



Westfield
Business
School



Maestría en

ADMINISTRACIÓN DE EMPRESAS

**Tesis previa a la obtención de título de
Magister en Administración de Empresas**

AUTORES: Ing. Lenin Fernando Bautista Moreta
Ing. Stefany Arline Aguilar Delgado
Lcda. Mónica Paola Medina Puetate
Ing. Karla Gabriela Suntasig Pachacama
Lcdo. Santiago David Carrillo Sánchez

TUTOR: Ing. José Luis Pérez

WayData Plataforma Analítica para Movilidad
Autónoma Sostenible

**PROYECTO DE TITULACIÓN – FIN DE MÁSTER
MAESTRÍA EN ADMINISTRACIÓN DE EMPRESA ENL**

WayData – Plataforma Analítica para Movilidad Autónoma

Por

Lenin Fernando Bautista Moreta

Stefany Arline Aguilar Delgado

Mónica Paola Medina Puetate

Karla Suntasig Pachacama

Santiago David Carrillo Sánchez

Febrero 2026

Aprobado

A handwritten signature in purple ink, appearing to read 'CRISTIAN', with a horizontal line underneath.

Cristian Melo
Presidente(a) del Tribunal
Universidad Internacional del Ecuador

Yo, Cristian Javier Melo González e Ignacio Maroto, declaramos que, personalmente conocemos que los graduandos: Lenin Fernando Bautista Moreta, Stefany Arline Aguilar Delgado, Mónica Paola Medina Puetate, Karla Suntasig Pachacama, Santiago David Carrillo Sánchez, son los autores exclusivos de la presente investigación y que ésta es original, auténtica y personal de ellos.



Cristian Melo
Coordinador MBA UIDE



Ignacio Maroto
Provost WBS

Autoría del Trabajo de Titulación

Nosotros, Lenin Fernando Bautista Moreta, Stefany Arline Aguilar Delgado, Mónica Paola Medina Puetate, Karla Suntasig Pachacama, Santiago David Carrillo Sánchez, declaramos bajo juramento que el trabajo de titulación titulado **WayData Plataforma Analítica para Movilidad Autónoma Sostenible**, es de nuestra autoría y exclusiva responsabilidad legal y académica; que no ha sido presentado anteriormente para ningún grado o calificación profesional, habiéndose citado las fuentes correspondientes y respetando las disposiciones legales que protegen los derechos de autor vigentes.

Lenin Bautista

Lenin Fernando Bautista Moreta

Correo electrónico: leferba114@gmail.com

Arline Aguilar

Stefany Arline Aguilar Delgado

Correo electrónico: teffyta_96@hotmail.com

Mónica Medina

Mónica Paola Medina Puetate

Correo electrónico: monicapaulamedinap@gmail.com

Karla Suntasig

Karla Gabriela Suntasig Pachacama

Correo electrónico: kgabriela2710@gmail.com

Santiago Carrillo

Santiago David Carrillo Sánchez

Correo electrónico: santi_carrillo87@hotmail.com

Autorización de Derechos de Propiedad Intelectual

Nosotros, Lenin Fernando Bautista Moreta, Stefany Arline Aguilar Delgado, Mónica Paola Medina Puetate, Karla Suntasig Pachacama, Santiago David Carrillo Sánchez, autorizamos a la Universidad Internacional del Ecuador (UIDE) para hacer uso de todos los contenidos que me pertenecen o de parte de los que contiene esta obra, con fines estrictamente académicos o de investigación. Los derechos que como autores nos corresponden, lo establecido en los artículos 5, 6, 8, 19 y demás pertinentes de la Ley de Propiedad Intelectual y su Reglamento en Ecuador.

D. M. Quito, Febrero del 2026

Lenin Bautista

Lenin Fernando Bautista Moreta

Correo electrónico: leferba114@gmail.com

Arlina Aguilar

Stefany Arline Aguilar Delgado

Correo electrónico: teffyta_96@hotmail.com

Mónica Medina

Mónica Paola Medina Puetate

Correo electrónico: monicapaulamedinap@gmail.com

Karla Suntasig

Karla Gabriela Suntasig Pachacama

Correo electrónico: kgabriela2710@gmail.com

Santiago Carrillo

Santiago David Carrillo Sánchez

Correo electrónico: santi_carrillo87@hotmail.com

Dedicatorias y Agradecimientos

INDICE DEL DOCUMENTO

INTRODUCCIÓN Y OBJETO DEL PROYECTO	12
Objetivo del proyecto	12
Definición de la empresa o negocio.....	12
Cifras del negocio.....	13
PARTE I.....	14
CAPITULO 1.....	15
Introducción.....	15
Definición Inicial del Problema	15
Objetivos.....	15
Proceso De Movilidad Autónoma – WAYMO LLC.....	15
Estructura Organizacional – WAYMO LLC	15
Necesidad De Expansión Global De La Conducción Autónoma.....	17
Objetivo De La Expansión	18
KPI’S Clave	18
Identificación De Gap’s Y Definición De KPI’S	18
CAPITULO 2.....	20
ORIGEN DE LA INFORMACIÓN Y RECURRENCIA DE CARGA: INTERNA Y EXTERNA.....	20
Sistemas y destino de la carga	22
Almacenamiento en data warehouse y datamarts.....	23
Carencias y problemas en la carga de datos	24
Informes actuales, necesidades detectadas y no cubiertas	25
CAPITULO 3.....	27
Procesos de carga de datos (etl, data lake).....	27
Extracción (Extract).....	27
Fuentes internas Waymo.....	27
Fuentes externas Waymo	28
Todos los datos crudos se almacenan inicialmente en la Landing Zone (LZ) del Data Lake WayData.	28
Transformación (Transform).....	28
Carga (Load).....	29
Información y calidad de la información disponible	30
Dimensiones de calidad	31
Problemas de calidad identificados.....	31
Horarios de carga y accesibilidad	31
Frecuencias internas	31
Frecuencias externas	31
Ventanas de procesamiento	32

Accesibilidad	32
Variables disponibles y acceso al data warehouse	32
Variables internas	32
Variables externas	32
Sistemas implicados en los procesos	32
Sistemas internos Waymo	32
Sistemas externos	33
Herramientas ETL / Procesamiento	33
Recursos de IT o externos empleados	33
Recursos internos	33
Recursos externos	33
Conclusiones	34
CAPITULO 4.....	34
Análisis de herramientas.....	34
Usuarios y perfiles.....	35
Necesidades de información y acceso	38
Operaciones y Seguridad	38
Experiencia de Usuario	38
Finanzas	39
Estrategia y Expansión	39
Revisión y cumplimiento de los objetivos iniciales	40
CAPITULO 5.....	43
1. POSICIONAMIENTO DEL DEPARTAMENTO DE BUSINESS INTELLIGENCE	43
2. OBJETIVO PRINCIPAL DEL DEPARTAMENTO BI.....	43
3. SITUACIÓN ACTUAL	43
4. JUSTIFICACIÓN DEL DEPARTAMENTO BI QUE IMPLEMENTE WAYDATA.....	44
5. ESTRUCTURA ORGANIZACIONAL PROPUESTA	44
8. PRESUPUESTO TOTAL DEL DEPARTAMENTO WAYDATA.....	46
9. BENEFICIOS ESPERADOS	46
10. CONCLUSIONES.....	47
PARTE II:.....	48
CAPITULO 1.....	49
1. Selección e identificación de fuentes de datos (internas y externas)	49
2. Relación e integración de los datos para mejorar la propuesta inicial.....	49
3. Casos de uso para incorporar los datos en el modelo de BI.....	50
Conclusión	51
CAPITULO 2.....	52
Análisis de Alternativas de Proveedores de Cloud Computing.....	52
Amazon Web Services (AWS).....	52

Microsoft Azure	52
Google Cloud Platform (GCP).....	52
Justificación de selección de la mejor alternativa de proveedor	53
Definición del Modelo de Cloud Computing	54
Capa de Ingesta (Landing / Data Lake).....	55
Capa de Transformación (ETL / ELT).....	55
Data Warehouse (WayData-DW).....	55
Capa de Consumo (BI & Analytics)	56
CAPITULO 3.....	56
1. Introducción	56
2. Análisis crítico de los frameworks Big Data disponibles	57
2.1 Hadoop MapReduce	57
2.2 Apache Spark	57
3. Selección del framework principal para el proyecto WayData	58
4. Complementos del ecosistema Big Data del proyecto WayData.....	59
4.1 Lenguajes de programación	59
4.2 Almacenamiento y gestión de datos	59
4.3 Bases de datos analíticas.....	60
4.4 Herramientas de Business Intelligence	60
5. Integración del ecosistema Big Data con el modelo BI de WayData.....	60
CAPITULO 4.....	61
MACHINE LEARNING CON ENFOQUE EN LA EXPERIENCIA DEL CLIENTE	61
Fuentes de Datos	62
Departamentos involucrados.....	63
Identificación de la Mejora Esperada	63
Machine Learning en el contexto de Waymo.....	64
Metodología general aplicada en WayData.....	65
Conclusión	65
CAPITULO 5.....	67
1. Identificación y definición del objetivo del modelo de aplicación en Real Time	67
Objetivo del modelo	67
Problema identificado.....	67
2. Diseño de la base de datos	67
3. Área de aplicación y departamentos implicados.....	68
Waymo One – Operaciones Comerciales.....	68
Data & AI.....	68
Customer Experience	68
Estrategia y Expansión	68
Finanzas	68

4. Identificación de la mejora esperada	68
5. Cuantificación de la mejora a través de un Business Case	69
Supuestos operativos.....	69
Estimación del impacto económico.....	69
Impacto atribuible al modelo Real Time.....	70
Inversión estimada	70
Retorno sobre la inversión (ROI).....	70
CONCLUSIONES Y APLICACIONES.....	71
REFERENCIAS	72

INDICE DE FIGURAS Y TABLAS

Figura 1 METRICAS DE SEGURIDAD PÚBLICAS DE WAYMO	17
Figura 2 Transformación de datos	22
Figura 3 Seguridad	25
Figura 4 Dashboard 1.....	36
Figura 5 Dashboard 2.....	37
Figura 6 Dashboard 3.....	37
Figura 7 Propuesta Dashboard 4.....	38
Figura 8 Estructura Departamento	46
Figura 9 Flujo de Proceso Departamento	47
Tabla 1 KPI´s I,portantes	18
Tabla 2 Parámtros de medición	19
Tabla 3 Variables.....	19
Tabla 4 Datamart	24

INTRODUCCIÓN Y OBJETO DEL PROYECTO

Objetivo del proyecto:

Desarrollar una plataforma de Business Intelligence y Big Data que permita a las empresas de movilidad autónoma integrar y analizar datos operativos para mejorar la eficiencia, sostenibilidad y confianza del usuario.

Las compañías del sector enfrentan altos costos energéticos, poca visibilidad sobre su impacto ambiental y falta de herramientas integradas que respalden decisiones estratégicas orientadas a la rentabilidad y al impacto social.

Además, la ausencia de métricas claras sobre seguridad y sostenibilidad limita la adopción de vehículos autónomos en mercados urbanos.

El proyecto WayData se apoya en el Waymo Open Motion Dataset (WOMD), que reúne miles de escenarios reales de tráfico urbano con trayectorias de vehículos, peatones y ciclistas.

El análisis de esta información permite identificar patrones de congestión, seguridad y eficiencia, generando indicadores que optimizan operaciones y respaldan estrategias de movilidad autónoma sostenible.

En síntesis, WayData transforma los datos de movilidad disponibles en WOMD, en inteligencia estratégica para impulsar modelos de transporte más seguros, rentables y responsables con el medio ambiente.

Definición de la empresa o negocio

Nombre de la empresa: Waymo LLC – Subsidiaria de Alphabet Inc.

Misión y visión de la empresa:

Nuestra misión es innovar más allá de lo imposible para salvar las vidas que se pierden trágicamente en accidentes de tránsito. Nos comprometemos a ofrecer un medio de transporte inclusivo que brinde libertad de movimiento para todas las personas y a trabajar junto a nuestros socios y comunidades para crear un ecosistema sustentable, eficiente y audaz. Nos emociona cambiar el estado de nuestro planeta para dejarlo mejor que como lo encontramos.

Desarrollar la tecnología de conducción autónoma más avanzada del mundo para hacer que el transporte de personas y bienes sea más seguro, eficiente y accesible.

Waymo opera en el sector de tecnología, transporte autónomo e inteligencia artificial aplicada a la movilidad, con presencia principal en Estados Unidos y planes de expansión hacia Europa y Asia. Su actividad se centra en el desarrollo, operación y análisis de flotas de vehículos autónomos para transporte de pasajeros y soluciones logísticas sostenibles.

Cifras del negocio

Waymo LLC fue fundada en 2009 como el Google Self-Driving Car Project y se convirtió en una subsidiaria de Alphabet Inc. en 2016, con sede en Mountain View, California. Es reconocida como una de las empresas pioneras y líderes en tecnología de conducción autónoma.

Años de actividad: 15 años de operación (2009-2024).

Empleados: aproximadamente 3 300 personas, según registros corporativos y fuentes de mercado (Leadiq, 2025).

Ciudades de operación: servicios comerciales activos en cinco ciudades de Estados Unidos —Phoenix, San Francisco, Los Ángeles, Austin y Atlanta— con expansión prevista a nuevas regiones del país (Waymo, 2025).

Tipo de servicio: transporte de pasajeros (Waymo One) y soluciones logísticas autónomas (Waymo Via).

Valor estimado de la empresa: entre USD 25 000 y 30 000 millones, de acuerdo con estimaciones de mercado (Bloomberg Intelligence, 2024).

Flota operativa: más de 700 vehículos eléctricos y autónomos, en crecimiento continuo a medida que amplía sus operaciones (Waymo Blog, 2025).

