

CLOUDERA

EBOOK

Driving Automotive Success

Creating innovation and value
by moving from connected to
agentic vehicles.



Table of Contents

Executive Summary	3
Connected Vehicles: An Opportunity for Differentiation	3
Data Monetization	4
Connected Vehicles Are a Big Data Challenge	4
Harnessing IoT Data for Driver and Business Value at Scale	5
The Connected Vehicle AI Lifecycle	6
Conclusion: Making the Leap from Connected to Agentic	7
Learn More	7
About Cloudera	8

Executive Summary

The automotive industry has experienced an intense hockey-stick acceleration of technological innovation. After nearly a century of relatively little change, modern automakers are experiencing both massive disruptions and exciting opportunities. Some of these include the proliferation of electric and hybrid engines and the infrastructure needed to support them, the slow but steady rollout of driverless vehicles, and the increased connectivity and digitalization of the driver experience.

Of course, with these opportunities comes increased complexity. Data is the key to success with connected and autonomous vehicles, and the primary challenge for automakers is to implement a data architecture that can support the massive volumes of telemetry and other data, process it all, and make decisions in real time.

This eBook explores the opportunities for automakers to differentiate based on their connected vehicle experience, as well as data monetization strategies. Then, we will discuss data management and AI challenges and opportunities presented by connected vehicles, and provide a blueprint for implementing a data architecture that can tackle those challenges. Finally, we will look ahead to see how generative and agentic AI will introduce new innovations and customer experiences that shape the next generation of connected vehicles.



Connected Vehicles: An Opportunity for Differentiation

The driving experience is increasingly powered by data, from cameras and sensors that alert drivers when they are drifting into another lane or automatically turn on the headlights, to GPS that tracks location and route information, to mobile apps that share performance and maintenance data and control the vehicle remotely.

Even as so many vehicle features become commoditized, there is still a massive opportunity for automakers to differentiate based on the connectivity and intelligence they offer.

Data Monetization

While the focus of the data is primarily on enhancing driver safety and convenience features, there is also an opportunity for data monetization.

- **Usage-Based Insurance (UBI):** Traditional insurance relies on static demographics and personal information. With real-time telematics measuring how a person actually drives, such as harsh braking, cornering speeds, and time-of-day usage, automakers can partner with insurers to offer risk-based premiums that lower costs for safe drivers and shift the relationship between the insurer and the driver from a one-time transaction to a value-added service that encourages safe driving behaviors.
- **Fleet as a Service (FaaS):** As the automotive economy shifts toward access over ownership, automakers are moving beyond selling hardware to providing FaaS. For commercial operators and delivery providers, a vehicle represents uptime. FaaS leverages connectivity to provide condition-based maintenance, where AI monitors a vehicle's health to predict failures before they happen, and smart routing optimizes fuel and battery life based on real-time traffic and weather. By providing an integrated ecosystem of hardware, software, and actionable insights, automakers can secure high-margin, recurring revenue streams from enterprise customers, and customers can shift heavy capital expenditures to operating expenditures that won't depreciate.

If automakers can successfully harness the massive volumes of data generated by connected vehicles, they can deliver innovative differentiation and create new revenue streams.

Connected Vehicles Are a Big Data Challenge

With so many different data sources, vehicles are quickly becoming data centers on wheels. Harnessing all of that data and making use of it requires some specific capabilities to successfully deliver a connected vehicle experience, and the stakes are incredibly high.

The primary challenges for automakers delivering connectivity are:

- **Scale and Performance:** A single vehicle can produce as much as one terabyte of data every day. Automakers need a platform that can scale to manage multiple petabytes of data and train machine learning and AI models on that data.
- **Edge Processing and AI:** Many connected vehicle use cases require autonomous, near-real-time decision making. Models must be deployed at the edge to operate on telemetry streams, and fault tolerance is a necessity, as vehicles may lose connectivity in remote locations, but they must still function properly.
- **Security and Governance:** Driver and vehicle data is extremely sensitive, and ensuring data and AI model quality is paramount. A data fabric is critical for ensuring unified and consistent security and governance across the data and AI lifecycle while ensuring compliance with data privacy regulations.

