de los datos:

- Analizarlos para entender cuántos datos se tienen, dónde se encuentran y cuántos tipos hay.
- Categorizarlos según la naturaleza general de la información que contienen.
- Clasificar la información en función de lo que pasará después. ¿Se utilizará? ¿Se destruirá? ¿Se estudiarán en mayor profundidad estos datos? Una vez tomadas estas decisiones, se pueden enviar los grupos de datos a distintos lugares para aislar la información que se quiera analizar con mayor detalle.

Una vez identificado el contexto relativo de los grupos de datos, hay que centrar en los datos que consideremos que pueden proporcionar conocimiento. También se tendrá una imagen más clara de todo el entorno de los datos relativos a la entidad para poder establecer políticas de gobernanza de la información que aliviarán la carga de los Datos oscuros, mientras empiezan a trabajar.

### Amenazas para el usuario - Contras

- Ignorancia respecto a lo que se tiene y no se usa y aprovecha.
- Costos de implementación de software o hardware.
- Ciberseguridad de datos oscuros antiguos.
- Los datos oscuros sobre sus clientes y operaciones almacenados en la nube, fuera de su control y gestión inmediatos que pueden tener los siguientes riesgos:
  - Riesgo legal y regulatorio: Datos cubiertos por mandato o regulación.
  - Riesgo de inteligencia: Información confidencial.
  - Riesgo de reputación: Cualquier tipo de violación de datos afecta a la imagen.
  - Costos de oportunidad: Decidir no invertir en el análisis y minería de datos oscuros, representa posibles pérdidas de inteligencia y valor en función de su contenido.
  - Exposición abierta: Por definición, los datos oscuros contienen información que es muy difícil o costosa de extraer para ser extraída, o que contiene fuentes de inteligencia desconocidas (y por lo tanto no evaluadas) y la exposición a pérdida o daño. Los secretos de los datos oscuros pueden ser muy oscuros y dañinos, pero uno no tiene forma de saberlo con certeza. Esto no puede cultivar la complacencia o la indiferencia en aquellos que contemplan esos riesgos en serio.

Hay estrategias y tecnologías útiles que pueden proporcionar cierto grado de protección contra tales riesgos tales como<sup>10</sup>:

- Inventario en curso y evaluación.
- Encriptación ubicua.
- Políticas de retención y eliminación segura.
- Auditoría de datos oscuros por razones de seguridad.

<sup>10</sup> <https://www.cio.com/article/2686755/data-analytics/the-dangers-of-dark-data-and-how-to-minimize-your-exposure.html>

## 5. CAUSANTES DEL CAMBIO

- La tendencia digital Big Data y la analítica de datos.
- La tendencia del Internet de las Cosas que traerá grandes volúmenes de datos de todas las cosas.
- La tendencia de medir y sensorizar los procesos.
- Auge de la concientización sobre el aprovechar los datos que se capturan en las entidades.
- El aumento de la demanda de las personas, las comunidades y los ciudadanos de participar en el control, planeación y ejecución de proyectos.

## 6. RETOS

“El principal reto que presenta el Dark Data no es solo el almacenamiento, sino determinar su valor real, si es que tiene alguno. De hecho, muchos datos no salen a la luz porque las organizaciones simplemente no saben lo que contienen. Destruirlos podría ser muy arriesgado, pero analizarlos puede ser costoso. Y resulta difícil justificar ese gasto si se desconoce el valor potencial de los datos. Para determinar si el Dark Data merece un análisis ulterior, las organizaciones necesitan contar con medios de selección rápida y rentable, que estructuren y visualicen los datos. Un hecho importante a la hora de comprender el Dark Data es entender que no es un acontecimiento aislado.” opina Ahmed Banafa en su artículo para BBVA Open Mind<sup>11</sup>.

## 7. INSIGHTS Y/O NECESIDADES BÁSICAS

- **Transparencia:** Una organización transparente es aquella que hace pública su información, generando confianza en la ciudadanía.
- **Optimización del gasto (reducción de costos):** Aprovechar las herramientas TIC para reducir los costos de las operaciones y colocar esos recursos liberados en áreas o proyectos más interesantes para la entidad.
- **Necesidad de información relevante:** Determinar las necesidades de los usuarios es importante para conocer sus deseos o necesidades.
- **Volumen:** Tamaño de los datos.
- **Velocidad:** A la que se generan los datos.
- **Variedad:** Los diferentes tipos.
- **Veracidad:** Confiabilidad de los datos.
- **Valor:** Deben generar valor para ser útiles.
- **Conocimiento:** Conocer lo que se tiene y no se usa y reducir esa incertidumbre.

