



# BIG DATA

## FOR SOCIAL GOOD

### Estudio de caso: Telefónica Brazil

## Predecir niveles de contaminación del aire de 24 a 48 horas antes en São Paulo, Brasil

Telefónica

LUCA  
Telefónica DATA UNIT

vivo

La GSMA representa los intereses de los operadores móviles de todo el mundo y reúne a casi 800 operadores móviles con más de 300 compañías de otras actividades del ecosistema móvil, entre las que se cuentan fabricantes de terminales y dispositivos, compañías de software, proveedores de equipo y compañías de Internet, además de organizaciones de otros sectores adyacentes.

La iniciativa Big Data for Social Good reúne a organizaciones públicas y privadas para acelerar el impacto de la industria móvil sobre los Objetivos de Desarrollo Sostenible de la ONU.

Las enfermedades contagiosas, la contaminación, los terremotos, las inundaciones y otros desastres son algunos de los principales desafíos que el mundo enfrenta actualmente. De acuerdo con el Banco Mundial, la contaminación del aire ha surgido como el cuarto factor de riesgo de muerte en todo el mundo.

Los operadores móviles pueden ofrecer información significativa y única basada en datos de red anonimizados

y agregados para ayudar a resolver estos problemas complejos. Los datos de movilidad pueden ayudar a las organizaciones de salud pública a responder de forma más efectiva a las epidemias o planificar intervenciones específicas de salud. Puede ayudar a las agencias de alivio de emergencias a dirigir sus recursos de forma más precisa y eficiente. Puede ayudar a los gobiernos a comprender mejor el impacto de la contaminación y el cambio climático en los ciudadanos.

Mediante la GSMA, los operadores móviles de las diferentes geografías se han reunido para acelerar y escalar la oportunidad de Big Data for Social Good. La GSMA ofrece una plataforma única para establecer un marco común y enfoques de mejores prácticas, al tiempo que respeta y protege la privacidad de las personas. El caso de estudio del proyecto ha sido desarrollado como parte de la iniciativa de GSMA Big Data for Social Good, para promover la implementación de modelos similares en otras grandes ciudades del mundo.

1. <http://www.worldbank.org/en/news/infographic/2016/09/08/death-in-the-air-air-pollution-costs-money-and-lives>

## Resumen

Telefónica Brasil (Vivo) está en contacto con las municipalidades de São Paulo para utilizar los datos de las redes móviles para ayudar a combatir el impacto adverso en la salud de la contaminación del aire. Telefónica utiliza datos móviles anonimizados, algoritmos de machine learning y herramientas, que las municipalidades de São Paulo pueden integrar en sus procesos de gestión del tránsito y la contaminación. Desarrollados por el equipo de LUCA, la unidad de datos de Telefónica, los algoritmos usan aprendizaje automático (*machine learning*) y datos anonimizados de la red móvil, combinados con datos de sensores del clima, del tránsito y de polución, para monitorizar y predecir niveles de polución en toda la ciudad.

La solución puede predecir niveles de contaminación con 24 a 48 horas de anticipación, y permitir a las autoridades locales de São Paulo tomar medidas preventivas si las

emisiones de dióxido de nitrógeno (NO<sub>2</sub>) pudieran poner en peligro la salud humana. Al emplear macrodatos (*Big Data*) móviles, Telefónica y las municipalidades pueden reducir significativamente los costes, aumentar la precisión y proporcionar información valiosa a los encargados de tomar decisiones sobre cómo definir las políticas de transporte humano o evaluar los niveles de contaminación.

La implementación en São Paulo ha demostrado que los macrodatos móviles son una poderosa herramienta para predecir y analizar los patrones de tránsito y la contaminación resultante del aire. Los datos móviles anonimizados permiten una monitorización más granular y rentable de la contaminación de la ciudad complementaria a una red de sensores estáticos que recopila datos de contaminación.

## El desafío de la contaminación del aire

Múltiples estudios han demostrado la relación entre la contaminación ambiental y los problemas de salud pública. La Organización Mundial de la Salud ha encontrado que la contaminación del aire se asocia con enfermedades crónicas y exacerba las patologías relacionadas, como la morbilidad y mortalidad cardiovascular y respiratoria<sup>2</sup>. Un estudio en Brasil<sup>3</sup> estimó que más de 20 000 muertes al año, en 29 regiones metropolitanas brasileñas, se pueden atribuir a la contaminación del aire.

