

5

DESCRIPCIÓN

PROCEDIMIENTOS PARA DATOS ABIERTOS EN TIEMPO REAL

Estado de la técnica:

10

Esta tecnología pertenece al sector tecnológico de la inteligencia artificial, las TICs y las telecomunicaciones.

15

Para los funcionarios públicos encargado de publicar datos abiertos estatales les resulta muy engorroso, demorado y complejo el poder llevar a cabo todo el proceso para recolectar la información, luego organizarla, quitarle los datos personales, luego pasarla a formatos procesables, luego verificarla para no cometer errores legales de publicaciones indebidas y también describirla adecuadamente para su identificación y utilidad cuando sea publicada.

20

25

Existen entidades públicas que producen millones de datos cada día, datos que son muy útiles para tomar decisiones si fueran accesibles y abiertos en tiempo real, pero los funcionarios públicos se demoran meses en poder publicar y actualizar los datos abiertos debido a la complejidad de procesos previos de alistamiento, descripción, caracterización y verificación de los datos antes de entrar a la plataforma de publicación que dispone el estado y usando credenciales cargar los archivos con especificaciones especiales.

30

35

Existen muchos datos que provienen por ejemplo de los semáforos inteligentes y los sensores que estos incorporan, estos son equipos que entregan miles de datos por minuto de variables como velocidad, riesgos viales, peatones, vehículos, bicicletas, condiciones ambientales en el sitio, daños de los equipos, sobre paso de semáforos en rojo, excesos de velocidad en el paso de semáforo, placas de los vehículos, peatones que no respetan las señales de semáforos, fallas eléctricas, no respeto de la cebrá etc. Muchas señales viales inteligentes registran millones variables desde el sitio en el que las variables se producen y una ciudad puede tener miles de estas señales inteligentes fuentes de datos abiertos que pueden ser muy útiles para temas de seguridad vial y movilidad si estuvieran disponibles como datos abiertos en tiempo real. El objetivo de los datos abiertos es que estén disponibles para todo el mundo en la plataforma del gobierno,

pero el proceso para hacer que eso pase tiene varias complicaciones para poder
40 entregarlos como abiertos, cumplir las normas y todo el proceso que eso requiere y mas
difícil aun hacerlo en tiempo real y sobre datos asociados a riesgo vial por ejemplo.

El Problema existente que se busca resolver:

45 Existe una debilidad e incapacidad humana frente a la cantidad de información millones
de datos que se generan en algunos procesos del estado y que la ley obliga a que esos
datos sean publicados como datos abiertos según ley 1712 de 2014 y subsiguientes., son
millones de datos que un funcionario público encargado del proceso debe documentar,
caracterizar, comentar, organizar en formatos procesables y verificarlos que cumplan con
50 la norma legal, luego entrar a una plataforma del estado y publicar los archivos de datos
abiertos de manera organizada y entendible para quien quiera consultarlos. Millones de
esos datos diarios provienen de otros procesos, otras dependencias o de otros
dispositivos de borde como señales viales inteligentes como los semáforos que toman
datos de las vías. Se hace casi imposible que muchos de estos procesos que generan
55 millones de datos diarios puedan ser atendidos, revisados, comentados, estructurados y
publicados en tiempo real por un funcionario. Lo que hace que las frecuencias de
publicación y actualización sea de semanas, meses o años y demande mucho esfuerzo
humano. Aun con el esfuerzo humano de publicación de datos abiertos, menos del 1% de
los datos abiertos son utilizados por terceros, no son conocidos, no están bien
60 documentados, no están licenciados para ser usados legalmente, no son útiles por la
forma en la que fueron elaborados y finalmente no cumplen el propósito de que los datos
abiertos sean de uso masivo de terceros.

65 Como esta invención resuelve el problema anterior:

Esta invención proporciona al menos cuatro procesos que se ejecutan en sistemas
informáticos locales o distribuidos en redes públicas y/o privadas y desde señales viales
como semáforos, procesos para automatizar, controlar, auditar, verificar, documentar,
70 estructurar, licenciar y publicar datos abiertos en tiempo real sin intervención humana, así
como las respectiva publicidad de los conjuntos de datos publicadas en redes sociales y
mensajería directa a los interesados mediante el uso de inteligencia artificial para que los

datos abiertos cumplan su objetivo de llegar a los interesados en esos datos y ser utilizados, aprovechados al máximo por terceros.

75

Para resolver el problema se dispone de un Procedimiento para DOCUMENTAR conjuntos de datos el cual se ejecuta en medios informáticos remotos o locales con al menos un procesador donde se ejecuta el procedimiento que lee diferentes fuentes de información como urls, escaneada, obtenida por cámaras, digitada, obtenida por Apis y/o bases de datos locales o remotas, luego con la información obtenida se crea los metadatos para un conjunto de datos o una pluralidad de conjuntos de datos los cuales contienen los aspectos básicos para entender el contenido, generada por inteligencia artificial las características que describen cada conjunto de datos y luego los guarda en archivo digital o base de datos local o remota.