## 8. APLICACIONES EN LA ALCALDÍA DE SANTIAGO DE CALI

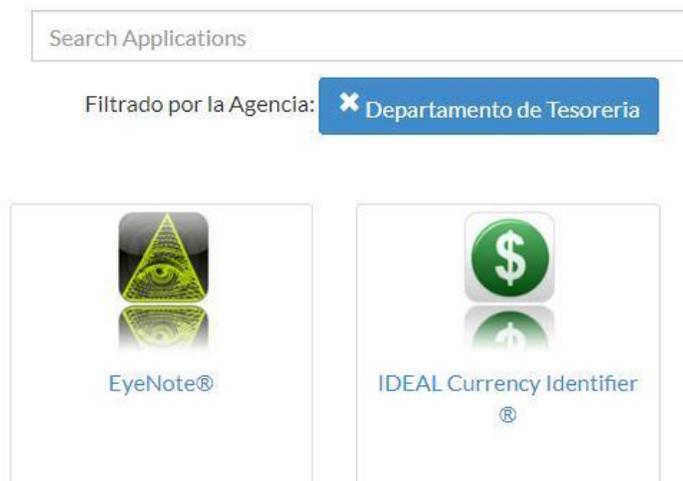
### **Apps para informar y mejorar las experiencias de los ciudadanos.**

Algunos ejemplos de Estados Unidos<sup>12</sup>

<sup>11</sup> <https://www.bbvaopenmind.com/entender-el-dark-data/>

<sup>12</sup> <https://www.data.gov/applications?>

La información gubernamental abierta impulsa las aplicaciones de software que ayudan a las personas a tomar decisiones informadas, desde elegir opciones de ayuda financiera para la universidad hasta encontrar los productos y vehículos de consumo más seguros. A continuación se incluyen algunos ejemplos de herramientas gubernamentales, comunitarias y comerciales que: usan datos gubernamentales abiertos de los Estados Unidos; son accesibles, verificables y disponibles; y son, para la mayoría, gratuitos y no requieren el registro para su uso.



**Figura 1:** Apps de datos abiertos del Departamento de Tesorería de Estados Unidos.

Fuente: data.gov

### En la Alcaldía estamos llegando a 500 teras de datos:

El Informe Global Databerg identificó que el 15% de la información de una organización podría clasificarse como Datos Críticos del negocio, el 33% como Datos ROT y el 52% como Datos Oscuros. Aplicando la relación de este informe, se proyecta que la clasificación de los datos de la Alcaldía sería:

15%	75 TB	Datos Críticos del negocio
33%	165 TB	Datos ROT
52%	260 TB	Datos Oscuros

Total de 500 TB de datos almacenados

Según la Evaluación de datos oscuros de Veritas<sup>13</sup>, esto representa para la Alcaldía:

- 75 TB de Datos Críticos del negocio costarían USD \$375.000
- 165 TB de ROT Data costarían USD \$825.000
- 260 TB de Datos Oscuros costarían USD \$1.300.000

Otra vez según la Evaluación de datos oscuros de Veritas, en 5 años habrá crecido...

<sup>13</sup> <https://veritas.valuestoryapp.com/DarkData/#>



Figura 2: Proyección del crecimiento de los datos almacenados en la Alcaldía.

### Departamento Administrativo de Planeación Municipal y Oficina Asesora de Transparencia

Estos organismos, como líderes del gobierno del dato y de la transparencia respectivamente, tienen una gran oportunidad de hacer visible de ese 65% de datos oscuros, la información municipal que pueda transformarse en datos abiertos para la ciudadanía, para impulsar el mejoramiento de los procesos, de los servicios y la transformación digital de la Alcaldía.

## CONCLUSIONES

Cuando se utilizan de manera efectiva, el análisis de datos puede ayudar a salvar vidas, mejorar la eficiencia, reducir costos y permitir que el municipio brinde mejores servicios a los ciudadanos. Pero cuando los datos útiles y relevantes utilizados para hacer esas mejoras quedan eclipsados por los datos oscuros, las eficiencias ya no existen y las decisiones pueden complicarse. A medida que crece la importancia del análisis de datos en el gobierno, se vuelve cada vez más importante reducir la cantidad de datos oscuros que oscurecen los datos que proporcionan un valor real y una perspectiva.

El rápido ritmo de crecimiento de los datos de hoy proporciona a los organismos una visión sin precedentes de las necesidades de los ciudadanos a los que sirven y las misiones que deben cumplir. Sin embargo, con más datos surgen nuevos desafíos en cuanto a organizarlos, almacenarlos y protegerlos. También existe el desafío de administrar los datos oscuros que inevitablemente ocurren junto con el crecimiento de los datos y la información existentes. Al trabajar con los socios adecuados y contar con las estrategias correctas de gestión de datos y rendición de cuentas en torno a sus datos, los organismos pueden proporcionar los conocimientos de información que necesitan para cumplir con sus misiones, mientras se protegen de una crisis de información.

Adicionalmente hoy en día la transformación digital está siendo promovida por una cultura impulsada por los datos. ¿Y qué es lo que la impulsa? Son los datos. Estamos viviendo la revolución digital, por lo que la gestión inteligente de los datos correctos es la clave para desbloquear nuevas oportunidades. Finalmente, los datos oscuros que son descartados por un proceso u organismo, pueden ser muy valiosos para otro.