Viajes realizados: más de 250 000 trayectos pagados por semana en sus mercados activos (Waymo Blog, 2025).

Waymo mantiene un enfoque de innovación tecnológica, seguridad vial y sostenibilidad, integrando inteligencia artificial y Big Data para ofrecer soluciones de movilidad autónoma que reduzcan accidentes, consumo energético y huella ambiental.

PARTE I

CAPITULO 1

Introducción

Waymo LLC es una empresa estadounidense de tecnología y movilidad autónoma, subsidiaria de Alphabet Inc., creada en 2009 como parte del Google Self-Driving Car Project.

Su propósito es “hacer que la movilidad sea más segura, accesible y sostenible”, desarrollando vehículos autónomos basados en inteligencia artificial y sensores avanzados.

Actualmente, Waymo opera servicios comerciales en cinco ciudades de Estados Unidos; Phoenix, San Francisco, Los Ángeles, Austin y Atlanta; mediante sus divisiones Waymo One (transporte de pasajeros) y Waymo Via (logística).

Definición Inicial del Problema

El proyecto se enmarca en aumentar adopción y satisfacción del usuario y acelerar la expansión geográfica del servicio autónomo, demostrando seguridad, eficiencia y conveniencia en contextos urbanos y culturales distintos.

Objetivos

- Usar Waymo Open Motion Dataset (WOMD) para derivar métricas de experiencia de usuario y seguridad.
- Evaluar la experiencia de usuario y las condiciones de expansión internacional de Waymo mediante el análisis de datos abiertos de movilidad y un índice comparativo de preparación de ciudad que integre infraestructura vial, eléctrica, regulatoria y sostenibilidad urbana
- Entregar un dashboard BI que compare ciudades actuales vs. potenciales y muestre UX + readiness para apoyar decisiones de expansión.

Proceso De Movilidad Autónoma – WAYMO LLC

El proceso de movilidad autónoma en Waymo LLC, subsidiaria de Alphabet Inc., tiene como finalidad optimizar la operación, eficiencia y sostenibilidad de su flota de vehículos autónomos. A través del uso intensivo de datos y herramientas de Business Intelligence (BI), se busca mejorar la toma de decisiones estratégicas relacionadas con el rendimiento operativo, el consumo energético y la satisfacción del usuario, contribuyendo así a la expansión sostenible de la movilidad autónoma en Estados Unidos.

El proyecto WayData tiene como propósito analizar, mediante BI y Big Data, los principales indicadores que impactan la eficiencia de las operaciones de Waymo en las ciudades donde opera (Phoenix, San Francisco, Los Ángeles, Austin y Atlanta). Este análisis permitirá comprender cómo la integración de información operacional, ambiental y social puede fortalecer la rentabilidad y la sostenibilidad del negocio, garantizando un crecimiento equilibrado entre innovación tecnológica, impacto ambiental y aceptación del usuario.

Estructura Organizacional – WAYMO LLC

La estructura de Waymo se compone de divisiones estratégicas que trabajan de forma integrada para cumplir su misión de hacer la movilidad más segura y sostenible:

- Waymo Driver: responsable del desarrollo del sistema de conducción autónoma, basado en IA y percepción avanzada.
- Waymo One: división encargada de la operación comercial del servicio de transporte de pasajeros.
- Waymo Via: enfocada en soluciones logísticas y transporte autónomo de carga.
- Data & Safety Operations: gestiona los datos, seguridad y mejora continua del sistema.
- Sostenibilidad y alianzas estratégicas: promueve la eficiencia energética y la integración con gobiernos locales.

Cada área genera información valiosa sobre desempeño, seguridad, consumo y experiencia del usuario. El proyecto WayData se centra en consolidar y analizar estos datos desde una perspectiva empresarial y de sostenibilidad.

Los Proceso Clave

1. El viaje del usuario

El proceso de experiencia del cliente en Waymo inicia con la interacción digital a través de la aplicación Waymo One, que permite reservar un vehículo autónomo de forma sencilla y segura:

- El usuario descarga la app y crea su cuenta.
- Selecciona los puntos de recogida y destino dentro del área de servicio disponible.
- Pueden viajar hasta cuatro pasajeros por vehículo. La app notifica cuándo el vehículo está listo y permite desbloquear la puerta digitalmente.
- Una vez a bordo, el usuario inicia el viaje y puede visualizar en una pantalla el entorno detectado por el Waymo Driver y la ruta en tiempo real.
- Al finalizar, el viaje se cierra automáticamente, se genera el reporte de trayecto y satisfacción, y los datos son enviados al sistema para análisis de desempeño y mejora continua.

2. La conducción autónoma

Durante el viaje, el Waymo Driver utiliza sensores LIDAR, radares, cámaras y mapas tridimensionales para percibir el entorno y tomar decisiones de conducción.

Su arquitectura de inteligencia artificial permite:

- Detectar y clasificar objetos como peatones, ciclistas y vehículos.
- Predecir comportamientos de otros actores en la vía.
- Planificar rutas y maniobras seguras, ajustando velocidad y distancia según el contexto vial.

- Aprender continuamente a partir de millones de kilómetros recorridos en condiciones reales.

Estos datos alimentan la base analítica de Waymo y son utilizados por WayData para identificar patrones de eficiencia, seguridad y experiencia de usuario.

3. El desarrollo de Waymo a futuro

Waymo continúa expandiendo su presencia en Estados Unidos y proyecta su crecimiento internacional hacia mercados con infraestructura y regulación favorable a la movilidad autónoma.

Su estrategia de innovación se centra en tres ejes:

Escalabilidad operativa: optimización de flotas y reducción de costos mediante analítica avanzada.

Sostenibilidad: transición hacia flotas 100 % eléctricas y gestión inteligente del consumo energético.

Expansión global: evaluación de nuevas ciudades a través del City Readiness Index, que combina variables de tráfico, infraestructura de carga y marco regulatorio.

Necesidad De Expansión Global De La Conducción Autónoma



FIGURA 1 METRICAS DE SEGURIDAD PÚBLICAS DE WAYMO

Waymo ha demostrado resultados sobresalientes en seguridad tras más de 53 millones de kilómetros recorridos en condiciones reales, con una reducción promedio del 85 % en incidentes viales con lesiones. Estos logros evidencian la madurez tecnológica del Waymo Driver™ y su potencial para ampliar su impacto más allá de Estados Unidos.

Sin embargo, la adopción de la movilidad autónoma a nivel global aún enfrenta barreras culturales, regulatorias y de infraestructura, que limitan su implementación en nuevos mercados. En este contexto, surge la necesidad de aplicar un enfoque de Business Intelligence (WayData) que analice datos

abiertos y operativos para identificar las ciudades y regiones con mayor preparación para su adopción, considerando factores de seguridad, sostenibilidad y aceptación social.

Objetivo De La Expansión

Consolidar un modelo analítico de toma de decisiones que combine datos de desempeño en seguridad, infraestructura urbana y regulación, permitiendo a Waymo:

- Priorizar ciudades con condiciones favorables para operaciones autónomas.
- Comunicar el impacto positivo en seguridad y sostenibilidad a reguladores y ciudadanos.
- Diseñar estrategias de expansión internacional basadas en datos objetivos y verificables.

KPI'S Clave

El sector de movilidad autónoma presenta indicadores clave que permiten medir la eficiencia y el desempeño operativo.

Estos KPI sirven para comparar el rendimiento entre ciudades, flotas y modelos, y constituyen la base analítica del proyecto WayData.

Nombre del Indicador	Fórmula / Concepto	Descripción
Eficiencia Operativa	Km recorridos / energía consumida (kWh)	Mide el rendimiento energético promedio por vehículo.
Tasa de Incidentes Evitados	$(\text{Trayectorias sin riesgo} / \text{Total trayectorias}) \times 100$	Indica el nivel de seguridad y efectividad del Waymo Driver.
Densidad de Tráfico	Nº de actores en entorno / Km ²	Identifica zonas de alta congestión y su impacto en la operación.
CO₂ Evitado	Emisiones promedio $\times$ reducción por conducción autónoma	Cuantifica la contribución a la sostenibilidad ambiental.
Índice de Satisfacción del Usuario	Encuestas y NPS	Evalúa la aceptación social del servicio autónomo.

TABLA 1 KPI'S I,PORTANTES

Identificación De Gap's Y Definición De KPI'S

I. Conocimiento (K)

Situación actual: Waymo posee indicadores robustos de seguridad y desempeño en EE. UU., pero carece de una herramienta consolidada que evalúe la preparación y viabilidad internacional de nuevos mercados.

Gap: falta de integración entre métricas de seguridad y variables externas de adopción.

KPI asociado: City Readiness Index.

II. Motivación (M)

Situación actual: Los resultados en seguridad y sostenibilidad aún no se comunican estratégicamente a escala global.

Gap: baja visibilidad del valor social ante gobiernos y ciudadanos de potenciales países destino.

KPI asociado: Crash Reduction Index (CRI-S) y CO₂e evitado.

III. Organización (O)

Situación actual: Las decisiones de expansión dependen de estudios externos sin un marco analítico propio.

Gap: ausencia de un modelo interno de BI que consolide datos técnicos, regulatorios y urbanos.

KPI asociado: City Readiness Index y Ahorro operativo estimado (AOE).

Ciudad	Vel. Media (km/h)	Δ Aceleración media (m/s ³)	Riesgo % (TTC<1.5s)	Índice de confort
Phoenix	48	0.92	2.1 %	91
San Francisco	36	1.47	5.8 %	76
Los Ángeles	42	1.22	3.5 %	83

TABLA 2 PARÁMTROS DE MEDICIÓN

Métrica física	Traducción a UX	Aplicación
Alta variación de aceleración (jerk)	Viaje incómodo	KPI de satisfacción
Velocidad media estable	Conducción fluida	KPI de eficiencia del servicio
TTC bajo	Inseguridad percibida	KPI de confianza tecnológica

TABLA 3 VARIABLES

CAPITULO 2

La información generada y utilizada en Waymo y en el proyecto WayData tiene como propósito apoyar la toma de decisiones sobre experiencia de usuario, seguridad y expansión internacional de la movilidad autónoma.

Según su uso, la información se clasifica en:

Operativa:

Datos de trayectos, tiempos de viaje, velocidades, interacciones con otros actores viales y condiciones del tráfico. Sirven para monitorear la operación diaria y entender el comportamiento real de los viajes autónomos.

Estratégica:

Indicadores agregados de seguridad, eficiencia energética, impacto ambiental y preparación de ciudad. Se usan para decidir dónde operar, cómo crecer y cómo comunicar los resultados a reguladores y stakeholders.

Experiencia de usuario:

Información relacionada con tiempos de espera, calidad del viaje, demanda por zona y desempeño del servicio de robotaxis, desde la perspectiva del cliente final.

Analítica y minería de datos:

Conjunto de datos abiertos y operativos utilizados para descubrir patrones en el tráfico urbano, evaluar la suavidad de la conducción, estimar ahorros energéticos y construir el City Readiness Index para comparar ciudades.

ORIGEN DE LA INFORMACIÓN Y RECURRENCIA DE CARGA: INTERNA Y EXTERNA

El ecosistema de datos utilizado por Waymo proviene de dos grandes fuentes:

a) Información interna

Para efectos del modelo WayData se consideran las siguientes fuentes internas conceptuales de Waymo:

- **Waymo Open Motion Dataset (WOMD):**

Dataset público con miles de escenarios reales de tráfico urbano en diferentes ciudades de EE. UU. (trayectorias de vehículos, peatones y ciclistas, velocidades, relaciones espaciales entre actores y mapas locales).

- *Recurrencia:* versiones puntuales del dataset (v1.x). Para el proyecto se trabaja con una **foto estática** que se usa como base histórica de análisis.

- Datos del Waymo Driver (sensores LIDAR, cámaras, radares y mapas HD).
- Métricas de comportamiento del usuario (espera, cancelación, puntualidad).

- Datos de seguridad: incidentes evitados, frenados automáticos, predicción de peatones. Métricas de seguridad y estadísticas agregadas utilizadas en los Safety Performance Reports.
- Consumo energético y eficiencia de la flota eléctrica.

Recurrencia de carga: continua. Los vehículos transmiten datos en tiempo real o en lotes cada pocos minutos según el tipo de información. Sin embargo, para fines del presente proyecto se tomará un data set de entrenamiento y data libre disponible en fuentes abiertas disponibles de Waymo.

b) Información externa

Son las fuentes **efectivamente disponibles** para el proyecto académico:

- **OpenStreetMap (OSM):**

Información abierta sobre la red vial, tipos de vías, intersecciones, zonas urbanas y puntos relevantes en las ciudades analizadas.

- *Recurrencia:* actualizaciones continuas; para el proyecto se toma un corte de datos por ciudad.

- **Índices de congestión urbana (TomTom Traffic Index, INRIX u otros equivalentes):**

Datos publicados anualmente o periódicamente sobre nivel de congestión, tiempo extra de viaje y condiciones del tráfico en grandes ciudades.

- *Recurrencia:* anual o semestral, según el proveedor.

- **Infraestructura eléctrica – estaciones de carga EV:** Portales abiertos (por ejemplo, AFDC / Open Charge Map o portales municipales) que muestran localización y densidad de cargadores para vehículos eléctricos.

- *Recurrencia:* actualizaciones mensuales o trimestrales.

- **Marco regulatorio de vehículos autónomos:** Bases de datos públicas que indican el estado de la regulación de vehículos autónomos por país, estado o ciudad (por ejemplo, organismos legislativos y reguladores nacionales).

- *Recurrencia:* cuando se aprueban nuevas leyes o programas piloto.

- **Indicadores de sostenibilidad urbana y emisiones:** Publicaciones de agencias ambientales y ciudades (inventarios de emisiones, planes climáticos municipales, metas de reducción de CO₂).