85

Además el procedimiento para documentar los datos comprende los metadatos con aspectos como nombre del conjunto de datos, descripción del conjunto de datos, la entidad o fuente de los datos, la categoría de los datos que pueden ser de diferentes aéreas como la salud, el transporte, seguridad vial, cultura, educación y otra pluralidad de categorías o áreas. También comprende los metadatos sobre la fecha de publicación, la periodicidad con la que se actualiza el conjunto de datos, el tipo de licencia bajo la cual se publica el conjunto de datos y otra pluralidad de metadatos generados automáticamente con inteligencia artificial.

90

95

Otro Procedimiento se dispone para ESTRUCTURAR conjuntos de datos y hacerlos procesables por un sistema informático, en el cual primero se toman los datos en una pluralidad de formatos físicos o digitales y de una pluralidad de fuentes de datos, luego transforma esos datos a formatos que permiten manipulación y conexión con otros datos para interoperabilidad en lenguaje común.

100

Este procedimiento para estructurar los datos comprende archivos digitales con bases de datos remotas o locales con formatos de datos tubulares como CSV, formatos de datos estructurados como JSON, XML y otra pluralidad de formatos estructurados., también comprende formatos de datos espaciales como SHP, GeoJSON, WMS, WFS, KML, GeoTIFF y otra pluralidad de datos espaciales.

105

Otro Procedimiento se dispone para PUBLICAR conjuntos de datos en el cual primero se verifican en sistema informático las credenciales de acceso previamente creadas y/o autorizadas para acceder a conjuntos de datos y publicarlos, se toman los datos
 110 previamente documentados y estructurados extraídos de en un archivo digital o base de datos local o remota, se envían las credenciales de acceso al portal o plataforma de datos donde quedaran publicados.

El procedimiento para publicar los datos automáticamente sin intervención humana,
 115 comprende un seleccionador automático de archivos de carga, una base de datos de usuarios, perfiles, restricciones y reglas de publicación, un archivo de instrucciones de procedimiento de conexión, direcciones de conexión, métodos de ingreso y pluralidad de formas de publicación, además del tipo de licencia de la base de datos a publicar y otros controles de legalidad configurables por el usuario desde una interfaz grafica o mediante
 120 comandos de voz. Usando conexiones inalámbricas para los proceso remotos.

El procedimiento para publicar los datos automáticamente usa inteligencia artificial y sin intervención humana selecciona los posibles interesados en los datos abiertos, publica en redes sociales, envía correos y mensajería de los análisis de Inteligencia Artificial que se
 125 hacen sobre los metadatos publicados, indicando predicciones, tendencias, estadísticas y alertas en tiempo real a modo de resumen útil de los conjuntos de datos publicados, para motivar el uso y consulta de los datos abiertos por parte de diferentes actores y en tiempo real.

130 El cuarto Procedimiento para publicar datos abiertos en tiempo real desde dispositivos de borde como semáforos y señales viales se ejecuta en al menos un procesador local donde se ejecuta el procedimiento de datos abiertos, conectado a una pluralidad de sensores y fuentes de datos digitales o análogos. Que se conectan a redes inalámbricas mediante una pluralidad de formas y genera datos abiertos verdes a partir de células
 135 fotovoltaicas. La definición de datos abiertos verdes, son datos abiertos ecológicos producidos sin el impacto ambiental ni la generación de gases efecto invernadero asociados al uso de fuentes de energía convencionales que para la generación de energía tienen impacto en el medio ambiente, como las hidroeléctricas y otras que impactan en ecosistemas acuáticos y en los ecosistemas por donde pasan todas las
 140 redes de transición de energía. Los datos abiertos verdes producidos con la energía

renovable de células fotovoltaicas en sitio, son datos producidos ecológicamente haciendo la transición energética al uso de energías limpias en la generación, transmisión y publicación de datos abiertos verdes producidos sin impacto ambiental de la energía tradicional.

145

El procedimiento para publicar datos abiertos en dispositivo de borde, comprende un dispositivo IoT para conectarse a internet u otras redes usando una pluralidad de formas de comunicación electromagnética, enviando una trama de datos que contiene al menos el identificador del dispositivo, lugar GPS, el tipo de dato, la fecha y hora.

150

El procedimiento para publicar datos abiertos en tiempo real desde dispositivos de borde comprende al menos un medio de almacenamiento local, donde primero almacena los datos de sensores y entrada digitales, luego el procesador los estructura y finalmente los publica automáticamente sin intervención humana como datos abiertos en tiempo real de variables viales, seguridad, riesgo vial, salud, comerciales y otra pluralidad de variables.

