

Akridata and Dell Technologies

Edge-to-core-to-cloud data pipelines for Advanced Driver Assistance Systems and Autonomous Driving (ADAS / AD)

Solution benefits

Easily manage AI data across edge core and cloud.

- Understand your data at source instantly. Extend the core/cloud to edges with “smart edges.”
- Unified, transparent access to data (active and archived) via global, user-defined catalog.

Eliminate “data blindness” with automated workflows.

- Evaluate data quality and relevance at source. Transfer it by priority to core data centers and/or cloud.
- Browse, search, visualize petabytes of data. Track data provenance and lineage.

Build pre-tested, customer-proven smart edges and core infrastructure optimized for ADAS / AD.

- Akridata software optimized on Dell EMC PowerEdge line of servers.
- Integrated with Dell EMC UDS solutions, ECS and PowerScale, at edge locations and in the core.

Data challenges of large scale AI/ML

Artificial intelligence (AI) and machine learning (ML) are transforming enterprises by increasing productivity, reducing costs, streamlining operations and fueling innovation. The rising demand for automation and remote-enabled operations have further highlighted the value of AI, which is growing rapidly in sectors such as automotive, manufacturing, health care and life sciences, retail and smart cities. Deploying AI at scale in these sectors requires an abundance of data for training AI algorithms and models. But managing AI data at scale presents many challenges. Besides the sheer volume, complexity and diversity of data, requirements and dependencies change constantly as data drifts, use cases evolve, and models have to adapt.

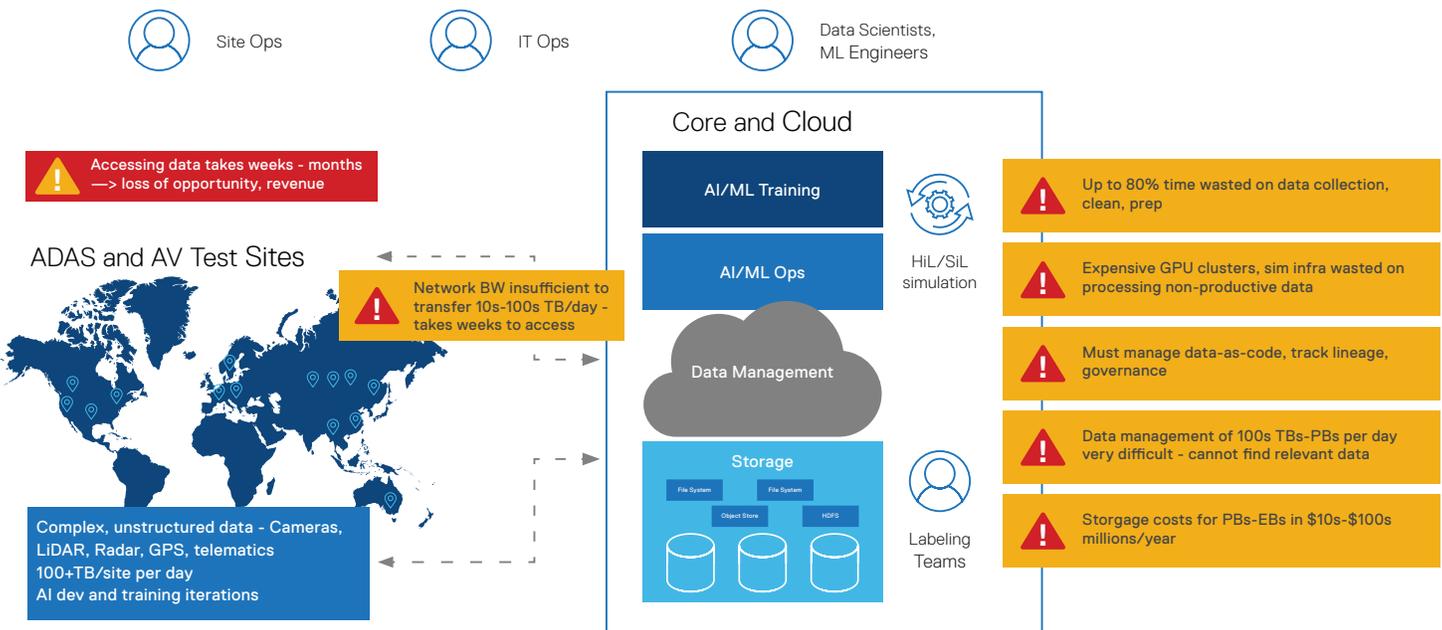
In the automotive industry, the Society of Automotive Engineers (SAE) has defined six levels of vehicle autonomy (0-5) with level 0 offering no automation. Levels 1-3 offer driver safety features (ADAS) and levels 4 and 5 offering full autonomous driving (AD) in predefined vs. any environment, respectively. AI is integral to delivering safety, comfort and autonomy in ADAS and AD vehicles. Vehicles have to navigate an incredibly dynamic environment, with varying traffic conditions, weather, lighting, and other variables that can change rapidly. AI models and ML algorithms used by the vehicle’s perception, prediction and planning systems have to be capable of adapting to such a changing environment. As ADAS/AD development