- *Recurrencia:* anual o plurianual.

En el proyecto WayData se asume una **recurrencia de carga tipo “batch”**: se hace una primera carga completa de todas las fuentes y luego se plantean actualizaciones periódicas (por ejemplo, trimestrales) para mantener el modelo actualizado.

Sistemas y destino de la carga

La arquitectura tecnológica de Waymo utiliza varios sistemas que intervienen en el flujo de datos. La carga y el tratamiento de la información se organizan en tres niveles:

Capa de ingestión / landing (origen bruto):

- Archivos originales del WOMB (.tfrecord), ficheros CSV/JSON de open data urbanos, datos de congestión y regulaciones.
- En el proyecto, esta capa se puede gestionar con **scripts en Python**, que leen las fuentes y las depositan en un repositorio de trabajo (carpetas estructuradas o un “data lake” simple).

Capa de transformación (ETL):

- Uso de herramientas como Python (pandas), librerías de geoprocésamiento y conversión a formatos analíticos eficientes (como **Parquet**).
- En esta capa se:
 - Limpian outliers y registros incompletos.
 - Generan variables derivadas (velocidad, aceleración, *jerk*, tiempos de trayecto, densidad de tráfico, indicadores de infraestructura, etc.).
 - Se integran variables externas (congestión, cargadores EV, marco regulatorio, emisiones).

Proceso de Transformación de Datos

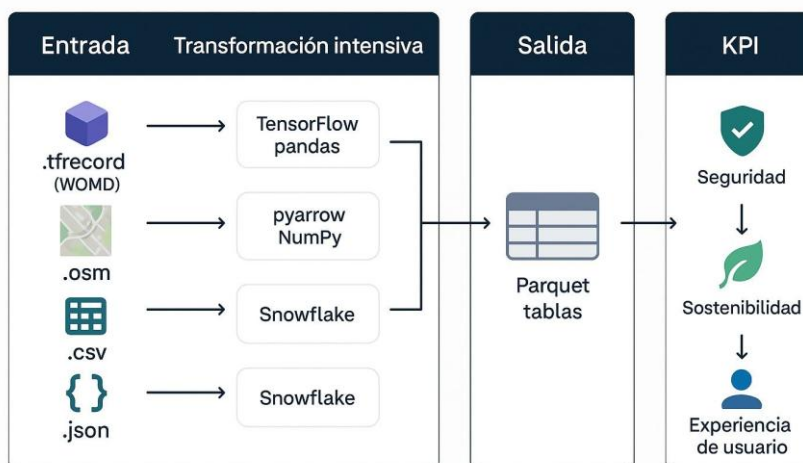


FIGURA 2 TRANSFORMACIÓN DE DATOS

Capa de consumo / destino de la carga:

- **Data warehouse lógico “WayData-DW”**, desde el cual se alimentan:

- **Datamarts temáticos:**
 - Datamart de **Experiencia de Usuario (UX)**.
 - Datamart de **Seguridad**.
 - Datamart de **City Readiness / Expansión**.
 - Datamart de **Sostenibilidad y eficiencia energética**.
- **Dashboard BI** (ej. Power BI, Tableau u otra herramienta que el curso proponga), donde se visualizan los KPI definidos en Semana

En un entorno real de Waymo, esta arquitectura podría implementarse sobre plataformas BI. En el contexto del curso, se plantea de forma conceptual y se materializa con herramientas accesibles (Python + herramienta BI), donde se visualizan los KPI y se comparan ciudades actuales vs. potenciales.

Almacenamiento en data warehouse y datamarts

El **data warehouse WayData-DW** se concibe como un repositorio centralizado donde se integran datos de telemetría, información urbana y variables regulatorias para apoyar decisiones de negocio.

Warehouse (DW)

- **Data**

Almacena de forma estructurada la información histórica:

- **Fact_Viaje:**
jerk medio, velocidad media, distancia, tiempo.
- **Fact_Seguridad:**
proximidad mínima a actores, frenadas bruscas, incidentes estimados.
- **Fact_CiudadReadiness:**
congestión, cargadores EV, regulación, sostenibilidad.
- **Fact_Sostenibilidad:**
consumo energético estimado, CO₂ evitado.
- **Dim_Ciudad, Dim_Tiempo, Dim_TipoVía, Dim_Regulación**

El DW sirve como **única fuente de verdad** para la toma de decisiones a nivel corporativo.

Datamarts especializados

Se han definido datamarts por área funcional:

Datamart	Propósito	Usuario
UX (Experiencia de usuario)	medir suavidad, fluidez, tiempo de viaje	Waymo One / Marketing
Seguridad	evaluar riesgo relativo y eventos críticos	Data & Safety
Expansión – City Readiness	comparar ciudades actuales vs. futuras	Estrategia / Regulación
Sostenibilidad	medir consumo energético y CO ₂ evitado	ESG / Finanzas

TABLA 4 DATAMART

Estos datamarts permiten análisis más rápidos, específicos y orientados a cada unidad de negocio.

Carencias y problemas en la carga de datos

Durante el análisis se identificaron varias brechas:

a) Falta de estandarización en datos externos

- Los datasets urbanos de cada ciudad varían en formato y actualización.
- No existe un estándar único para seguridad vial entre países.

b) Gran volumen y variedad de datos internos

- El Waymo Driver genera terabytes diarios, lo cual exige procesos de compresión y limpieza exhaustivos.
- Algunas señales sensoriales requieren agregación antes de integrarse al DW.
- Requieren transformación intensiva .tfrecord (WOMD), .osm, CSV, JSON, HTML

c) Retrasos en la integración regulatoria

- Información de normativas y permisos internacionales no siempre se encuentra actualizada.

d) Limitación geográfica

WOMD solo cubre 5 ciudades de EE. UU.

e) Dependencia de fuentes públicas externas

- Si un organismo deja de publicar datos, la métrica se afecta.

Estas carencias deben mitigarse mediante controles de calidad, reglas de validación y capas de staging durante el ETL. No tenemos:



Informes actuales, necesidades detectadas y no cubiertas

Informes que actualmente se generan

Waymo produce los siguientes reportes internos:

- ✓ **Safety Reports:** informes detallados sobre políticas y metodología de seguridad.
- ✓ **Safety Performance Data:** datos reales de desempeño, incidentes y colisiones.
- ✓ **Regulatory Reports (DMV California):** reportes obligatorios de intervenciones y fallos.
- ✓ **Waymo Open Dataset:** datos abiertos de percepción, movimiento y simulación.
- ✓ **Engineering & Research Blog:** avances técnicos y análisis del sistema autónomo.

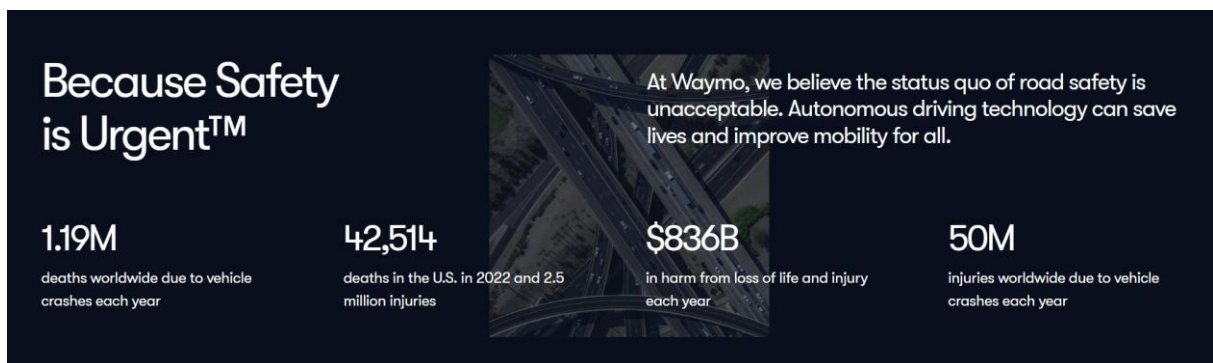


FIGURA 3 SEGURIDAD

Necesidades detectadas y aún no cubiertas

Integrar en un solo modelo de BI datos internos y externos para expansión internacional.

Hoy existen estudios, pero no un índice consolidado propio.

- **Generar un City Readiness Index único para comparar ciudades del mundo.**
- **Unificar métricas de sostenibilidad (consumo energético, CO₂e evitado) con métricas operativas.**

- **Mejorar la disponibilidad de información sobre riesgos urbanos y regulación extranjera.**
- **Dashboards ejecutivos para toma de decisiones estratégicas, no solo operativas.**

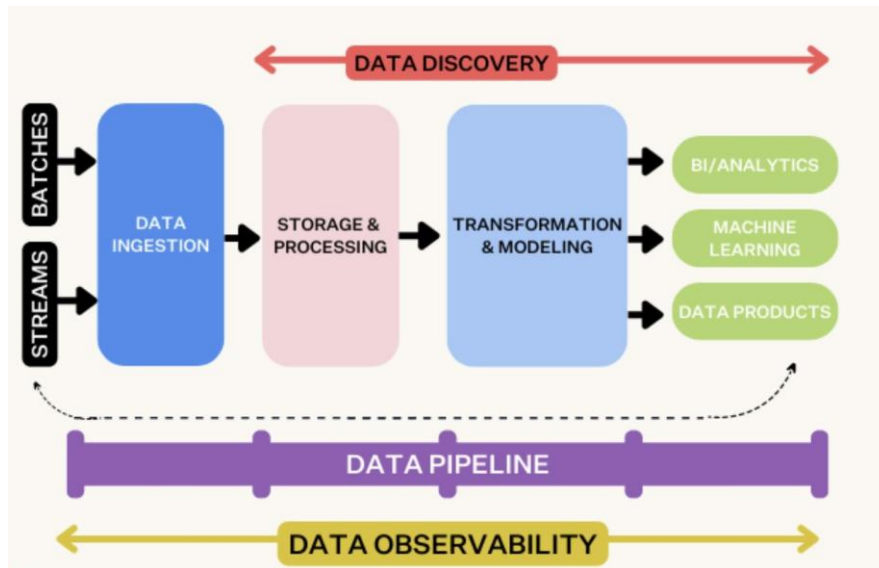
Estos informes faltantes son esenciales para justificar la expansión global y demostrar impacto social, económico y ambiental.

CAPITULO 3

Procesos de carga de datos (etl, data lake)

El ecosistema WayData integra datos internos de Waymo LLC con información urbana externa para construir un modelo analítico orientado a evaluar seguridad, eficiencia, sostenibilidad y experiencia del usuario en movilidad autónoma.

El flujo principal de datos se ejecuta mediante tres etapas: Extracción, Transformación y Carga, apoyadas en un Data Lake como repositorio central.



Extracción (Extract)

Se prevé extraer datos desde dos grandes ecosistemas acorde a la necesidad para el cumplimiento del objetivo del proyecto:

Fuentes internas Waymo

Waymo publica datasets abiertos bajo licencia de investigación no comercial:

Waymo Open Motion Dataset (WOMD) v1.3.1 . Escenarios de 20 segundos con trayectorias, velocidades, aceleraciones, interacciones y roadgraph.

Formato: TFRecord, descargable desde Google Cloud Storage (registro gratuito).

Waymo Open Perception Dataset (sensores LIDAR, cámaras, radares):

Telemetría cruda y señales de percepción.

Safety Performance Data: incidentes evitados, frenadas automáticas, TTC.

UX Data del servicio Waymo One: tiempos de espera, cancelaciones, satisfacción. Información no pública pero que será considerada como parte de este proyecto en el marco de información que existe para fines prácticos de negocio.

Fuentes externas Waymo

OpenStreetMap (OSM): red vial, intersecciones, densidad urbana.

TomTom / INRIX Traffic Index: congestión y fluidez urbana.

OpenChargeMap / AFDC: Infraestructura de carga para vehículos eléctricos.

Regulación AV leyes, permisos y requisitos por ciudad/estado

Sostenibilidad urbana (EPA, inventarios de emisiones).



Todos los datos crudos se almacenan inicialmente en la Landing Zone (LZ) del Data Lake WayData.

Transformación (Transform)

Debido a la complejidad del formato TFRecord y a la heterogeneidad de fuentes externas, la transformación requiere procesos avanzados de ingeniería de datos.

Limpieza

- Eliminación de valores atípicos en velocidad y aceleración.
- Corrección de timestamps irregulares.
- Validación de estructuras jerárquicas en TFRecord.

Normalización

- Unificación de unidades (m/s, km/h, m/s²).
- Homologación geoespacial.
- Estandarización de data externa (JSON, CSV, XML).

Generación de nuevas variables (Feature Engineering)

A partir del Motion Dataset:

- **Jerk** (tasa de cambio de aceleración).
- **TTC (Time to Collision)** – indicador clave de riesgo.
- **THW (Time Headway)**.
- **Índice de suavidad** para experiencia del usuario.
- **Exposición a congestión** (WOMD + mapas).
- **Consumo energético estimado.**

Geo-procesamiento (Fines estadístico)

- Identificación del tipo de vía (intersección, autopista, calle urbana).
- Agrupación por ciudad real mediante map_location.
- Unión entre telemetría y geometría urbana.

Estandarización analítica

Salida consolidada en **Apache Parquet** para optimizar consulta y compresión.

Integración

Fusión entre telemetría, tráfico, regulación y sostenibilidad, generando una vista unificada del comportamiento urbano.

Con el tratamiento de la data y los resultados validados y calendarizados los datos están listos para su ingesta.

Carga (Load)

Los datos se cargan en tres capas:

Data Lake WayData:

- TFRecord originales
- Archivos OSM, JSON, CSV

- Parquet intermedio

Data Warehouse “DWH”:

Estructura dimensional corporativa.