A medida que hay más datos ambientales disponibles, existe una tendencia global hacia enfoques más orientados a los datos para manejar la contaminación del aire y su impacto, en línea con los objetivos de desarrollo sostenible (ODS) establecidos por las Naciones Unidas:

- **ODS 3** Garantizar una vida sana y promover el bienestar para todos en todas las edades: El ODS 3 busca aumentar la expectativa de vida, en parte mediante medidas para fortalecer la capacidad de los países, en particular los países en desarrollo, para una advertencia temprana, reducción del riesgo y gestión de riesgos nacionales y mundiales de salud.
- **ODS 13** Adoptar medidas urgentes para combatir el cambio climático y sus efectos.
- **ODS 11** Lograr que las ciudades y los asentamientos humanos sean inclusivos, seguros, resilientes y sostenibles: El ODS 11 busca reducir el impacto ambiental adverso per cápita de las ciudades para 2030, en parte prestando atención especial a la calidad del aire y la gestión de desechos.

El mapeo de precisión de los niveles de contaminación en las áreas metropolitanas se puede utilizar para:

- Identificar los potenciales niveles pico de contaminación local que podrían afectar la salud de la población local y, a más largo plazo, vincular esta información con otros datos de la población (por ej., asma) para informar proactivamente a los ciudadanos sobre el impacto específico.
- Informar a los sistemas de gestión de tránsito para reducir la cantidad y velocidad de vehículos en áreas con niveles elevados o peligrosos de contaminación.

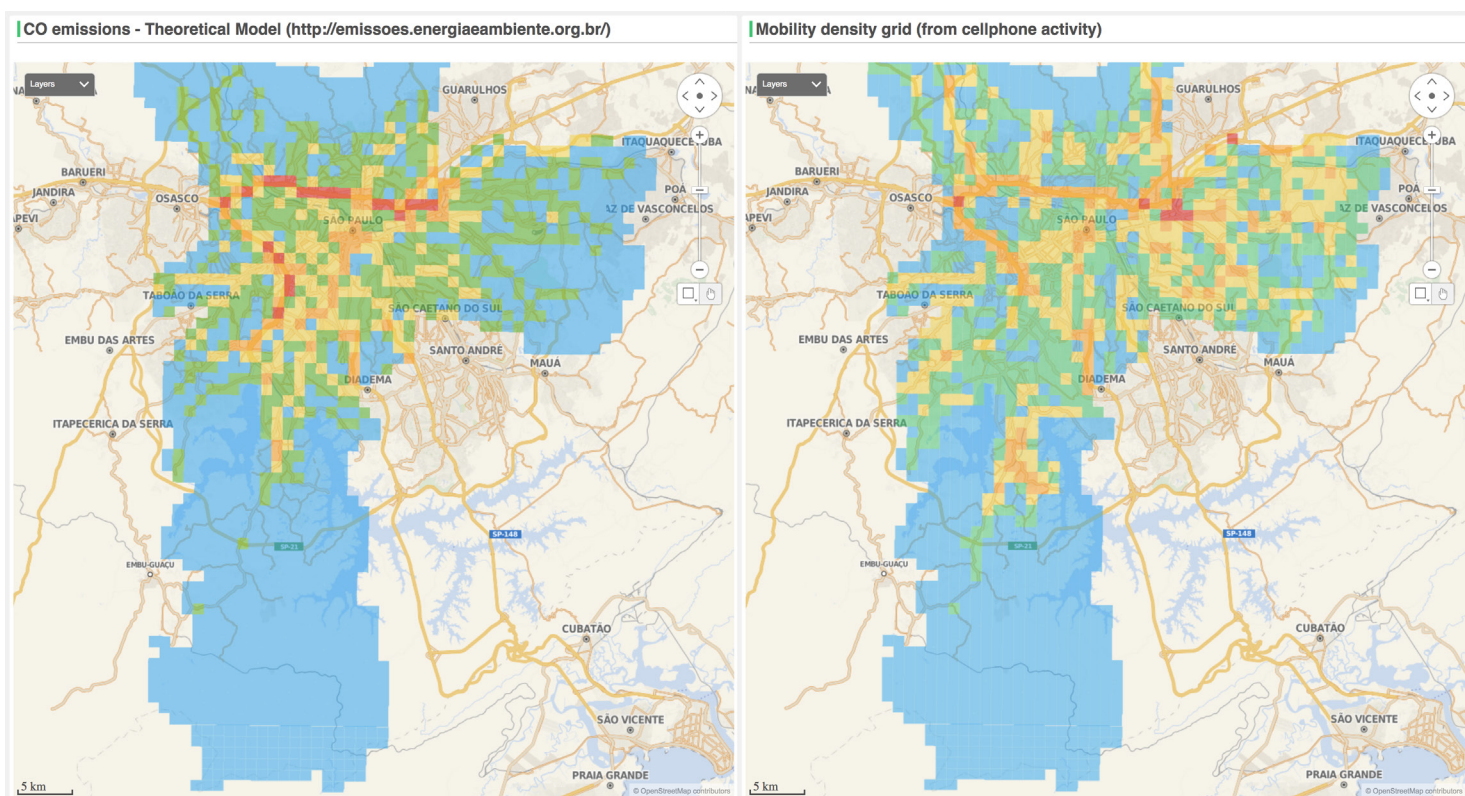
Sin embargo, la gestión de la calidad del aire puede ser costosa y el análisis de tránsito puede ser relativamente manual, usando datos de entrevistas in situ y contadores manuales. Existe una infraestructura limitada para el monitoreo de la calidad del aire en la mayoría de las ciudades debido a los altos costes de instalación y operación de las estaciones fijas de medición. Esto significa que los datos son dispersos, lo que dificulta que las administraciones públicas recopilen los datos precisos que necesitan para identificar la ubicación de los principales puntos de contaminación del aire.