matures, developers' requirements evolve from needing more data to needing more relevant data, which captures complex and unique scenarios. For example, developers may be interested in collecting data about specific traffic conditions, such as the presence of a large truck blocking visibility at an intersection. While the total data collected may be several petabytes, often the relevant subset (scenes with trucks in intersections) is a small portion, typically between 1%-10%. Finding such subsets is extremely difficult, like looking for a specific needle in a haystack of needles.

Auto OEMs and Tier 1s face the following key data challenges when they get into large-scale development and deployment of ADAS and AD capabilities:

- Ingest of hundreds of terabytes to petabytes of data per day from each test site (edge locations) across multiple regions nationally and internationally.
- Processes that are “data blind,” meaning the value of data is unknown until all data is processed, annotated and analyzed.
- Data scientists and ML engineers that must wait several weeks to access the data collected at the edge locations.
- Petabytes of data that are stored every month, incurring millions of dollars in storage costs alone, even though only 1%-10% is really relevant (as mentioned above).
- Data science teams that must spend 80% of their time on data management issues, such as accessing, cleaning and converting data, leaving little time for algorithmic development.
- Continuously evolving data requirements that force data pipeline changes, which are complex and time-consuming, and delay data availability by 1-3 months or more.
- Tracking and governance of data with current processes that are extremely difficult.