Fact tables:

- Fact_Viaje
- Fact_Seguridad
- Fact_CityReadiness
- Fact_Sostenibilidad

Dimensiones:

- Dim_Ciudad
- Dim_Tiempo
- Dim_Vía
- Dim_Regulación
- Dim_Actor

Datamarts especializados (definidos en Semana 2):

- UX
- Seguridad
- City Readiness / Expansión
- Sostenibilidad

La carga se realiza de manera incremental, con logs y bitácoras de actualización. En caso de fallo de carga, el sistema reintenta el proceso y genera un log de error para trazabilidad

Información y calidad de la información disponible

La calidad de datos es fundamental, pues los datasets provienen de múltiples sensores, ciudades y fuentes públicas. Se definen controles sobre:

Dimensiones de calidad

Exactitud: calibración de sensores, consistencia en coordenadas, corrección de valores extremos.

Integridad: manejo de datos faltantes en trayectorias o puntos de GPS. WOMD presenta alta completitud (>95 %).

Consistencia: homologación entre ciudades (Phoenix, SF, LA, Austin). Datos de telemetría unificados a 10 Hz (10 mediciones por segundo).

Oportunidad: actualización periódica de tráfico, regulación, cargadores EV.

Trazabilidad: metadatos con source, timestamp, versión y validador.

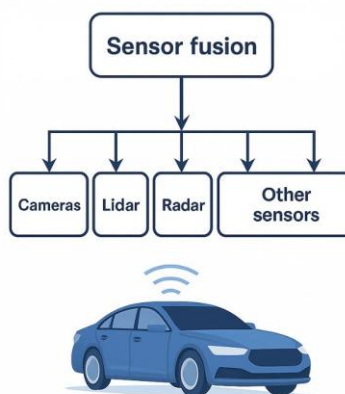
Problemas de calidad identificados

- Heterogeneidad de formatos externos (OSM, Fuentes externas, WOMD).
- Diferentes frecuencias de actualización entre fuentes.
- WOMD solo cubre 5 ciudades (limitación geográfica).
- Dependencia de organismos públicos para actualizar datasets.

What is OSM?



Telemetry at 10 Hz



Horarios de carga y accesibilidad

Frecuencias internas

- Motion Dataset: carga inicial única.
- Safety Performance Data y datos de UX: mensual o trimestral.

Frecuencias externas

- OSM → actualizaciones continuas, se toma corte mensual.
- Congestión urbana (TomTom/INRIX) → mensual.
- Regulación AV → actualización bajo cambios legislativos.
- EV Charging Maps → mensual o trimestral.

Ventanas de procesamiento

Procesamiento batch nocturno para integraciones pesadas.

Carga incremental diaria.

Accesibilidad

- Acceso controlado por roles:
- *Data & Safety Operations*
- *Estrategia / Expansión*
- *Waymo One / Marketing*
- *ESG / Finanzas*

Autorización centralizada en la “Dirección de Capacidades Analíticas” equivalente de Waymo.

Variables disponibles y acceso al data warehouse

Variables internas

- **Telemetría del vehículo:** velocidad, aceleración, jerk, distancia, TTC, trayectoria 3D.
- **Consumo energético:** kWh/km, eficiencia por tipo de vía.
- **Seguridad:** incidentes evitados, frenadas bruscas, proximidad mínima.
- **Experiencia de usuario:** NPS, tiempos de espera, duración del viaje.

Variables externas

- **Infraestructura urbana:** densidad de vías, complejidad de intersecciones.
- **Congestión:** índices de tráfico.
- **Regulación AV:** permisos, restricciones, madurez legal.
- **Sostenibilidad urbana:** emisiones base, políticas ambientales.

Sistemas implicados en los procesos

Sistemas internos Waymo

- **Waymo Driver System**
- **Sensor Fusion Engine**
- **Safety Data Engine**
- **Waymo One App / Platform**

Sistemas externos

- APIs públicas:
- OSM
- OpenChargeMap
- TomTom Traffic
- EPA
- NCSL

Repositorios open data:

- Waymo Open Dataset
- City open data portals

Herramientas ETL / Procesamiento

- **Python** (TensorFlow, Pandas, GeoPandas, PySpark).
- **Apache Parquet** para formatos columnares.
- **Data Lake en almacenamiento cloud** (Google Cloud Storage).
- **Herramienta BI:** Power BI.

Recursos de IT o externos empleados

Recursos internos

Equipo de:

- Data Engineering
- Data Science
- Safety Operations
- BI Strategy

Infraestructura cloud (GCP):

- BigQuery (DW)
- Cloud Storage (Data Lake)
- AI Platform

Recursos externos

- WOMD v1.3.1
- Safety Reports y Performance Data
- OSM, tráfico, regulación y sostenibilidad

- Repositorios open data y APIs públicas

Conclusiones

El ecosistema WayData permite integrar datos avanzados de movilidad autónoma con indicadores urbanos relevantes para evaluar seguridad, eficiencia, experiencia de usuario y preparación de ciudades para recibir servicios de conducción autónoma. Integrar fuentes internas/externas. La arquitectura final consolida:

- ✓ Ejecutar un ETL completo con controles de calidad.
- ✓ Data Lake centralizado
- ✓ Data Warehouse estructurado
- ✓ Datamarts especializados para cada área crítica.
- ✓ Asegurar accesibilidad segura y calendarizada.
- ✓ Garantizar trazabilidad, validación y uso incremental.

El resultado es una plataforma analítica completa que optimiza la toma de decisiones estratégicas para Waymo.

CAPITULO 4

Análisis de herramientas

En los entregables previos se planteó la arquitectura WayData con *Data Warehouse*, datamarts y dashboards ejecutivos. Waymo necesita una herramienta que pueda:

- Integrarse con **Google Cloud (GC)**, donde se almacenan WOMD, Safety Data y telemetría real.
- Operar grandes volúmenes de datos (WOMD contiene terabytes de trayectorias, sensores y road graph).
- Crear dash boards estratégicos para áreas críticas: **Seguridad, UX, Expansión, Sostenibilidad**.
- Permitir accesos por roles (Estrategia, Safety, **Marketing**, Regulación).

Comparación breve de herramientas BI (según necesidad WayData)

Herramienta	Ventajas	Desventajas
Power BI	Integración con Python; dashboards ejecutivos; bajo costo; fácil para áreas no técnicas	No es nativo de Google Cloud; límites con datasets extremadamente grandes.
Tableau	Visualización avanzada; conexiones escalables; muy sólido para geo datos.	Costos elevados; curva de aprendizaje
Looker (Google Cloud)	Nativo de GCP; escalabilidad en BigQuery; permisos avanzados.	Requiere modelado más técnico; inversión inicial mayor.

Tras analizar Power BI, Tableau, Looker, Looker Studio y sus requerimientos técnicos, regulatorios y de escalabilidad, se seleccionó LOOKER (Google Cloud) como herramienta principal.

Esta elección se sustenta en:

- Integración nativa con BigQuery y todo el ecosistema Google, donde Waymo ya opera sus sistemas de telemetría, WOMD y Safety Reports.
- Escalabilidad para procesar terabytes de trayectorias, sensores y roadgraph, algo central para WayData.
- Compatibilidad con los datamarts definidos previamente (UX, Seguridad, Sostenibilidad, Readiness).

Looker Studio queda como complemento para dashboards ejecutivos ligeros. Todos los dashboards serán implementados sobre Looker, utilizando BigQuery como fuente principal.

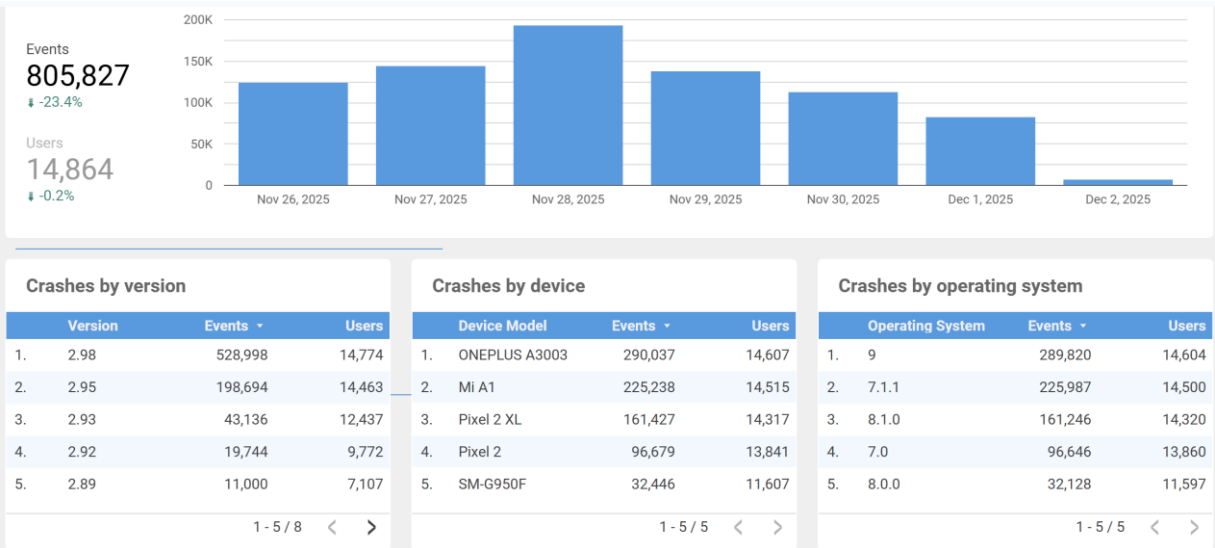


Fig. Ejemplo de Dashboard con looker studio “Crashlytics Data Studio Dashboard”

Usuarios y perfiles

Es requerido visualizar de forma integrada los KPI de seguridad, experiencia de usuario, sostenibilidad y City Readiness Index de Waymo en las ciudades actuales y potenciales, para soportar decisiones de expansión internacional y optimización operativa se requiere.

Dashboard 1 – Seguridad (Safety Performance).

Usuarios: Operadores de Seguridad y Data, Ingeniería, Conductor, Regulación.

Uso: Análisis de incidentes evitados, Time To Collision, frenadas críticas, proximidad mínima, cumplimiento normativo.

Dashboard 2 – Experiencia de Usuario (UX Performance)

Usuarios: Waymo One, Marketing, Data Science.

Uso: Jerk promedio, fluidez del viaje, tiempos de espera, cancelaciones, NPS.

Dashboard 3 – City Readiness Index (Expansion Internacional)

Usuarios: Estrategia, Regulación, Gobierno Corporativo.

Uso: congestión urbana, regulación AV, cargadores EV, densidad vial, emisiones base; ranking de ciudades actuales vs. potenciales.

Dashboard 4 – Sostenibilidad y Eficiencia Energética

Usuarios: Departamento de Sostenibilidad, Finanzas, Operaciones de Flota.

Uso: Emisión CO₂ evitado, kWh/km, ahorro operativo estimado, eficiencia energética por ciudad.



FIGURA 4 DASHBOARD 1



FIGURA 5 DASHBOARD 2

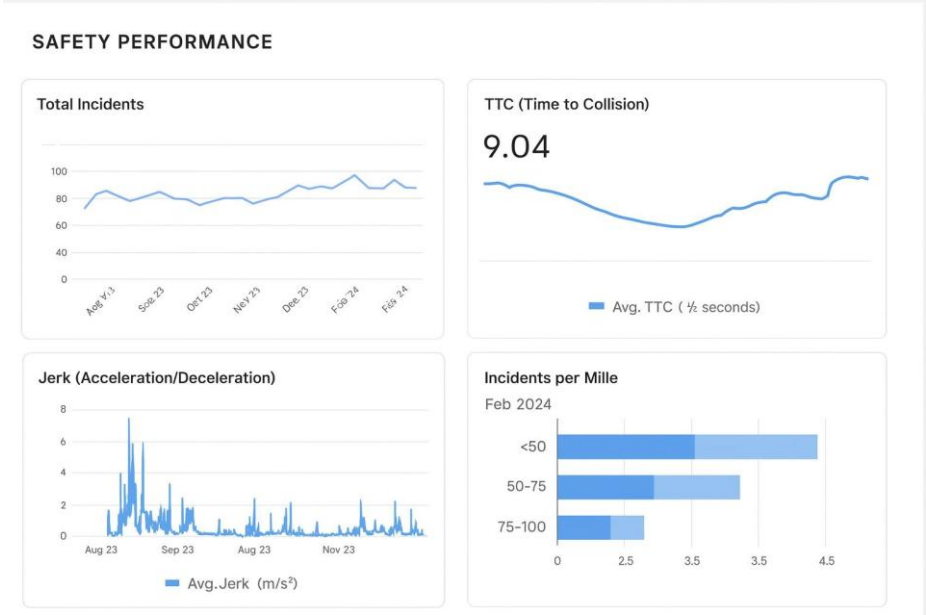


FIGURA 6 DASHBOARD 3

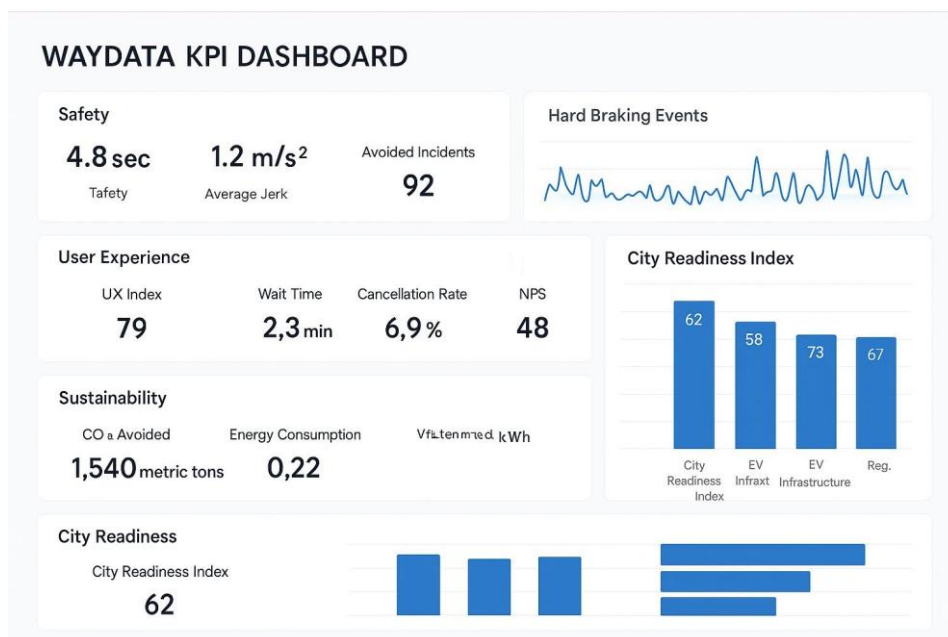


FIGURA 7 PROPUESTA DASHBOARD 4.