En Sao Paulo, el gobierno local utiliza un modelo teórico de emisiones de CO<sub>2</sub> basado en mediciones de tráfico limitadas a horas fijas del día. Los datos móviles muestran una gran concordancia con este modelo y pueden generalizarse y perfeccionarse para ofrecer predicciones más útiles y ajustadas a la realidad (ver la Figura 1).

2. <http://www.who.int/airpollution/en/>

3. Miraglia S., Gouveia N. Costs of air pollution in Brazilian metropolitan regions. Cienc. Saude Colet. 2014;19:4141-4147

**Figura 1** El tablero de Telefónica muestra una fuerte concordancia entre las mediciones de tráfico realizadas a partir de datos móviles (derecha) y las estimaciones teóricas de emisiones de CO del gobierno.



## Cómo pueden ayudar los datos móviles

Telefónica Brasil trabaja con las municipalidades de São Paulo para correlacionar la información de los sensores de calidad del aire, los sensores de tránsito y las estaciones climáticas con los datos generados por las redes móviles. Tal enfoque puede generar información de forma más rentable que la observación directa de los niveles de contaminación, y de forma más regular y dinámica que los métodos más tradicionales de recopilación de datos.

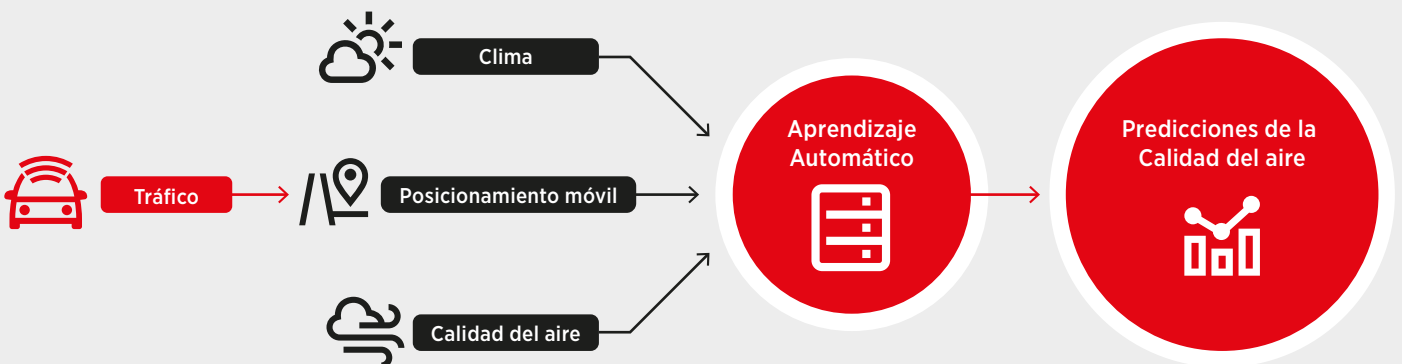
Para compensar el número limitado de sensores en la ciudad, Telefónica utiliza los datos de movilidad como representantes de las observaciones del volumen y flujo de vehículos. Con su plataforma LUCA, la unidad de datos de Telefónica ha creado un conjunto de algoritmos para calcular estimaciones de movilidad y tránsito, y un tablero que presenta los niveles de contaminación de toda la ciudad. Los algoritmos usan aprendizaje automático y datos del clima, el tránsito y la contaminación para extrapolar el mapeo a toda la ciudad (vea la Figura 2).

Los datos móviles de Telefónica pueden abarcar eventos activos, como una llamada de voz o SMS, y eventos pasivos, como dispositivos sincronizados con torres de celulares. Estos datos se pueden agregar y escalar en un índice para representar a toda la población de la ciudad. De acuerdo con Telefónica, este índice en general ha estado altamente correlacionado (hasta el 94 %) con el tránsito real observado. Para proteger la privacidad de las personas, se logra la anonimización al eliminar los MSISDN, añadir ruido, agregar y extrapolar a partir de la participación del mercado de Telefónica a toda la población y normalizar dichos datos<sup>4</sup>.

