aiData

Data Quality in Multi-sensor Integration for Physical AI

Lessons from Automotive and Real-World Data applications

Frida Kobor



18 September 2026
aimotive.com



Who are we?

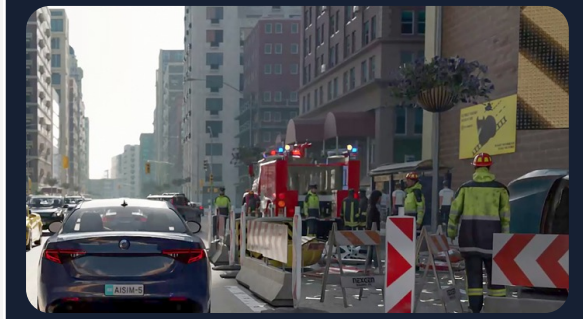
10+ years of experience in developing an ADAS/AD software stack and tooling for development and validation

Today, our technology portfolio covers the key building blocks of ADAS/AD development:

- **aiDrive** — aiDrive, an ADAS/AD solution, offering modular, hardware-agnostic software stack for SAE Level 2-4 ADAS & AD.
- **aiSim** — aiSim, the world's first ISO 26262-certified automotive simulator, enables AI-based digital twin creation and highly optimized sensor simulation - while ensuring deterministic, reproducible environments.
- **aiData** — aiData, an end-to-end data pipeline, enables a streamlined workflow for collecting, curating, auto annotating, managing and evaluating data to for ADAS/AD development and validation.
- **aiWare** — aiWare, an automotive NN acceleration hardware IP that support efficient deployment of AI workloads.

This experience has shown us that the quality of our solutions is fundamentally determined by the quality of the underlying data.





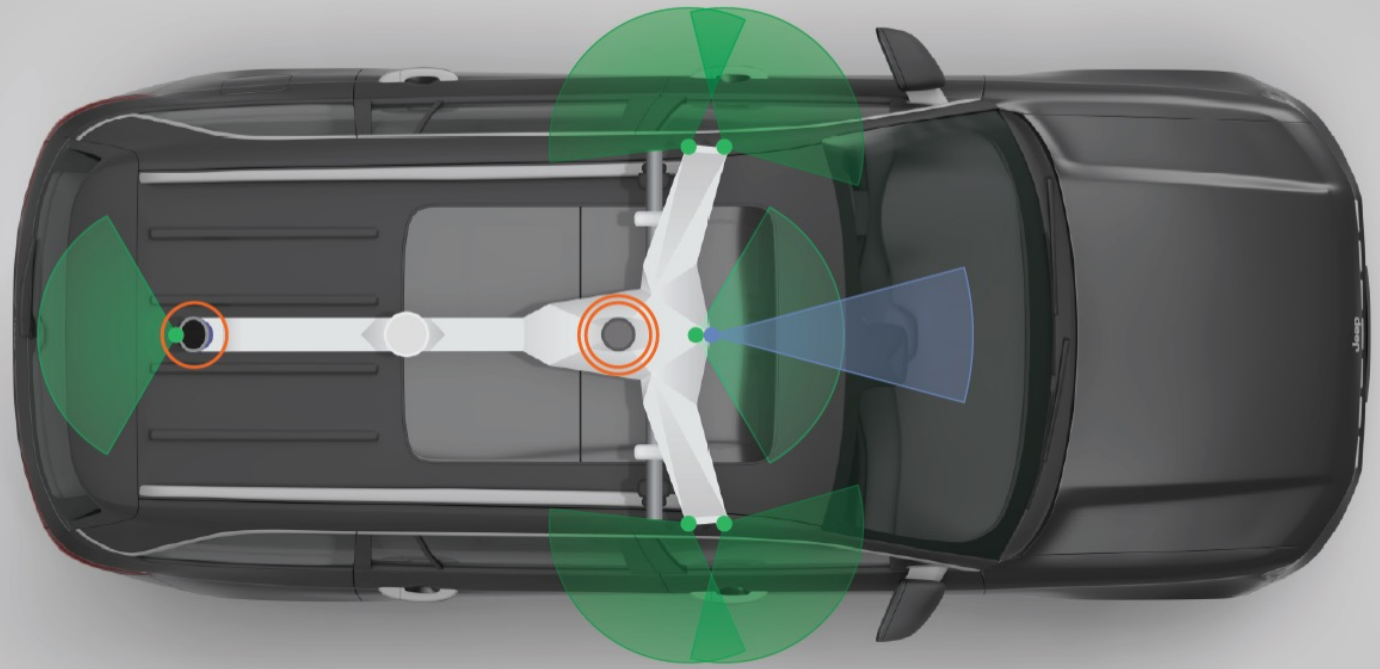




What have we learned?