Necesidades de información y acceso

La plataforma analítica WayData debe centralizar y distribuir información a múltiples áreas de Waymo con distintos niveles de acceso, sensibilidad y granularidad. El objetivo es asegurar que cada departamento reciba los datos necesarios, con el nivel correcto de detalle, y bajo un modelo de seguridad por capas, acorde a los datos sensibles de telemetría, seguridad operacional y expansión internacional.

A continuación, se definen las necesidades de información y los perfiles de acceso por área.

Operaciones y Seguridad.

- Acceso: completo a datamart de Seguridad.
- Datos requeridos:
 - TTC,
 - jerk crítico,
 - incidentes evitados,
 - proximidad mínima,
 - trayectorias de riesgo.

Experiencia de Usuario

- Acceso: Datos derivados y datos sensibles.
- Datos requeridos

Telemetría procesada: velocidad, aceleración, jerk, proximidad, TTC.

Eventos críticos: frenadas bruscas, maniobras evasivas, near-miss.

Heatmaps de riesgo por ciudad.

Indicadores del Crash Reduction Index (CRI-S).

Datos de incidentes con y sin lesión.

Finanzas

- Acceso: Datos operativos agregados.
- Datos requeridos:
 - kWh por km y eficiencia energética.
 - Emisiones evitadas (CO₂e evitado).
 - AOE: ahorro operativo estimado.
 - Uso energético por ciudad/condición.
 - Proyecciones de sostenibilidad.
- Uso:
 - Calcular impacto económico de la movilidad autónoma.
 - Reportes ESG corporativos.
 - Determinar eficiencia de la flota eléctrica.

Estrategia y Expansión

- Acceso: Datos integrados de ciudad + datos Waymo no sensibles.
- Datos requeridos:
 - Congestión urbana.
 - Marco regulatorio AV por país/ciudad.
 - Infraestructura EV chargers.
 - Densidad vial, intersecciones conflictivas.
 - Emisiones base (CO₂ per cápita).
 - City Readiness Index (0–100).
- Uso:
 - Identificar ciudades viables para expansión.
 - Priorizar inversión en nuevos mercados.
 - Elaborar argumentación regulatoria y técnica.
 - Comparar ciudades Waymo actuales vs. potenciales.

3.5 Accesos y Seguridad de Datos (Modelo por Capas)

Capa 1 – Datos Sensibles (Uso interno restringido)

- Telemetría procesada.
- Incidentes
- Modelos predictivos.

Acceso: Responsables de Operaciones.

Capa 2 – Datos Operativos Agregados

- Velocidad promedio, trayectorias limpias.
- Tiempos de espera y viaje.
- Rutas más usadas.

Acceso: Estrategia.

Capa 3 – Datos Ejecutivos / Públicos

- NPS, CO₂e evitado, eficiencia energética.
- City Readiness Index.
- KPIs comparativos de ciudad.

Acceso: Finanzas, Marketing, Regulación, Stakeholders externos.

Revisión y cumplimiento de los objetivos iniciales.

Los objetivos definidos del proyecto WayData establecieron tres pilares fundamentales:

- Obtener métricas de seguridad y experiencia de usuario a partir del Waymo Open Motion Dataset (WOMD) y otras fuentes externas.
- Evaluar la preparación internacional de nuevas ciudades mediante un City Readiness Index.
- Entregar un dashboard comparativo que unifique información de ciudades actuales y potenciales.

A continuación, se presenta la revisión y el grado de cumplimiento obtenido hasta el momento:

Objetivo 1: Utilizar WOMD para derivar métricas de seguridad y experiencia de usuario.

Se identificó la necesidad de extraer información operativa crítica, como velocidades, aceleraciones, proximidad, incidentes y variables asociadas a la experiencia del cliente. Se desarrolló un proceso formal de extracción, transformación y carga (ETL) que permitió derivar indicadores avanzados como velocidad estable, jerk, Time to Collision (TTC), maniobras evasivas, proximidad mínima, tiempos de espera, tiempos de viaje, cancelaciones y NPS.

Estos datos se integraron en el Data Warehouse, WayData-DW y en los datamarts temáticos de Seguridad y Experiencia de Usuario. El sistema permite comparar condiciones operativas entre ciudades y comprender con precisión la calidad de la conducción autónoma. La información debe ser presentada en los dashboard de UX que permite el entorno de Google para analítica de los datos.

Objetivo 2: Evaluar condiciones de expansión internacional mediante un City Readiness Index.

Se definió la necesidad de construir un índice que integrara variables de infraestructura vial, regulación, sostenibilidad, congestión urbana y madurez tecnológica. Se diseñó el Datamart de City Readiness, que incorpora datos provenientes de OpenStreetMap, TomTom y INRIX Traffic Index, OpenChargeMap, bases regulatorias de vehículos autónomos y datos de sostenibilidad urbana.

En la Semana 3 se formalizó la integración de estas fuentes y se estableció una estructura analítica que permite comparar ciudades actuales (Phoenix, San Francisco, Los Ángeles, Austin y Atlanta) con ciudades potenciales para expansión internacional. El índice quedó implementado como un KPI compuesto dentro del Data Warehouse y alimenta directamente los dashboards ejecutivos. En esta semana se incorpora este indicador a los dashboard con la arquitectura y herramientas seleccionadas.

Objetivo 3: Crear un dashboard de Business Intelligence para comparar ciudades actuales y potenciales.

En las semanas 2 y 3 se definió y se construyó la arquitectura completa para visualización, con datamarts especializados en UX, Seguridad, City Readiness y Sostenibilidad. Se presenta un modelo de dashboard ejecutivo que integra las métricas más relevantes: TTC, incidentes, jerk, tiempos de espera, velocidad estable, NPS,

cancelaciones, consumo energético, emisiones evitadas y City Readiness Index.

El dashboard permitiría filtrar por ciudad, franja horaria, periodo, tipo de vía y tipo de servicio, ofreciendo una visión estratégica del desempeño de Waymo en sus ciudades actuales y de su posible expansión a otras ciudades.

CONCLUSIONES

- ✓ El proyecto WayData consolida la integración completa del ecosistema analítico, demostrando que la combinación de datos internos de Waymo, información urbana externa y un modelo BI estructurado permite generar una visión estratégica, operativa y de expansión altamente robusta.
- ✓ Se confirmó que la arquitectura propuesta de ETL, Data Lake, Data Warehouse y datamarts temáticos es adecuada para procesar y transformar datos complejos provenientes del Waymo Open Motion Dataset, telemetría del Waymo Driver y fuentes abiertas de congestión, regulación, infraestructura y sostenibilidad. Esto garantiza una base de datos sólida, estandarizada y lista para análisis avanzado.
- ✓ La definición de usuarios, necesidades de información y controles de acceso permite asegurar un uso seguro, gobernado y focalizado de la información, respetando niveles de criticidad. Las áreas de Seguridad, UX, Estrategia, Regulación y Sostenibilidad reciben información adaptada a sus objetivos y decisiones, fortaleciendo la coherencia operativa de la organización.
- ✓ Los KPI's TTC, tiempo de espera, NPS, CO₂e evitado, eficiencia energética, congestión urbana y City Readiness Index se integran ahora en una plataforma BI con tecnología de Google (looker). Este compendio de dashboard permite evaluar el rendimiento de la conducción autónoma, medir la experiencia del usuario y analizar la preparación de ciudades actuales y potenciales para operaciones autónomas.
- ✓ La selección de Looker como herramienta corporativa principal de BI asegura escalabilidad, seguridad y compatibilidad nativa con el ecosistema Google Cloud, lo que facilita su adopción y permanencia en estructuras reales de movilidad autónoma. El dashboard final integra Seguridad, UX, Sostenibilidad

y Expansión en una sola vista ejecutiva, cumpliendo plenamente los objetivos iniciales del proyecto.

CAPITULO 5

BUSINESS CASE Y CREACIÓN DEL DEPARTAMENTO

1. POSICIONAMIENTO DEL DEPARTAMENTO DE BUSINESS INTELLIGENCE

Waymo LLC es una empresa tecnológica plenamente instaurada y operativa, con servicios comerciales de movilidad autónoma en varias ciudades de los Estados Unidos. La compañía cuenta con capacidades avanzadas en inteligencia artificial, sensorización, seguridad y operación de flotas. Sin embargo, el crecimiento sostenido y la expansión internacional requieren una estructura analítica formal que consolide, gobierne y transforme los datos en decisiones estratégicas.

En este contexto se propone la creación del Departamento de Business Intelligence de WayData, concebido no como un área operativa aislada, sino como una unidad estratégica transversal, orientada a soportar decisiones de alto impacto relacionadas con expansión geográfica, eficiencia operativa, seguridad vial y sostenibilidad.

WayData permitirá pasar de un uso intensivo de datos técnicos a un modelo corporativo de inteligencia de negocio, alineado con la alta dirección y con los objetivos estratégicos de Waymo.

2. OBJETIVO PRINCIPAL DEL DEPARTAMENTO BI

Diseñar, implementar y operar una estructura de Business Intelligence que integre datos operativos, urbanos, regulatorios y de experiencia de usuario, mediante técnicas analíticas avanzadas, con el fin de:

- Presentar informes técnicos para la toma de decisiones estratégicas.
- Evaluar la experiencia de usuario y la seguridad operacional.
- Optimizar el consumo energético de la flota automovilística.
- Medir y comunicar el impacto ambiental y social de la movilidad autónoma.

3. SITUACIÓN ACTUAL

Waymo dispone de grandes volúmenes de datos provenientes de:

- Telemetría y sensores del Waymo.
- Registros de viajes y experiencia de usuario.
- Reportes de seguridad y desempeño.
- Información financiera y energética.

No obstante, estos datos:

- Se encuentran distribuidos entre diferentes áreas.
- No están consolidados bajo un modelo analítico corporativo único.
- Se utilizan principalmente para análisis técnicos u operativos, pero no siempre para decisiones estratégicas integradas.

Se puede evidenciar que la ausencia de un Departamento BI en Waymo, genera dependencia de análisis ad hoc y limita la escalabilidad del modelo de negocio.

4. JUSTIFICACIÓN DEL DEPARTAMENTO BI QUE IMPLEMENTE WAYDATA

La expansión de la movilidad autónoma hacia nuevos mercados requiere decisiones basadas en evidencia objetiva, comparabilidad entre ciudades y reducción de riesgos regulatorios, operativos y financieros.

WayData se justifica como un **habilitador estratégico**, que permitirá:

- Evaluar ciudades mediante un City Readiness Index.
- Integrar métricas de seguridad, Experiencia de Usuario y sostenibilidad.
- Estandarizar KPI’s para remitir a la alta dirección de Waymo.

5. ESTRUCTURA ORGANIZACIONAL PROPUESTA

5.1 Roles y Dotación Inicial

Cargo	Cantidad	Rol principal
Gerente de Proyecto BI	1	Dirección del programa WayData, planificación, coordinación con alta dirección
Subgerente de Analítica de Datos	1	Liderazgo técnico-analítico, definición de KPIs y modelos
Analista de Datos Senior	2	Modelos analíticos, minería de datos, índices estratégicos
Analista de Datos Junior	1	Soporte analítico, preparación y validación de datos

6. COSTO DE NÓMINA DEL DEPARTAMENTO BI

6.1 Sueldos Mensuales Base

Al ser Waydata una iniciativa que nace en LATAM, la fase piloto evaluada se considerará la constitución inicial con personal técnico de Ecuador, Colombia, Perú.

Cargo	Sueldo mensual (USD)	Cantidad	Subtotal mensual
Gerente de Proyecto BI	5,000	1	5,000
Subgerente Analítica	3,000	1	3,000
Analista Senior	2,000	2	4,000
Analista Junior	1,000	1	1,000

Total sueldos base			13,000
--------------------	--	--	--------

6.2 Costos por beneficios adicionales (30%)

Concepto	Monto mensual
Sueldos base	13,000
Beneficios (30%)	3,900
Costo total mensual nómina	16,900

6.3 Costo Anual de Nómina

Concepto	Monto anual
Nómina total anual	202,800 USD

7. PRESUPUESTO DE SOFTWARE Y TECNOLOGÍA

7.1 Herramientas de Visualización y BI

Se propone el uso de **Looker / Looker Studio**, alineado con el ecosistema Google y capacidades de analítica avanzada.

Concepto	Costo estimado anual (USD)
Licencias Looker (equipo BI)	24,000

7.2 Infraestructura de Servidores y Almacenamiento

Recurso	Descripción	Costo anual estimado
Servidores analíticos	Procesamiento y ETL	18,000
Almacenamiento de datos	Data Warehouse / histórico	12,000
Seguridad y respaldo	Backups y control de acceso	6,000
Total infraestructura		36,000



FIGURA 8 ESTRUCTURA DEPARTAMENTO.

