Data pipeline

A managed, high-speed data pipeline that dramatically accelerates the ingestion and transformation of raw fleet data into the high-quality, AI-ready data necessary to help develop the next generation of autonomous vehicle technology

THE CHALLENGE

- The data pipeline for AV is an important enabler for the efficient development of high-quality AI–"the AI is only as good as the data"
- As the AV industry pushes toward L3 autonomy, data volumes are expected to dramatically increase by 10X, model complexity by 4X, and compute/storage costs by 4X
- However, today's data pipeline architectures are not built for the level of scalability that may be required to ingest 2 PB of data per day while meeting critical data quality standards
- AI-ready data needs high-quality labels for DNN development, and manually annotating can be slow and cumbersome
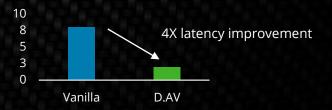
SOLUTION HIGHLIGHTS

- Production-grade, AI-enabled, fully-automated L3 data management platform from data ingestion through curation
- Observability and built-in intelligence for anomaly detection and remediation for quality assurance
- Full-featured, AI-powered auto-labeling platform using pretrained AV-specific models to accelerate labeling

VALUE PROPOSITION

- Up to 4X latency improvement
- Up to 99% data quality SLA
- Flexibility (deployable on cloud/hybrid cloud/on-prem)

LATENCY IMPROVEMENT

4X latency improvement

DEPLOYMENT OPTIONS

PLATFORM-AS-A-SERVICE	BESPOKE IMPLEMENTATION + OPERATE	IMPLEMENTATION
• Ready-to-use Deloitte-hosted service with data processing and curation functionality • Reporting and analytics on platform usage and performance	• Custom development + existing data processing and curation functionality • Seamless integration into existing toolchain • Reporting and analytics on platform usage and performance • Prelabeling functionality included • Ongoing training and support • Monitoring, troubleshooting, and maintenance	• Custom development + existing data processing and curation functionality • Seamless integration into existing toolchain • Reporting and analytics on platform usage and performance