Standardized sensor setup for data collection from training through validation to neural reconstruction

aiMotive reference sensor setup



Sensors:

- **6 x Sony IMX728 120° cameras**
- Optional: 1 x Sony IMX728 30° camera for traffic sign and traffic light detection
- **1 x Hesai OT128 LiDAR**
- **1 x Hesai XT32M2X LiDAR**
- **1 x Novatel PwrPak7D-E2 GNSS/IMU**

Yaw camera offsets:

- **Front Midrange camera 0°**
- **Cross traffic cameras +/- 60°**
- **Neighborlane cameras +/- 120°**
- **Back Midrange camera 180°**

aiData by aiMotive

EVALUATION TEMPLATE FOR

Auto Annotation & Neural Reconstruction

Recording available

- Minimum 1 minute
- Vehicle moving
- Vehicle standing still
- Vehicle turning
- Multiple pass
- *Drive coverage check (NR)*
- Sensor rigidity check
- Environment check

Camera available

- Format converted
- Color space check
- *360 degree with non-fisheye (optional) - NR*
- Frame rate check
- No lens undistortion
- No motion blur
- Frame drop check
- Camera settings check

Lidar available

- Format converted
- Compatible device
- Rotation speed check
- Unprocessed points check
- Lidar range check
- Frame drop check

GNSS available

- Format converted
- Compatible device
- Unprocessed GNSS data
- Signal check (N/A for processed data)
- Correction signal coverage (N/A for processed data)
- Frame drop check (N/A for processed data)
- Egomotion precision check

Calibration available

- Camera intrinsics available
- Sensor extrinsics available
- Calibration check

Sync available

- Camera timestamps available
- Lidar timestamps available
- GNSS timestamps available
- Synchronization check

Legend | Complete / Checked | Warning / Non-ideal | Not available / info needed | To be done

EVALUATION TEMPLATE FOR

Auto Annotation & Neural Reconstruction

Recording available

- Minimum 1 minute
- Vehicle moving
- Vehicle standing still
- Vehicle turning
- Multiple pass
- *Drive coverage check (NR)*
- Sensor rigidity check
- Environment check