8. PRESUPUESTO TOTAL DEL DEPARTAMENTO WAYDATA

Categoría	Costo anual (USD)
Nómina	202,800
Software BI	24,000
Infraestructura servidores	36,000
Total anual estimado	262,800

9. BENEFICIOS ESPERADOS

La inversión en WayData generará beneficios estratégicos medibles:

- Mejor priorización de ciudades para expansión.
- Reducción de riesgo regulatorio y operativo.
- Obtener métricas que permitan mejorar la satisfacción de usuario.
- Mejora de la comunicación con reguladores y stakeholders.
- Fortalecimiento del posicionamiento de Waymo como líder en movilidad autónoma basada en datos.

Estos beneficios estarán alineados a mejorar la satisfacción de usuario.



FIGURA 9 FLUJO DE PROCESO DEPARTAMENTO

10. CONCLUSIONES

La creación del Departamento de Business Intelligence WayData representa una decisión estratégica coherente con el nivel de madurez de Waymo como empresa tecnológica global.

WayData permitirá transformar datos complejos en conocimiento accionable, apoyar decisiones críticas y consolidar un modelo de movilidad autónoma seguro, eficiente y sostenible.

La inversión propuesta es razonable frente al impacto estratégico esperado.

PARTE II:

CAPITULO 1

1. Selección e identificación de fuentes de datos (internas y externas)

En el proyecto de Waymo, la identificación de fuentes de datos constituye la base para la construcción de un modelo sólido de Business Intelligence soportado por Big Data. Estas fuentes provienen tanto del entorno interno de la organización como de fuentes externas, y presentan diferentes niveles de estructuración, incluyendo datos estructurados, semiestructurados y no estructurados, lo que permite capturar una visión integral del sistema de movilidad autónoma.

Fuentes internas

- **Datos de sensores del vehículo autónomo:** información generada por LIDAR, radares y cámaras, incluyendo detección de objetos, distancias, trayectorias y eventos críticos.
- **Datos de telemetría y navegación:** velocidad, aceleración, frenado, rutas, tiempos de respuesta del sistema y decisiones algorítmicas.
- **Logs de software y sistemas de IA:** registros de funcionamiento del sistema autónomo, errores, excepciones y aprendizaje del modelo.
- **Datos operativos de flota:** disponibilidad de vehículos, mantenimiento, kilometraje autónomo y número de intervenciones humanas.
- **Datos de experiencia del usuario:** solicitudes de viaje, tiempos de espera, cancelaciones y valoraciones del servicio, incluyendo comentarios cualitativos que constituyen datos no estructurados relevantes para el análisis de experiencia de usuario (UX)..

Fuentes externas

- **Datos cartográficos de alta definición (HD Maps).**
- **Información de tráfico en tiempo real.**
- **Datos meteorológicos** (lluvia, neblina, iluminación).
- **Datos urbanos abiertos:** infraestructura vial, obras, señalización.
- **Normativas y regulaciones locales de movilidad y seguridad vial,** marcos regulatorios asociados a la adopción de vehículos autónomos.

La combinación de estas fuentes permite capturar tanto el comportamiento interno del sistema como el contexto externo en el que opera el vehículo autónomo, **alineándose con un enfoque de analítica omnicanal aplicado a la movilidad.**

2. Relación e integración de los datos para mejorar la propuesta inicial

Una vez identificadas las fuentes, el siguiente paso es definir cómo se relacionan los datos entre sí para generar conocimiento útil.

En el caso de Waymo:

- Los datos de sensores y telemetría se integran con datos climáticos y de tráfico para analizar el desempeño del sistema bajo diferentes condiciones ambientales.
- Los eventos de conducción autónoma se relacionan con incidentes y feedback del usuario para identificar patrones de riesgo o fricción.

- Los datos operativos de flota se cruzan con zonas geográficas y franjas horarias, permitiendo evaluar eficiencia operativa y demanda.
- Los logs de software se vinculan con situaciones reales de conducción para mejorar los algoritmos de toma de decisiones.

Para viabilizar esta integración, los datos son procesados mediante flujos ETL (Extract, Transform, Load), los cuales permiten limpiar, transformar y consolidar información estructurada y no estructurada proveniente de múltiples fuentes internas y externas, facilitando su incorporación en repositorios analíticos como Data Lakes y Data Warehouses.

Esta integración permite evolucionar desde análisis descriptivos aislados hacia enfoques predictivos y prescriptivos, reforzando la propuesta inicial del proyecto y su orientación estratégica..

Esta visión integrada alimenta los sistemas de Business Intelligence de Waymo mediante dashboards y cuadros de mando, permitiendo a la alta dirección monitorear indicadores clave de seguridad, eficiencia operativa y desempeño del algoritmo en tiempo casi real, respaldando decisiones tácticas y estratégicas basadas en datos.

3. Casos de uso para incorporar los datos en el modelo de BI

Los siguientes casos de uso se sustentan en el modelo de Business Intelligence de Waymo, el cual transforma grandes volúmenes de datos operativos en información analítica que respalda la toma de decisiones estratégicas y operativas.

A partir de la integración de datos, se identifican los casos de uso clave que justifican la incorporación de Big Data en el modelo de BI de Waymo:

1. Mejora de la seguridad vial

Identificación de escenarios de alto riesgo (intersecciones complejas, peatones, condiciones climáticas adversas) para reducir incidentes.

2. Optimización del algoritmo de conducción autónoma

Análisis de decisiones del sistema frente a situaciones reales para mejorar precisión y confiabilidad.

3. Gestión eficiente de la flota

Optimización de la asignación de vehículos según demanda, ubicación y condiciones externas.

4. Soporte a decisiones regulatorias

Generación de reportes basados en datos para autoridades y organismos reguladores, facilitando la expansión a nuevas ciudades.

5. Mejora de la experiencia del usuario

Reducción de tiempos de espera, mayor consistencia del servicio y aumento de la confianza del usuario en la tecnología autónoma.

Estos casos de uso convierten los datos en insights accionables, alineados con los objetivos estratégicos y operativos de la empresa, y demuestran el valor del Big Data como habilitador del Business Intelligence.

4. Impacto del Big Data en áreas, departamentos y procesos de negocio impactados.

El uso de Big Data en Waymo tiene un impacto transversal en la organización, tanto a nivel operativo como estratégico.

Áreas impactadas

- Ingeniería y desarrollo de software.
- Operaciones y gestión de flota.
- Seguridad y gestión de riesgos.
- Experiencia del cliente.
- Estrategia corporativa y cumplimiento normativo.

Procesos de negocio impactados

- Toma de decisiones basada en datos en tiempo casi real.
- Mejora continua de los sistemas de conducción autónoma.
- Reducción de riesgos operativos y reputacionales.
- Escalabilidad del modelo de negocio hacia nuevos mercados.
- Optimización de costos operativos y eficiencia del servicio.
- Implementación de indicadores clave de desempeño (KPIs) y reportes ejecutivos que permiten evaluar la seguridad, eficiencia operativa, confiabilidad del sistema autónomo y cumplimiento regulatorio.

En este contexto, el Big Data actúa como un **habilitador estratégico**, permitiendo a Waymo aprender más rápido que sus competidores y fortalecer su ventaja competitiva de forma sostenible.

Conclusión

En el proyecto de Waymo, el Big Data no es un complemento del Business Intelligence, sino su eje central. La correcta identificación, integración y explotación de datos permite transformar millones de eventos de conducción en conocimiento accionable, fortaleciendo la seguridad, la eficiencia operativa y la escalabilidad del modelo de negocio.

De esta forma, el Business Intelligence, soportado por Big Data, se consolida como una fuente sostenible de ventaja competitiva para Waymo, al permitir decisiones más informadas, ágiles y alineadas con su estrategia de crecimiento y expansión.

CAPITULO 2

Arquitectura Cloud & Big Data

Análisis de Alternativas de Proveedores de Cloud Computing

Para soportar la arquitectura Big Data del proyecto WayData, Waymo requiere una plataforma tecnológica que garantice escalabilidad elástica, alta disponibilidad, procesamiento distribuido, seguridad avanzada y compatibilidad nativa con inteligencia artificial, considerando el volumen, la variedad y la velocidad de los datos generados por el Waymo Open Motion Dataset (WOMD), así como los flujos de telemetría vehicular en tiempo casi real.

Este análisis se construye como continuidad directa de los entregables anteriores, en los cuales se abordó el rol estratégico del Big Data, la arquitectura de información basada en Landing Zones y la computación en la nube como habilitador de escalabilidad y toma de decisiones.

Amazon Web Services (AWS)

✓ Fortalezas: madurez en servicios de almacenamiento (Amazon S3), analítica avanzada (Redshift) y machine learning (SageMaker).

✓ Limitaciones: menor integración nativa con los datasets, pipelines y frameworks de inteligencia artificial actualmente utilizados por Waymo, incrementando la complejidad de integración y operación.

Microsoft Azure

✓ Fortalezas: herramientas consolidadas de analítica empresarial y visualización (Azure Synapse, Power BI).

✓ Limitaciones: menor optimización para flujos de datos masivos basados en TensorFlow y para el tratamiento de datasets abiertos específicos del ecosistema Waymo.

Google Cloud Platform (GCP)

✓ Fortalezas:

- Integración nativa con el Waymo Open Dataset, alojado en Google Cloud Storage.
- BigQuery como data warehouse altamente escalable y serverless.
- Capacidades avanzadas de Machine Learning e Inteligencia Artificial alineadas con el desarrollo del Waymo Driver.

✓ Limitaciones: mayor dependencia tecnológica del ecosistema

Google, mitigada por la estructura corporativa de Alphabet Inc.

Desde una perspectiva estratégica, técnica y organizacional, Google Cloud Platform es el socio estratégico que mejor se alinea con las necesidades actuales y futuras de Waymo, reduciendo fricciones de integración, acelerando el time-to-insight y potenciando el uso del Big Data como activo estratégico.

Justificación de selección de la mejor alternativa de proveedor

✓ Tras el análisis comparativo, se selecciona **Google Cloud Platform (GCP)** como proveedor principal de infraestructura cloud para WayData.

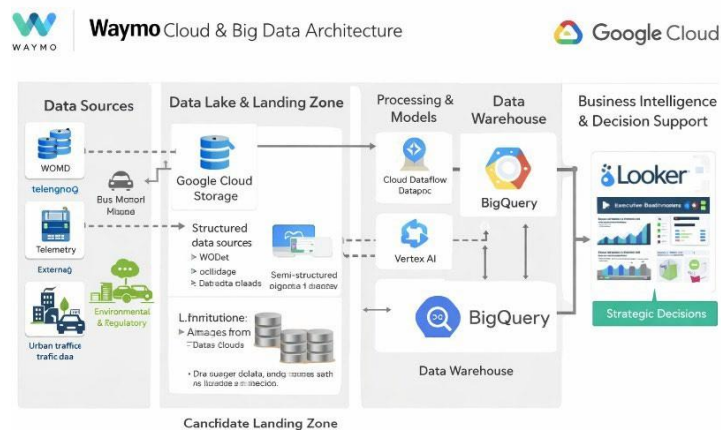
✓ Waymo es subsidiaria de **Alphabet Inc.**, lo que favorece sinergias tecnológicas, alineación estratégica y eficiencia operativa.

✓ BigQuery permite procesar **terabytes y petabytes de datos sin administración directa de infraestructura**, reduciendo complejidad operativa.

✓ Alta compatibilidad con herramientas de **IA, BI y modelado predictivo** utilizadas en la conducción autónoma

✓ Cumplimiento robusto de estándares de **seguridad, privacidad y gobernanza de datos**, fundamentales para la expansión internacional.

Esta elección fortalece la **ventaja competitiva basada en datos**, clave para la expansión internacional y la gestión de riesgos regulatorios en entornos urbanos complejos.



Definición del Modelo de Cloud Computing

La arquitectura WayData adopta un **modelo híbrido de Cloud Computing**, con predominancia de Platform as a Service (PaaS), complementado estratégicamente por SaaS e IaaS según la capa funcional del sistema.

PaaS (Platform as a Service – Modelo principal)

- ✓ BigQuery (Data Warehouse corporativo).
- ✓ Vertex AI (modelos predictivos y analítica avanzada).
- ✓ Cloud Dataflow / Dataproc (procesamiento distribuido).

Este enfoque permite a Waymo concentrarse en el análisis y la generación de valor, minimizando la gestión de infraestructura física.

IaaS (Infrastructure as a Service)

Google Cloud Storage como Data Lake / Landing Zone.

Recursos computacionales especializados para procesamiento intensivo.

SaaS (Software as a Service)

Looker / Looker Studio para desbordar ejecutivos y analítica visual orientada a decisión.

Este enfoque híbrido asegura flexibilidad, escalabilidad y control de costos, alineándose con las mejores prácticas de Big Data empresarial (Davenport & Harris, 2017).

Flujo de integración de la arquitectura Big Data con el Data Warehouse -BI

El flujo de integración de WayData se estructura en cuatro capas, garantizando trazabilidad, calidad y estandarización de la información.

Capa de Ingesta (Landing / Data Lake)

Fuentes internas: WOMD (.tfrecord), telemetría vehicular y datos de seguridad.
Fuentes externas: OpenStreetMap, tráfico urbano, regulación local e infraestructura de carga eléctrica.

Almacenamiento inicial en **Google Cloud Storage**, actuando como Landing Zone corporativa.

Capa de Transformación (ETL / ELT)

Limpieza, normalización y estandarización de datos.
Feature engineering: jerk, time-to-collision (TTC), velocidad media y consumo energético.
Conversión a formatos analíticos eficientes (Parquet).