Telefónica también está probando la solución de predicción de contaminación en Madrid para crear un enfoque robusto que podría escalarse a diferentes geografías.

4. Por ej., proporcionando desviaciones de la movilidad media en cada ubicación - Telefónica computa el índice de movilidad de cada una de las estaciones de medición. La correlación oscila entre el 89 % y el 97 %, de acuerdo con la estación\*. El 94 % es el promedio entre las diferentes correlaciones.

**Figura 2 Combinación de los datos de contaminación, datos de tránsito y datos de telecomunicaciones para mapear y predecir la calidad del aire.**



Los datos de posicionamiento móvil se utilizan como un proxy para los datos de tráfico, lo que permite medir el tráfico en toda la ciudad en lugar de las ubicaciones de la red limitada de sensores de tráfico. Los datos móviles, junto con los datos meteorológicos y de calidad del aire históricos, se combinan en un modelo de aprendizaje automático para generar predicciones sobre la calidad del aire.

## Impacto de la solución de datos móviles

La implementación en São Paulo ha demostrado que los macrodatos móviles son una poderosa herramienta para predecir y analizar los patrones de tránsito y la contaminación del aire en una ciudad. En general, los datos móviles permiten una monitorización más granular de la contaminación en la ciudad que una red de costosos sensores estáticos.

La implementación de São Paulo destaca cómo los datos de las redes móviles, correlacionados con la calidad del aire, el clima y otras fuentes de datos, pueden ayudar a predecir niveles de contaminación con 24 a 48 horas de

anticipación, y permitir a las autoridades locales tomar medidas preventivas si los niveles de NO<sub>2</sub> pudieran poner en peligro la salud humana. Tal enfoque puede reducir significativamente los costos, aumentar la precisión y proporcionar información valiosa a los encargados de tomar decisiones sobre cómo definir las políticas de transporte humano o evaluar los niveles de contaminación. Asimismo, los terceros podrían usar una API para obtener acceso directo a datos agregados y anonimizados de telecomunicaciones e información de movilidad para cumplir con sus propias necesidades.





## Desarrollo de una propuesta sostenible

Una lección clave aprendida es que para que este tipo de implementación sea sostenible, los gobiernos locales y los operadores móviles deben trabajar en conjunto, para combinar sus conjuntos de datos y analizar nuevos modelos que generen mejores predicciones, y definir políticas que podrían desarrollarse o informarse en base a estos datos.

Para desbloquear por completo el potencial de los grandes datos móviles, un paso necesario es que las administraciones públicas modifiquen su presupuesto

del proyecto anual para incluir la gestión de estos servicios e invertir en las habilidades relevantes. Esto ayudará directamente a reducir los costos relacionados con la atención médica, el desarrollo de infraestructura local y la movilidad urbana. Algunos administradores ya están reservando parte del presupuesto para las estadísticas de datos y los estudios de caso, incluidos específicamente los conjuntos de datos móviles que mejoran sus modelos<sup>5</sup>.

5. Plan de innovación para el transporte y las infraestructuras de 2017 – Ministerio de Obras Públicas y Transporte – Gobierno de España



## El futuro

Telefónica ha identificado una serie de posibles próximos pasos:



Mejorar las herramientas en términos de solidez y relevancia, p. ej. aprovechar las fuentes de datos adicionales dentro de los procesos operativos de las autoridades locales, y mejorar la calidad de los modelos predictivos.



Ayudar a las municipalidades a continuar optimizando la red de estaciones que controlan los niveles de contaminación y tránsito en las ciudades y más en general en el país para mejorar los algoritmos y las capacidades de predicción.



Desarrollar un mapa de ruta para transformar la solución de São Paulo en productos más escalables que luego podrían implementarse en otras geografías. Telefónica está desarrollando un enfoque similar para cubrir el municipio de Madrid, la capital de España.



Continuar desarrollando los flujos de ingresos de terceros para acceder a datos agregados de telecomunicaciones mediante una API para otros fines que requieran e involucren información de movilidad o ambiental para crear valor económico (por ej., campañas específicas de marketing) y social (por ej., preparación para desastres).

Mire nuestro video, aprenda más sobre la iniciativa y contacto nosotros para más información: [bd4sg@gsma.com](mailto:bd4sg@gsma.com)

[gsma.com/betterfuture/bd4sg](https://gsma.com/betterfuture/bd4sg)



@GSMA

#BetterFuture #BD4SG