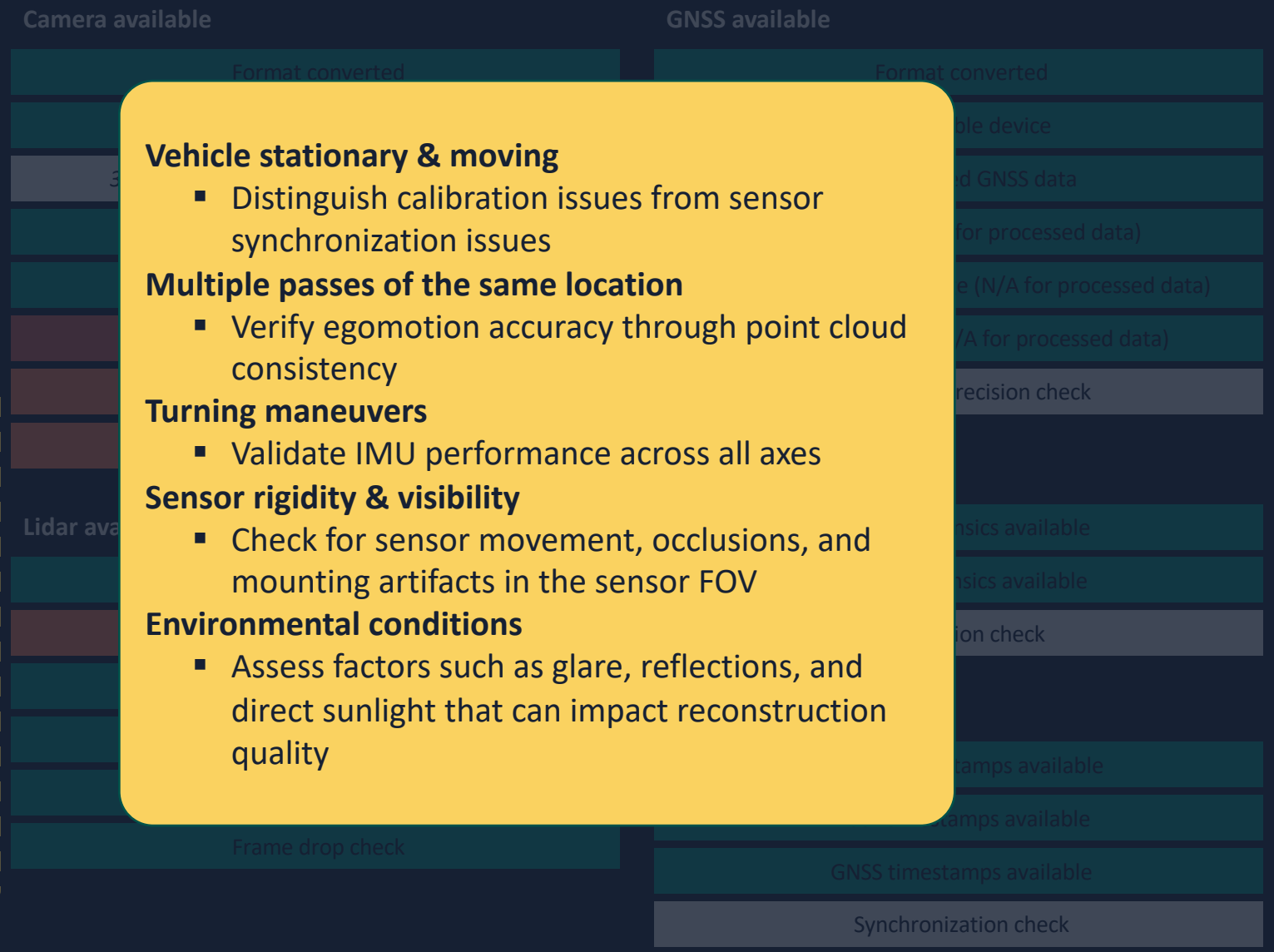
Vehicle stationary & moving
- Distinguish calibration issues from sensor synchronization issues

Multiple passes of the same location
- Verify egomotion accuracy through point cloud consistency

Turning maneuvers
- Validate IMU performance across all axes

Sensor rigidity & visibility
- Check for sensor movement, occlusions, and mounting artifacts in the sensor FOV

Environmental conditions
- Assess factors such as glare, reflections, and direct sunlight that can impact reconstruction quality



EVALUATION TEMPLATE FOR

Auto Annotation & Neural Reconstruction

Recording available

- Minimum 1 minute
- Vehicle moving
- Vehicle standing still
- Vehicle turning
- Multiple pass
- *Drive coverage check (NR)*
- Sensor rigidity check
- Environment check