- **Hybrid and Multi-Cloud Complexity:** Automotive data doesn't live in one place. R&D might happen in a private cloud, while consumer-facing applications run on a public cloud. Moving petabytes of sensor data between these environments without paying prohibitively expensive egress fees or managing gaps in data lineage is an architectural hurdle.
- **Data Silos and Interoperability:** Historically, telematics data, manufacturing logs, and customer CRM data lived in separate silos. To address many connected vehicle use cases, automakers must find a way to integrate these datasets for a 360-degree view of the vehicle and the driver.

To deliver the next generation of intelligent, connected vehicles, automakers must adopt a distributed data architecture that supports data management, processing, and analytics and AI at the edge and data science in the cloud or data center.

Harnessing IoT Data for Driver and Business Value at Scale

Cloudera provides a platform that delivers the capabilities needed to manage data from vehicles at scale, build AI and ML models for a wide range of use cases, and deliver real-time intelligence at the edge. Automakers around the world are leveraging Cloudera for their connected vehicle workflows.

Streaming Data at the Edge: MiNiFi collects data from cameras, radar, LiDAR, GPS, and other sensors directly on the vehicle's gateway so automakers can shift intelligence to the source. MiNiFi filters telemetry noise before passing critical signals to the data center.

Analytics and Data Science at the Core: Data passes from vehicles to the data center or to the cloud for a variety of use cases:

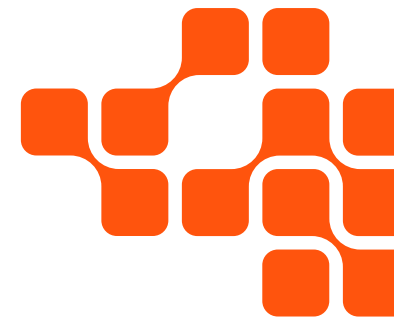
- Apache Flink performs complex, stateful cross-fleet analysis in real time.
- Data is stored in the data lakehouse for enrichment and analysis.
- Data science teams build, train, and retrain models based on fresh and historical data.
- R&D experiments with real data to develop new features and functionality.



Distributing Intelligence Back to the Edge:

Updated models are passed back to the edge via an Over-the-Air (OTA) software update.

A Unified Data Fabric: A consistent framework for security and governance, including end-to-end lineage, is critical for ensuring the quality and safety of customers and their data throughout the data and AI lifecycle. Cloudera Shared Data Experience (Cloudera SDX) ensures the integrity of the entire data estate, from creation to deletion.



The Connected Vehicle AI Lifecycle

Looking Ahead: The Generative and Agentic AI Future of Connected Vehicles

The previous generation of connected vehicles relied on a cycle of capturing data and waiting for the intelligence. The next generation will transform vehicles into proactive agents. By moving from passive ML to generative and agentic AI, vehicles will partner with drivers for a truly connected driving experience.

Advanced Driving Intelligence

Rather than a single, monolithic black box model for self-driving, newer architectures will use a squad of specialized agents. With selective activation, only agents relevant to the current situation, such as a Heavy Rain Agent or a Pedestrian-Heavy Intersection Agent, will be active. This architecture will reduce the massive compute load necessary for level 3 and 4 autonomy, enabling advanced safety features to run on standard mass-market vehicle hardware.

The Digital Concierge

Infotainment is evolving into an agentic orchestrator. AI-powered systems now complete tasks in addition to answering questions. An orchestrator may give a contextual answer to a question based on route and weather information, and an agent may make a gas station recommendation based on fuel level and pricing.