## BIBLIOGRAFÍA

- Analyzing dark data for hidden opportunities | Deloitte Insights. (s. f.). Recuperado 6 de julio de 2018, de <https://www2.deloitte.com/insights/us/en/focus/tech-trends/2017/dark-data-analyzing-unstructured-data.html>
- Applications. (s. f.). Recuperado 10 de julio de 2018, de <https://www.data.gov/applications>
- Cómo buscar oro eficazmente (Dark Data vs Big Data). (s. f.). Recuperado 6 de julio de 2018, de <https://www.businessintelligence.info/definiciones/big-data-vs-dark-data.html>
- Dark Data – Complexity Labs. (s. f.). Recuperado 6 de julio de 2018, de <http://complexitylabs.io/dark-data/>
- Dark Data - Gartner IT Glossary. (s. f.). Recuperado 6 de julio de 2018, de <https://www.gartner.com/it-glossary/dark-data/>

Dark Data Management: The Next Frontier for Government Data. (s. f.). Recuperado 9 de julio de 2018, de <https://issuu.com/govloop/docs/dark-data-management>

Dark Data: tu empresa tiene «datos oscuros» que puedes convertir en oro - Gestión - Emprendedores - Webs. (s. f.). Recuperado 6 de julio de 2018, de <http://www.emprendedores.es/gestion/que-es-dark-data-utilidad-empresa>

Data-Driven Development - Reports - World Economic Forum. (s. f.). Recuperado 6 de julio de 2018, de <http://reports.weforum.org/data-driven-development/the-role-of-communities-and-the-individual-digital-identity-and-human-centred-design/>

Datumize. (s. f.). Datumize. Recuperado 9 de julio de 2018, de <https://www.datumize.com>

datumize. (s. f.). The Evolution of Dark Data and how you can harness it to make your business Smarter. Recuperado 9 de julio de 2018, de <https://blog.datumize.com/evolution-dark-data>

Digging up dark data: What puts IBM at the forefront of insight economy | #IBMinsight - SiliconANGLE. (s. f.). Recuperado 6 de julio de 2018, de <https://siliconangle.com/blog/2015/10/30/ibm-is-at-the-forefront-of-insight-economy-ibminsight/>

Entender el Dark Data - OpenMind. (s. f.). Recuperado 6 de julio de 2018, de <https://www.bbvaopenmind.com/tecnologia/mundo-digital/entender-el-dark-data/>

Harvard Business Review: From data to disruption: Innovation through digital intelligence. (s. f.). Recuperado 6 de julio de 2018, de <https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=IML14573USEN>

How data becomes knowledge, Part 3: Extracting dark data. (s. f.). Recuperado 6 de julio de 2018, de <https://www.ibm.com/developerworks/library/ba-data-becomes-knowledge-3/index.html>

How to Create a Data Strategy for Machine Learning-Powered Artificial Intelligence. (s. f.). Recuperado 6 de julio de 2018, de <https://www.gartner.com/doc/3734817?ref=unauthreader&srcId=1-3478922254>

Shedding light on dark data | CIO. (s. f.). Recuperado 6 de julio de 2018, de <https://www.cio.com/article/2926089/data-analytics/shedding-light-on-dark-data.html>

The API Briefing: the Challenge of Government's Dark Data / DigitalGov - Building the 21st Century Digital Government. (s. f.). Recuperado 6 de julio de 2018, de <https://digital.gov/2015/06/17/the-api-briefing-the-challenge-of-governments-dark-data/>

The Dangers of Dark Data and How to Minimize Your Exposure | CIO. (s. f.). Recuperado 6 de julio de 2018, de <https://www.cio.com/article/2686755/data-analytics/the-dangers-of-dark-data-and-how-to-minimize-your-exposure.html>

The Data Briefing: Tales from the Dark Side of Data / DigitalGov - Building the 21st Century Digital Government. (s. f.). Recuperado 6 de julio de 2018, de <https://digital.gov/2016/04/27/the-data-briefing-tales-from-the-dark-side-of-data/>

Todo lo que necesitas saber para sacar partido al Dark Data en tu negocio - Gestión - Emprendedores - Webs. (s. f.). Recuperado 6 de julio de 2018, de <http://www.emprendedores.es/gestion/dark-data-utilidad-empresa>

## CRÉDITOS

### Aprobado por

Esteban Oliveros Montoya  
Subdirector de Innovación Digital  
Departamento Administrativo de las Tecnologías de la Información y las Comunicaciones  
Alcaldía de Santiago de Cali - Colombia

### Revisado por

Tulio Enrique Tafur  
Contratista  
Departamento Administrativo de las Tecnologías de la Información y las Comunicaciones

### Elaborado por

Carlos Eduardo Jiménez Bacca  
Contratista  
Departamento Administrativo de las Tecnologías de la Información y las Comunicaciones

Jimmy Fernando Montaña Hernández  
Contratista  
Departamento Administrativo de las Tecnologías de la Información y las Comunicaciones