155

Ventajas de esta invención:

- Aumenta el conocimiento de los datos publicados por parte de terceros interesados.

160

- Aumenta el uso de datos abiertos por parte de terceros, mediante publicidad automática con inteligencia artificial.

- Permite controlar, auditar, verificar, documentar, estructurar, licenciar y publicar datos abiertos en tiempo real sin intervención humana.

165

- Disminuye los tiempos de publicación de datos abiertos y la cantidad de datos que puede manejar un humano.

170

- Permite licenciar el uso de datos abiertos para su publicación de manera automática.

- Desde señales viales, semáforos permite en tiempo real controlar, auditar, verificar, documentar, estructurar, licenciar y publicar datos abiertos sin intervención humana.

175 **Descripción de los dibujos:**

Se muestra en Dibujo 1 la imagen de diagrama de bloques de la invención.

Mejor manera de realizar la invención:

180

Esta invención se realiza mediante procedimientos que se cargan en un sistema informático local o distribuido en una red pública y/o privada. Los procedimientos los ejecuta al menos un procesador existente en el sistema informático. Estos procedimientos escritos se almacenan en una unidad de memoria local o remota de
 185 donde el procesador los lee para ejecutarlos. A nivel de dispositivo de borde en una señal vial semáforo que mide datos de riesgo vial, integra al menos un procesador donde se ejecuta el procedimiento para obtener datos de riesgo vial y estado de semáforo mediante sensores de velocidad, de ToF de vehículos y peatones, de vehículos, unidad de almacenamiento, cámaras y sistema de comunicación que trasmite los datos abiertos de
 190 riesgo vial identificados por el semáforo en tiempo real. Datos de riesgo de atropellar a peatón según velocidad, de no respetar la cebra los vehículos, de daños del semáforo, de paso en rojo, de exceso de velocidad en paso por el semáforo etc. Datos que procesa el propio semáforo y que puede publicar como datos abiertos en tiempo real.

195 **Ejemplo de aplicación:**

Una ciudad requiere publicar datos abiertos en tiempo real de cuál es el nivel de riesgo de que peatones sean atropellados en los 1500 cruces semaforizados de la ciudad Bogotá, por vehículos con exceso de velocidad que no respetan el semáforo y cuál es la placa de
 200 los vehículos que mayor riesgo representan.

En este caso la invención le puede decir a varias instituciones que miren el conjunto de datos abiertos publicados sobre el nivel de riesgo de atropello de peatones en las 1500 intersecciones de semáforos de Bogotá., mediante publicidad directa a compañías de
 205 seguros, autoridades de tránsito, aplicaciones para conductores, sistemas de transporte público, dueños de los vehículos que desconocen su uso, asociaciones de personas con discapacidad, víctimas de accidentes, organizaciones de prevención vial, desarrolladores de aplicaciones para personas vulnerables etc. La invención con inteligencia artificial

210 produce los datos abiertos en tiempo real, licencias, verificaciones, publicaciones y
publicidad a los interesados en los datos abiertos para asegurar su conocimiento, su uso y
el impacto de los datos en la sociedad. Un humano sería incapaz de lograr estas
publicaciones de datos abiertos en tiempo real sobre los riesgos en 1500 sitios de la
ciudad. Si la aseguradora conociera que su cliente está en riesgo o es un conductor
215 peligroso, o el dueño del vehículo o empresa de transporte supiera que su conductor
genera riesgos de accidente, sería información valiosa., igual que los peatones que
podrían saber los sitios donde mayor atención deben prestar por el nivel de riesgo de ser
atropellados etc.

Otro ejemplo es que una entidad manda 1000 encuestadores a los barrios para saber en
220 esas zonas quienes tienen acceso a servicios de salud y qué tipo de servicio. Son 1000
formatos con diferentes campos de información, alguna reservada y no publicable por ser
personal. Con la información dispuesta en el DANE y la información de las encuestas
anteriores se debe publicar datos abiertos de cómo ha cambiado la cobertura en salud.

225 A un funcionario público le tomaría días recolectar la información de las encuestas y
organizarla, recolectar la información del DANE y organizarla, recolectar la información de
encuestas anteriores y organizarla, hacer los análisis de la información a publicar, realizar
los comentarios a cada conjunto de datos de interés, poner todos los datos en un mismo
formato procesable, verificar los datos, generar la licencia de que datos son usables por
230 terceros y bajo qué condiciones, organizar los archivos para cargar al sistema nacional de
datos abiertos y finalmente entrar a la plataforma y publicarlos. De allí en adelante motivar
a terceros a su uso, identificar que terceros serían los interesados en esos datos y
contactarlos para que entren a consumir los datos abiertos publicados indicándoles que
beneficio tienen para ellos. Este proceso lo haría en posos segundos la presente
235 invención.