Autonomous Service and Maintenance

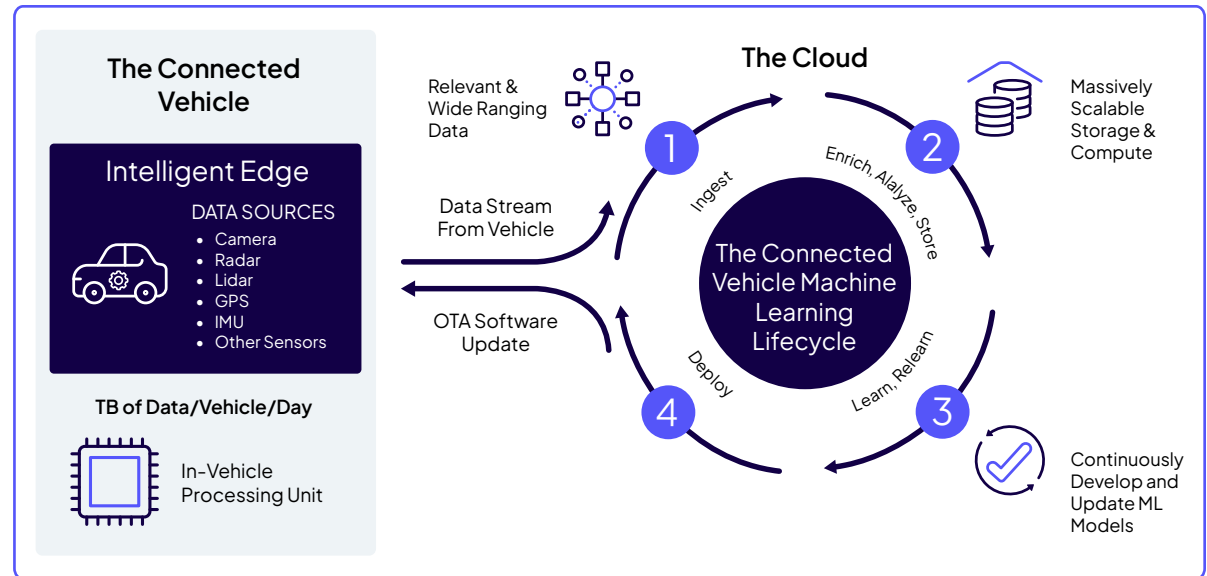
Connected vehicles can use agents to close the loop between detecting a maintenance issue and resolving it. Agents can monitor sensors in real time for maintenance issues, and schedule service for the driver.

Vehicle-to-Everything (V2X) Coordination

In a connected ecosystem, agents can act as negotiators between the vehicle and the city. For electric vehicles, agents can interface with smart grids to sell surplus energy back during

peak hours or delay charging until electricity is cheapest, effectively turning the car into an autonomous energy trader. For traffic flow, agents can coordinate with infrastructure to suggest speeds that ensure optimal travel, reducing idling and emissions.

By harnessing big data and leveraging an edge-to-AI architecture, automakers can develop the generative and agentic AI functionality needed to deliver the next wave of connected innovation.

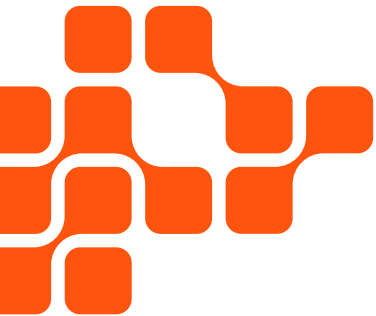


Conclusion: Making the Leap from Connected to Agentic

The transition from connected to agentic vehicles requires a fundamental shift in how automakers manage their most valuable asset: their data. By harnessing massive volumes of telemetry data from their vehicles and building a unified, edge-to-cloud data architecture with Cloudera, vehicle manufacturers are delivering innovation and value with a truly differentiated connected vehicle experience today, while building the resilient, scalable foundation necessary to lead the next century of mobility.

Learn more about how Cloudera supports manufacturing and automotive companies here:

trycloudera.com



About Cloudera

Cloudera is the only hybrid data and AI platform company that large organizations trust to bring AI to their data anywhere it lives. Unlike other providers, Cloudera delivers a consistent cloud experience that converges public clouds, on-prem data centers, and the edge, leveraging a proven open-source foundation. As the pioneer in big data, Cloudera empowers businesses to apply AI and assert control over 100% of their data, in all forms, improving security, governance, and real-time and predictive insights. The world's largest brands across all industries rely on Cloudera to transform decision-making and ultimately boost bottom lines, safeguard against threats, and save lives.

To learn more, visit [Cloudera.com](https://cloudera.com) and follow us on [LinkedIn](#) and [X](#).

CLUDERA

Cloudera, Inc. | 6220 America Center Dr, 5th Fl, San Jose, CA 95002 USA | cloudera.com

© 2026 Cloudera, Inc. All rights reserved. Cloudera and associated marks are trademarks or registered trademarks of Cloudera, Inc. All other company and product names may be trademarks of their respective owners.
9128-001 April 28, 2026