Autonomous driving data challenges

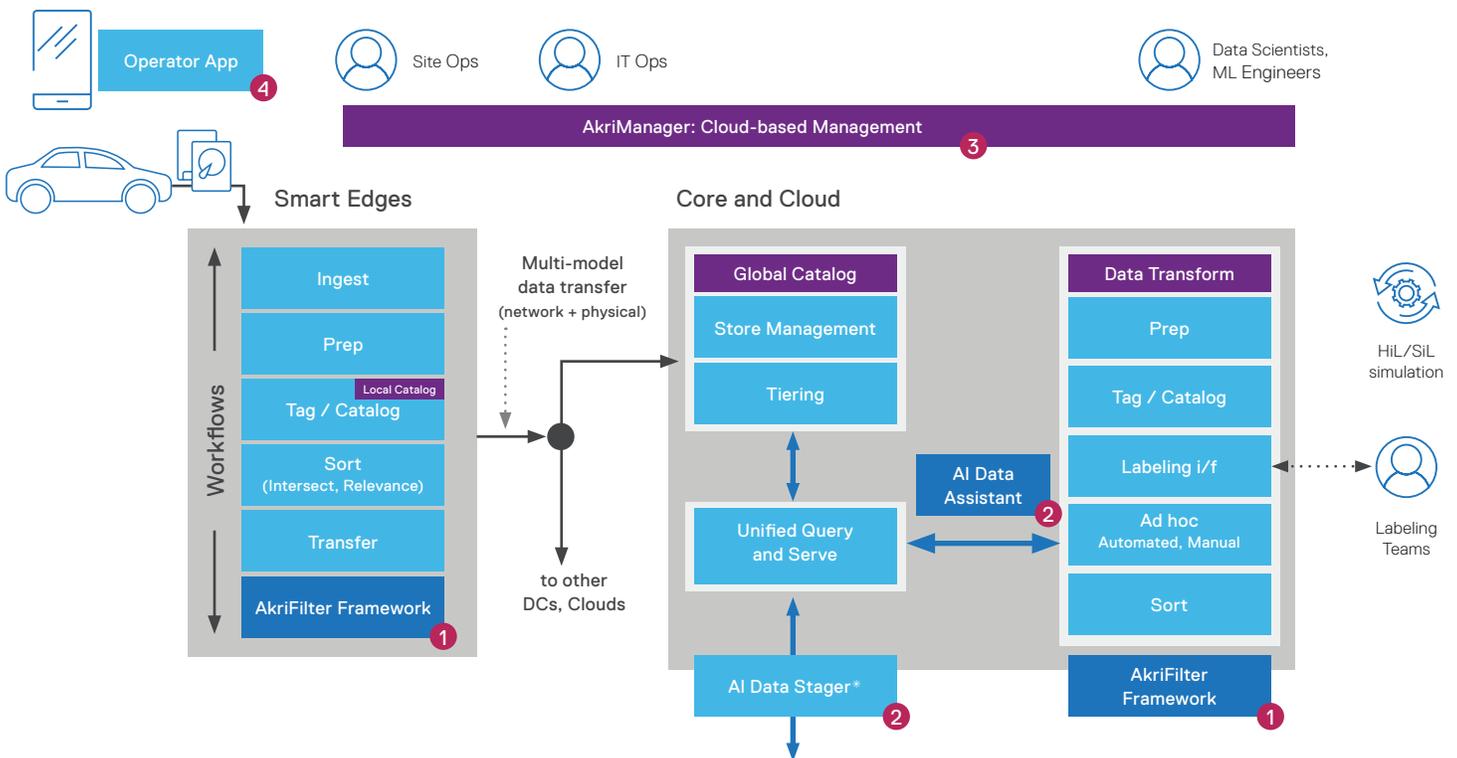


These challenges are magnified as multiple teams (vehicle/site operators, IT operation managers, and data science/ML developers) at automotive OEMs and Tier 1s work on multiple products across multiple locations simultaneously. These problems are further compounded as vehicles move from development and testing into production.

Additionally, as the use of AI systems proliferates development progresses, and data becomes code, we have to bring the same degree of discipline and sophistication as MLOps used for machine learning to data management.

Akridata and Dell Technologies: Edge-to-Core-to-Cloud DataOps at scale

Akridata helps automotive OEMs and Tier 1s overcome their data challenges with a platform for efficient, agile and scalable edge-to-core data pipelines, which meshes tightly with their MLOps setup. Built on Dell Technologies infrastructure, the Akridata DataOps platform provides benefits for data scientists, ML engineers, simulation engineers, external and internal labeling teams, IT operations and vehicle/site operations.



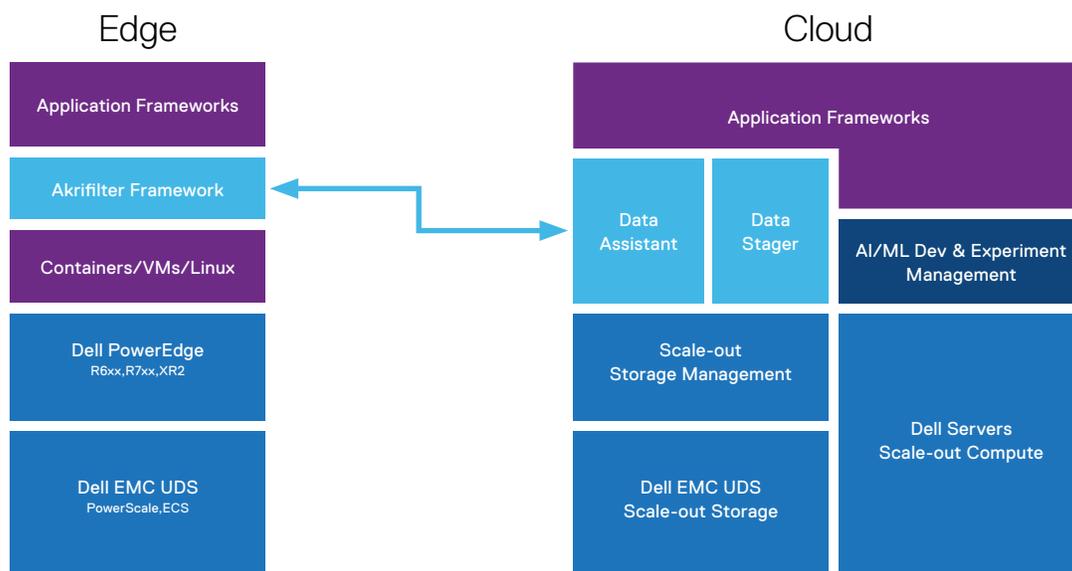
The DataOps platform allows you to:

- Streamline processing from test sites/edge locations to data centers and cloud with user-composable, automated workflows. This includes data cleaning, pre-processing, tagging, prioritizing data based on content, user-specified relevance and quality parameters.
- Reuse existing software assets, including code (java, C/C++, python), with any ML framework, libraries, executables and containers (e.g. Tier 1s sensor drivers).
- Run server-side data processes at edge locations so that users can understand their data instantly, within minutes to hours of ingest.
- Utilize powerful capabilities for statistical sampling, similarity, clustering and visualizing data.

- Browse, search, and access data in the system from any edge location, data center or cloud tier with a user-defined, global data catalog.
- Support data lineage and versioning for governance and forensic analysis.
- Deploy, monitor and manage edge sites, users, edge assets and workflows with cloud-based management tools.
- Provide “Self-service” for AI/ML users (data science/ML teams) and enhanced collaboration between various teams.
- Leverage an agile and scalable platform for test vehicles (prototype, data collection, FOT) and production vehicles.
- Deploy across hybrid environments, spanning on-premises data centers as well as public clouds.

Dell Technologies and Akridata – delivering DataOps for automotive AI development

Akridata’s solution is optimized to work with Dell Technology’s best-in-class data center infrastructure for compute and storage. Dell EMC PowerEdge servers provide a variety of operating configurations with powerful CPUs and GPUs. Ranging from 1 to 4 U, the PowerEdge family provides the flexibility to deploy the server that’s right for you, from compact edge deployments to dense data center compute. On the storage side, Dell Technologies provides the best options for users with high-performance, scale-out NAS solutions (Dell EMC PowerScale) and S3-compatible, cloud-friendly object storage solutions (Dell EMC ECS) that can scale from terabytes to petabytes without disruption – critical for time-critical applications like autonomous driving where sensor data is being ingested 24x7.



About Dell Technologies

Dell Technologies has been a leader in the advanced computing space for over a decade, delivering proven products, solutions and expertise. Dell Technologies has a team of data analytics, HPC and AI experts dedicated to staying on the cutting edge, testing new technologies and tuning solutions to your applications to help you keep pace with this constantly evolving landscape. With an extensive portfolio, years of experience, and an ecosystem of curated technology and service partners, Dell Technologies provides innovative solutions, workstations, servers, networking, storage and services that reduce complexity and enable you to capitalize on the promise of data analytics, HPC and AI.

Click below to learn more about Dell Technologies storage solutions for Automotive Applications

Delltechnologies.com/automotive

About Akridata

Akridata provides a unique edge-to-core-to-cloud DataOps solution for large-scale AI/ML development in sectors like automotive, manufacturing, health care, life sciences and retail. Akridata's cloud-managed platform helps data scientists, ML engineers, IT departments and vehicle/site operations teams. Benefits include shortening time to "good models," multiplying data science productivity, and lowering IT costs for large-scale AI/ML training, deployment and analytics. Akridata is based in Silicon Valley and backed by leading venture firms like Accel and Menlo Ventures.

Click below to learn more about Akridata

Akridata.com