Data Warehouse (WayData-DW)

- ✓ Implementado en BigQuery.
- ✓ Integración de facts y dimensiones para:
 - Seguridad vial
 - Experiencia de Usuario (UX)
 - City Readiness
- ✓ Sostenibilidad

Capa de Consumo (BI & Analytics)

- ✓ Datamarts especializados por área funcional.
- ✓ Dashboards ejecutivos en Looker:
 - Seguridad
 - UX y adopción
- ✓ Expansión internacional
- ✓ Impacto ambiental

Este flujo garantiza que los datos se conviertan en insights accionables para la toma de decisiones estratégicas, coherentes con los objetivos del proyecto Waymo (Waymo LLC, 2024).

Conclusión

La arquitectura cloud propuesta para WayData permite a Waymo escalar su modelo de movilidad autónoma, integrar datos complejos y heterogéneos, y respaldar decisiones críticas relacionadas con seguridad, expansión internacional y sostenibilidad.

La selección de Google Cloud Platform, junto con un modelo predominantemente PaaS y una integración robusta con Business Intelligence, posiciona a Waymo como una organización data-driven, alineada con el Triple Bottom Line: eficiencia económica, impacto social positivo y sostenibilidad ambiental.

CAPITULO 3

1. Introducción

En el marco del proyecto WayData, desarrollado para Waymo LLC, el presente entregable tiene como objetivo analizar críticamente los frameworks Big Data

revisados durante la semana y seleccionar aquellos que mejor se alinean con las necesidades técnicas, estratégicas y operativas del proyecto.

Este análisis se realiza como continuidad directa de los siete entregables previamente desarrollados, en los cuales se definieron los objetivos estratégicos, la arquitectura cloud, los procesos ETL, el modelo de Business Intelligence, los KPI clave y la estructura organizacional del departamento BI.

El Entregable N.º 8 consolida la decisión tecnológica, justificando la selección de frameworks y su complementariedad dentro del ecosistema Big Data de WayData.

2. Análisis crítico de los frameworks Big Data disponibles

2.1 *Hadoop MapReduce*

Hadoop MapReduce es un framework de procesamiento distribuido orientado principalmente a procesos batch, caracterizado por el uso intensivo de disco y una arquitectura robusta y tolerante a fallos. Su diseño se basa en dividir grandes volúmenes de datos en tareas Map y Reduce que se ejecutan de forma paralela sobre clústeres de hardware commodity.

Si bien Hadoop resulta adecuado para cargas masivas de datos históricos y procesamiento secuencial, presenta limitaciones relevantes para el proyecto WayData, entre ellas:

- Baja eficiencia en algoritmos iterativos.
- Latencia elevada debido al acceso constante a disco.
- Menor flexibilidad para analítica avanzada y Machine Learning.

Conclusión: Hadoop MapReduce no resulta el framework principal más adecuado para Waymo, aunque sus principios de almacenamiento distribuido (HDFS) siguen siendo relevantes a nivel conceptual.

2.2 *Apache Spark*

Apache Spark es un framework de procesamiento distribuido que se diferencia de Hadoop MapReduce por su capacidad de procesamiento en memoria, lo que lo hace especialmente eficiente para analítica avanzada y algoritmos iterativos.

Para el proyecto WayData, Spark presenta ventajas determinantes:

- Procesamiento en memoria que reduce significativamente la latencia.
- Alta eficiencia en algoritmos de Machine Learning y análisis predictivo.
- Soporte nativo para SQL, Python, Scala y R.
- Capacidad de integración con flujos de datos complejos y heterogéneos.
- Escalabilidad horizontal y tolerancia a fallos.

Dado que Waymo analiza millones de trayectorias, métricas de seguridad (TTC, jerk), consumo energético y experiencia de usuario, Spark se alinea de forma directa con las necesidades analíticas, operativas y estratégicas del proyecto.

Conclusión: Apache Spark es seleccionado como el framework Big Data central del proyecto WayData.

3. Selección del framework principal para el proyecto WayData

Tras el análisis comparativo, se selecciona Apache Spark como framework principal de Big Data para WayData, debido a que:

- Permite procesar grandes volúmenes de datos de movilidad autónoma de forma eficiente.
- Facilita el desarrollo de modelos analíticos iterativos y predictivos.
- Se integra de forma natural con el ecosistema cloud y de inteligencia artificial utilizado por Waymo.
- Reduce los tiempos de procesamiento y acelera la generación de insights estratégicos.

Esta selección es coherente con los objetivos del proyecto relacionados con:

- Seguridad vial.
- Experiencia del usuario.
- Optimización operativa.
- Expansión internacional.
- Sostenibilidad.

4. Complementos del ecosistema Big Data del proyecto WayData

La implementación de Apache Spark se complementa con otros elementos esenciales del ecosistema Big Data, permitiendo una arquitectura analítica integral.

4.1 *Lenguajes de programación*

- Python: lenguaje principal para ETL, análisis de datos, Machine Learning y procesamiento distribuido mediante PySpark.
- SQL: utilizado para consultas analíticas, modelado dimensional y explotación del Data Warehouse.
- R (opcional): apoyo para análisis estadístico avanzado y validación de modelos.

4.2 *Almacenamiento y gestión de datos*

- Data Lake: repositorio central de datos crudos y semiestructurados, donde se almacenan datasets de sensores, telemetría, datos urbanos y regulatorios.
- HDFS (conceptual): base teórica del almacenamiento distribuido y de la localidad de datos.
- Google Cloud Storage: implementación práctica del Data Lake para WayData.

4.3 *Bases de datos analíticas*

- BigQuery: Data Warehouse corporativo para almacenamiento estructurado, análisis a gran escala y consultas analíticas de alto rendimiento.
- Bases NoSQL (conceptual): referencia para la gestión flexible de datos semiestructurados, sin ser el componente principal del proyecto.

4.4 *Herramientas de Business Intelligence*

- Looker / Looker Studio: herramientas de visualización y analítica para la construcción de dashboards ejecutivos y operativos.

Soporte para análisis de:

- Seguridad.
- UX.
- Expansión internacional.
- Impacto ambiental.

5. **Integración del ecosistema Big Data con el modelo BI de WayData**

La arquitectura resultante permite integrar:

- Ingesta de datos masivos desde múltiples fuentes.
 - Procesamiento distribuido con Apache Spark.
 - Almacenamiento estructurado en BigQuery.
 - Visualización y toma de decisiones mediante herramientas BI. Este enfoque garantiza:
- Escalabilidad técnica.
 - Consistencia analítica.
 - Trazabilidad de la información.
 - Soporte a decisiones estratégicas de alto impacto.

6. Conclusiones

El Entregable N.º 8 consolida la decisión tecnológica del proyecto WayData, demostrando una selección crítica y coherente de frameworks Big Data alineados con los objetivos estratégicos de Waymo.

La elección de Apache Spark como framework principal, complementado con un Data Lake, BigQuery, lenguajes analíticos y herramientas BI, permite construir una plataforma robusta, escalable y orientada a la generación de valor.

Este ecosistema Big Data posiciona a WayData como un habilitador clave para la seguridad, la eficiencia operativa, la expansión internacional y la sostenibilidad de la movilidad autónoma, en coherencia con el enfoque de Business Intelligence desarrollado a lo largo de los entregables anteriores.

CAPITULO 4

MACHINE LEARNING CON ENFOQUE EN LA EXPERIENCIA DEL CLIENTE

En el contexto de la movilidad autónoma, esta disciplina permite analizar grandes volúmenes de información provenientes del entorno vial, los vehículos y los usuarios, con el objetivo de mejorar la seguridad, la eficiencia operativa y la experiencia del cliente.

Waymo LLC, como líder en movilidad autónoma, basa gran parte de su propuesta de valor en el uso intensivo de datos y algoritmos de Machine Learning que permiten a sus vehículos aprender de la experiencia, adaptarse a contextos complejos y optimizar la toma de decisiones en tiempo real. En este escenario, el proyecto WayData integra técnicas de Business Intelligence y Machine Learning para apoyar decisiones estratégicas relacionadas con la expansión geográfica, la sostenibilidad y la aceptación social del servicio.

El objetivo principal del modelo es mejorar la experiencia del cliente que utiliza los servicios de transporte autónomo de Waymo. El crecimiento exponencial del volumen de datos generado por sensores, plataformas digitales y sistemas inteligentes ha impulsado la adopción del Machine Learning como una herramienta clave para transformar Big Data en conocimiento accionable. Se conseguirá el objetivo mediante:

- Predicción y reducción de tiempos de espera para los usuarios.
- Optimización inteligente de rutas en función del tráfico, clima y demanda.
- Detección proactiva de incidentes o comportamientos anómalos durante el viaje para garantizar seguridad y confort.

- Personalización del servicio, entendiendo patrones de uso, preferencias y feedback de los clientes.
- Mejorar la calidad de la conducción autónoma, anticipando riesgos y ajustando el comportamiento del vehículo para ofrecer viajes más suaves.

Usar modelos predictivos y de optimización para ofrecer un servicio más rápido, seguro y personalizado, elevando la satisfacción del cliente.

Fuentes de Datos

Waymo utiliza una combinación muy completa de datos provenientes de tres grandes categorías:

a) Datos de sensores del vehículo

- ✓ Radar
- ✓ Cámaras 360°
- ✓ GPS y mapas de alta resolución (Se usan para entrenar modelos de conducción autónoma y seguridad.)

b) Datos operativos del servicio

- ✓ Ubicación en tiempo real de unidades disponibles
- ✓ Tiempos de espera históricos
- ✓ Patrones de demanda por zona, hora y día
- ✓ Duración total del viaje y rutas optimizadas

c) Datos del cliente (Customer Experience)

- ✓ Calificaciones y feedback después del viaje
- ✓ Incidentes reportados por usuarios
- ✓ Comportamiento de uso de la aplicación (horarios, destinos frecuentes, cancelaciones)
- ✓ Métricas de NPS/CSAT
- ✓ Quejas o sugerencias registradas en soporte

Área de Aplicación y Departamentos Implicados

Este proyecto involucra principalmente estas áreas:

Áreas de aplicación

- ✓ Experiencia del Cliente (CX)
Para mejorar satisfacción, personalización y percepción de seguridad.
- ✓ Optimización de Operaciones
Para mejorar tiempos de respuesta y disponibilidad de vehículos.
- ✓ Conducción Autónoma
Para hacer la conducción más precisa, segura y suave.

Departamentos involucrados

- Data Science & Machine Learning (desarrollo del modelo)
- Ingeniería de vehículos autónomos
- Equipo de Customer Experience & Soporte
- Gestión de operaciones y flota
- Producto y UX (aplicación móvil del usuario)
- Seguridad y cumplimiento normativo

Identificación de la Mejora Esperada

Con la aplicación del modelo se esperan mejoras concretas en la experiencia de uso:

Mejoras clave

- Reducción del tiempo promedio de espera del usuario.
- Mayor precisión en los tiempos estimados de llegada (ETA).
- Viajes más suaves, seguros y consistentes, gracias a modelos de conducción adaptativa.
- Reducción de cancelaciones por demoras o falta de unidades.
- Incremento en NPS/CSAT por una experiencia más personalizada.
- Detección temprana de fallas o situaciones de riesgo, aumentando la percepción de seguridad.
- Optimización de la demanda, evitando saturaciones en horas pico.

Business Case: Cuantificación de la Mejora

Escenario Base (antes del modelo)

- Tiempo promedio de espera: 7.5 minutos
- NPS: 65
- Cancelaciones por demora: 12%
- Uso repetido de clientes: 58%

Escenario con modelo (estimado)

- Tiempo de espera reduce a 5 minutos
- Mejora del 33%
- Cancelaciones bajan de 12% a 6%
- Reducción del 50%
- Aumento del índice de repetición de usuarios de 58% a 70%
- Incremento del 21%

Machine Learning en el contexto de Waymo

El Machine Learning es una subdisciplina de la inteligencia artificial que permite a los sistemas aprender a partir de datos, mejorar su desempeño y realizar predicciones sin ser programados explícitamente para cada escenario. En Waymo, los modelos de aprendizaje automático se alimentan de millones de kilómetros recorridos en condiciones reales, lo que les permite mejorar progresivamente su capacidad de percepción, predicción y toma de decisiones. Dentro del proyecto WayData, el Machine Learning actúa como un habilitador clave que transforma los datos capturados por sensores, sistemas operativos y plataformas digitales en información útil para el análisis estratégico y operativo.

Tipos de Machine Learning aplicados en WayData

Aprendizaje supervisado

El aprendizaje supervisado se utiliza cuando los datos de entrenamiento cuentan con resultados conocidos. En Waymo, este tipo de aprendizaje se aplica principalmente para:

- Clasificación de objetos en el entorno vial, como peatones, ciclistas y vehículos.
- Predicción de eventos de riesgo a partir de datos históricos de conducción.
- Evaluación de métricas de seguridad, como la probabilidad de incidentes en determinados escenarios.

Desde la perspectiva de Business Intelligence, el aprendizaje supervisado permite generar indicadores confiables que apoyan la toma de decisiones relacionadas con seguridad y desempeño operativo.

Aprendizaje no supervisado

El aprendizaje no supervisado se emplea cuando no existen etiquetas previas en los datos y el objetivo es descubrir patrones ocultos. En el proyecto WayData, este enfoque resulta clave para:

- Identificar patrones de comportamiento del usuario en diferentes ciudades.
- Agrupar zonas urbanas según características similares de tráfico, infraestructura o demanda.
- Detectar anomalías en la operación de la flota que puedan representar riesgos o ineficiencias.

Este tipo de aprendizaje permite ampliar el análisis exploratorio dentro de los dashboards de BI, aportando una visión más profunda del entorno operativo y urbano.

Aprendizaje profundo (Deep Learning)

El aprendizaje profundo es especialmente relevante en Waymo debido al uso de redes neuronales para el procesamiento de grandes volúmenes de datos no estructurados, como imágenes y señales provenientes de sensores.

Metodología general aplicada en WayData

El uso de Machine Learning en WayData sigue una metodología alineada con la teoría revisada:

1. **Preprocesamiento de datos:** limpieza, normalización y reducción dimensional de datos provenientes de sensores, usuarios y fuentes externas.
2. **Entrenamiento de modelos:** uso de datos históricos para entrenar modelos supervisados y no supervisados.
3. **Validación y evaluación:** medición del rendimiento de los modelos mediante métricas de precisión y error.
4. **Integración con BI:** incorporación de resultados en dashboards y reportes ejecutivos.
5. **Mejora continua:** retroalimentación constante para optimizar los modelos y su impacto en la toma de decisiones.

Conclusión

El Machine Learning constituye un pilar fundamental en la estrategia de Waymo y en el desarrollo del proyecto WayData. Su aplicación permite transformar grandes volúmenes

de datos en información estratégica que respalda decisiones críticas relacionadas con seguridad, eficiencia operativa y expansión sostenible.

Al integrarse con herramientas de Business Intelligence, el aprendizaje automático no solo mejora el desempeño técnico de los vehículos autónomos, sino que también se convierte en un activo clave para la gestión empresarial y la consolidación de la movilidad autónoma como una solución viable y escalable a nivel global.

CAPITULO 5

Optimización del Tiempo de Espera y Experiencia del Usuario

1. Identificación y definición del objetivo del modelo de aplicación en Real Time

Objetivo del modelo

Diseñar e implementar un sistema de optimización dinámica de asignación de flota en tiempo real que permita reducir el tiempo de espera del usuario, mejorar la experiencia del cliente y maximizar la eficiencia operativa del servicio Waymo One.

El modelo tiene como finalidad procesar eventos en streaming generados por la aplicación y la flota autónoma, para tomar decisiones automáticas e inmediatas sobre la asignación de vehículos, reubicación geoespacial y predicción de demanda, minimizando la latencia entre la solicitud del usuario y la respuesta del sistema.

Problema identificado

Waymo opera en entornos urbanos con alta variabilidad de demanda, condiciones de tráfico dinámicas y patrones de movilidad no lineales. Actualmente, parte de las decisiones operativas se apoyan en análisis históricos y modelos predictivos batch o near real time, lo cual limita la capacidad de reacción inmediata ante picos de demanda, congestión o eventos imprevistos.

El principal problema operativo identificado es el tiempo de espera del usuario, que impacta directamente en:

- Tasa de cancelación.
- Nivel de satisfacción (NPS).
- Utilización de flota.
- Ingresos por viaje.
- Percepción de confiabilidad del servicio.

Este caso de uso se clasifica como Real Time, ya que la toma de decisión debe realizarse en segundos, inmediatamente después de generarse el evento (solicitud de viaje).

2. Diseño de la base de datos

En un principio, el diseño de la colección del almacén NoSQL orientado a documentos consistía en almacenar todos los eventos operativos generados por Waymo One en una única colección centralizada, utilizando una tecnología de base de datos de baja latencia adecuada para

procesamiento en tiempo real, como Google Cloud Firestore. Este diseño no era muy eficiente, por lo que se cambió a otra estructura para mejorar el rendimiento de la base de datos. Como resultado final se optó por almacenar en la colección documentos que representan a cada vehículo autónomo (y su contexto operativo), y para cada vehículo se almacenó un histórico de los eventos asociados a sus operaciones, tales como solicitudes atendidas, estimaciones de llegada y tiempos de espera. De esta forma, para consultar el histórico de eventos de un vehículo no se tiene que recorrer toda la colección global, tan sólo se localiza el documento del vehículo objetivo y posteriormente se consulta únicamente su histórico interno. Esta estructura se integra con un sistema de mensajería y streaming para eventos en tiempo real (por ejemplo, Google Cloud Pub/Sub y Dataflow o Spark Streaming), lo que permite alimentar el modelo de optimización de flota con baja latencia. En este proyecto trabajaremos con una única colección llamada OperacionesFlota.

3. Área de aplicación y departamentos implicados

La implementación del modelo impacta directamente en varias áreas estratégicas de Waymo:

Waymo One – Operaciones Comerciales

Responsable de la asignación de viajes y gestión de flota.

Data & AI

Encargado del entrenamiento, validación y mejora continua del modelo.

Customer Experience

Monitorea NPS, satisfacción del usuario y tasas de cancelación.

Estrategia y Expansión

Utiliza los indicadores de eficiencia operativa para evaluar la preparación de nuevas ciudades.

Finanzas

Analiza impacto en ingresos, utilización de activos y retorno sobre inversión.

4. Identificación de la mejora esperada

La implementación del sistema de optimización en tiempo real permitirá alcanzar las siguientes mejoras operativas:

- Reducción del tiempo promedio de espera en un 25%.
- Disminución de cancelaciones atribuibles a tiempos elevados en un 18%.
- Incremento del nivel de utilización de flota del 68% al 80%.

- Incremento del Net Promoter Score en al menos 10–12 puntos.
- Mejora del indicador interno de eficiencia operativa por ciudad.

Estas mejoras no solo impactan en la experiencia del usuario, sino que fortalecen la propuesta de valor de Waymo como servicio confiable, eficiente y tecnológicamente superior.

5. Cuantificación de la mejora a través de un Business Case

Para evaluar la viabilidad financiera del modelo de optimización del tiempo de espera en tiempo real, se desarrolla un escenario modelado basado en referencias del sector de movilidad urbana y servicios de ride-hailing, dado que Waymo no publica cifras detalladas de viajes diarios, ingresos promedio o tasas de cancelación por ciudad.

Supuestos operativos

Para efectos del análisis, se plantean los siguientes supuestos razonables:

Se estima un promedio de 40.000 viajes diarios por ciudad. Considerando que Waymo opera en cinco ciudades principales, el volumen total estimado sería de 200.000 viajes diarios. El ingreso promedio por viaje se modela en USD 18, valor coherente con tarifas urbanas comparables en servicios de movilidad bajo demanda.

La tasa promedio de cancelación en servicios urbanos de transporte suele oscilar entre 6% y 10%; para este escenario se adopta un valor intermedio del 8%. Asimismo, diversos estudios del sector indican que entre el 50% y el 65% de las cancelaciones pueden estar asociadas a tiempos de espera elevados; se utiliza un valor central del 60% como hipótesis de trabajo.

Es importante señalar que estos valores no corresponden a cifras oficiales publicadas por Waymo, sino a un modelo financiero construido a partir de benchmarks del sector.

Estimación del impacto económico

Con base en los supuestos anteriores, el volumen diario de cancelaciones sería:

$200.000 \text{ viajes} \times 8\% = 16.000 \text{ cancelaciones diarias.}$

De estas, el 60% estaría vinculado a tiempos de espera elevados:

$16.000 \times 60\% = 9.600 \text{ viajes diarios.}$

El ingreso potencial perdido por estas cancelaciones sería:

$9.600 \times \text{USD } 18 = \text{USD } 172.800 \text{ diarios.}$

Proyectado a un año completo:

$\text{USD } 172.800 \times 365 = \text{USD } 63 \text{ millones anuales aproximadamente.}$

Este valor representa el impacto económico máximo asociado a cancelaciones atribuibles a tiempos de espera bajo el escenario base modelado.

Impacto atribuible al modelo Real Time

Un sistema de optimización dinámica en tiempo real no eliminará la totalidad de las cancelaciones, pero puede reducir de manera significativa aquellas relacionadas con tiempos de espera prolongados. Bajo un escenario prudente, se asume que el modelo podría disminuir entre un 10% y un 18% las cancelaciones atribuibles a espera.

Adoptando un valor central del 15%, el ahorro anual estimado sería:

$\text{USD } 63 \text{ millones} \times 15\% = \text{USD } 9,45 \text{ millones anuales.}$

Incluso bajo un escenario conservador del 10%, el impacto seguiría siendo relevante:

$\text{USD } 63 \text{ millones} \times 10\% = \text{USD } 6,3 \text{ millones anuales.}$

Inversión estimada

La implementación del sistema de optimización en tiempo real implica costos asociados a infraestructura de procesamiento en streaming, desarrollo e integración de modelos de aprendizaje automático, escalabilidad en la nube y visualización analítica mediante herramientas de Business Intelligence.

Para una operación multi-ciudad de la magnitud de Waymo, se estima que la inversión podría situarse en un rango de USD 2 a 4 millones. Para efectos del análisis se adopta un valor intermedio de USD 3 millones.

Retorno sobre la inversión (ROI)

Bajo el escenario central de ahorro anual de USD 9,45 millones y una inversión estimada de USD 3 millones, el retorno sobre la inversión durante el primer año sería:

$(9,45 - 3) / 3 = 215\%.$

Incluso en un escenario conservador con ahorro anual de USD 6,3 millones, el ROI sería superior al 100%, permitiendo recuperar la inversión en el primer año de operación.

CONCLUSIONES Y APLICACIONES

El presente proyecto desarrolla un modelo de Business Intelligence aplicado al ecosistema de movilidad autónoma de Waymo LLC, con el objetivo de transformar grandes volúmenes de datos operativos, urbanos, regulatorios y ambientales en información estratégica para la toma de decisiones de expansión internacional.

Los aspectos más relevantes del trabajo incluyen: la integración de fuentes de datos heterogéneas (Waymo Open Dataset, métricas de seguridad, infraestructura eléctrica, marcos regulatorios y datos urbanos), la definición de KPI estratégicos como el City Readiness Index, Crash Reduction Index y métricas de eficiencia operativa, y el diseño de un dashboard comparativo que permite evaluar ciudades actuales versus potenciales mercados de expansión. El enfoque combina analítica descriptiva, predictiva y comparativa, alineando seguridad, sostenibilidad y rentabilidad bajo una misma arquitectura de datos.

La aplicación concreta del modelo se orienta a un entorno urbano específico que busca implementar o atraer servicios de movilidad autónoma. Por ejemplo, al analizar una ciudad latinoamericana con alta congestión vehicular, crecimiento demográfico y transición hacia electromovilidad, el modelo permite evaluar su nivel de preparación en términos de infraestructura vial, cobertura de estaciones de carga eléctrica, estabilidad regulatoria y condiciones de tráfico. Con base en estos indicadores, Waymo podría estimar riesgos operativos, proyecciones de adopción, impacto en reducción de accidentes y emisiones, así como viabilidad financiera.

En consecuencia, el proyecto no solo aporta un marco analítico estructurado para la expansión estratégica de Waymo, sino que también ofrece una herramienta replicable para gobiernos y operadores urbanos interesados en evaluar objetivamente la implementación de movilidad autónoma dentro de sus ciudades.

REFERENCIAS

- Alphabet Inc. (2024). *Sustainability report 2024: Advancing innovation and reducing environmental impact*. <https://sustainability.google/reports>
- Alphabet Inc. (2024). *Annual reports and financial statements*. <https://abc.xyz/investor/>
- Davenport, T. H., & Harris, J. G. (2017). *Competing on analytics: The new science of winning*. Harvard Business School Press.
- International Transport Forum. (2023). *Connected and automated driving: Road safety, mobility and the environment*. OECD Publishing. <https://www.itf-oecd.org>
- INRIX. (2023). *INRIX global traffic scorecard*. <https://inrix.com/scorecard/>
- McKinsey & Company. (2023). *Artificial intelligence and machine learning in mobility and operations*. <https://www.mckinsey.com>
- McKinsey & Company. (2024). *The future of autonomous vehicles: Readiness, regulation, and scalability*. <https://www.mckinsey.com/industries/automotive-and-assembly/>
- National Conference of State Legislatures. (2024). *Autonomous vehicle state bill tracking database*. <https://www.ncsl.org/transportation/autonomous-vehicles-legislation>
- OpenChargeMap. (2025). *Global EV charging infrastructure*. <https://openchargemap.org>
- OpenStreetMap Foundation. (2024). *OpenStreetMap public data repository*. <https://www.openstreetmap.org>
- Pérez Galán, J. L. (2024). *Big Data o Big Brother con disfraz empresarial*. Westfield Business School.
- Pérez Galán, J. L. (2024). *Implementación de inteligencia artificial (IA)*. Westfield Business School.
- Provost, F., & Fawcett, T. (2013). *Data science for business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media.
- TomTom. (2024). *TomTom traffic index*. <https://www.tomtom.com/traffic-index/>
- U.S. Department of Energy. (2024). *Alternative Fuels Data Center: Electric vehicle charging station locations*. <https://afdc.energy.gov>
- U.S. Environmental Protection Agency. (2023). *Greenhouse Gas Inventory Data Explorer*. <https://cfpub.epa.gov/ghgdata/>
- UN-Habitat. (2023). *World cities report*. <https://unhabitat.org>
- UN-Habitat. (2024). *Sustainable urban mobility indicators*. <https://unhabitat.org>
- Waymo LLC. (2023). *Waymo Open Dataset: Motion dataset*. <https://waymo.com/open/>
- Waymo LLC. (2024). *Safety performance report 2024*. <https://waymo.com/safety/>
- Waymo LLC. (2025). *Safety performance data*. <https://waymo.com/safety/performance/>
- Waymo LLC. (2025). *Waymo Open Dataset*. <https://waymo.com/open/>
- World Economic Forum. (2023a). *Autonomous vehicles readiness index*. <https://www.weforum.org/publications>

- World Economic Forum. (2023b). *Global Future Council on Autonomous Mobility: Data and ethics framework for self-driving vehicles*.
<https://www.weforum.org/publications>