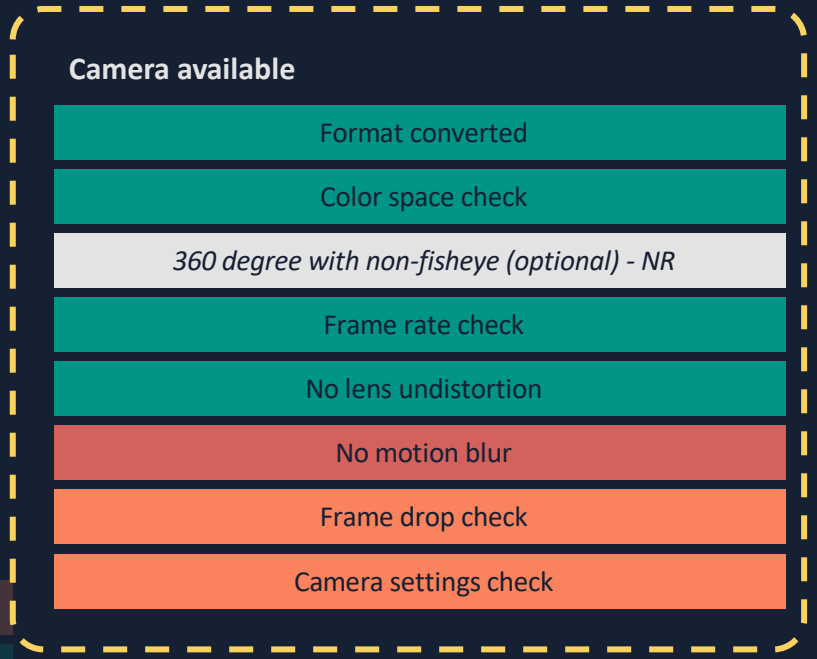




Legend: Complete / Checked | Warning / Non-ideal | Not available / info needed | To be done

EVALUATION TEMPLATE FOR

Auto Annotation & Neural Reconstruction



Raw camera data
- Eliminates ambiguity from unknown preprocessing steps
- Provides a transparent baseline across teams, suppliers, and projects

Color space
- Verify compatibility with downstream algorithms
- Ensure critical features remain visible (e.g., traffic lights)

Frame rate & temporal consistency
- Check FPS and ensure no jitters
- Supports reliable tracking and temporal reasoning

No lens undistortion
- Avoid pre-applied lens undistortion where possible
- Retain maximum information and minimize interpolation artifacts

Motion blur & frame drops
- Motion blur causes irreversible information loss
- Dropped frames create temporal gaps that impact perception and reconstruction

Camera settings
- Review exposure, focus, HDR, and image processing parameters
- Prevent inconsistencies due to automatic processing algorithms within and between recordings

Recording available
Minimum 1 minute
Vehicle moving
Lidar available
GNSS timestamps available
Synchronization check

Legend: Complete / Checked | Warning / Non-ideal | Not available / info needed | To be done

EVALUATION TEMPLATE FOR

Auto Annotation & Neural Reconstruction

Compatible device
- Verify sensor capabilities and suitability for the target application
- Review beam count, angular resolution, coverage, and intensity characteristics

Rotation speed check
- Ensure sufficient and stable scan frequency
- Minimize motion-induced distortions

Unprocessed points check
- Preserve original point measurements
- Retain per-point timing information for motion compensation

LiDAR range check
- Verify detection range meets application requirements
- Assess coverage and potential blind spots

Frame drop check
- Detect missing scans or scan sectors
- Prevent gaps in spatial and temporal coverage

Recording available
Minimum 1 minute
Vehicle moving
Vehicle standing still
Vehicle turning
Multiple pass
Drive coverage check (NR)
Sensor rigidity check
Environment check

Camera settings check

Lidar available
Format converted
Compatible device
Rotation speed check
Unprocessed points check
Lidar range check
Frame drop check

GNSS available
Format converted
GNSS timestamps available
Synchronization check

Legend: Complete / Checked | Warning / Non-ideal | Not available / info needed | To be done

EVALUATION TEMPLATE FOR

Auto Annotation & Neural Reconstruction

Recording available

- Minimum 1 minute
- Vehicle moving
- Vehicle standing still
- Vehicle turning
- Multiple pass
- *Drive coverage check (NR)*
- Sensor rigidity check
- Environment check

Compatible device:
- Assess the capabilities of the hardware components
- Assess the firmware configurations

Unprocessed GNSS data:
- Avoid black-box positioning outputs
- Enable independent post-processing

Signal check:
- Verify tracked satellite constellations
- Check for interference

Correction signal coverage:
- Check base station availability
- Enable high-accuracy localization

Frame drop check:
- Detect missing GNSS and IMU measurements
- Avoid drift in trajectory

Egomotion precision check:
- Validate localization accuracy after post-processing
- Asses trajectory consistency and precision

GNSS available

- Format converted
- Compatible device
- Unprocessed GNSS data
- Signal check (N/A for processed data)
- Correction signal coverage (N/A for processed data)
- Frame drop check (N/A for processed data)
- Egomotion precision check

Calibration available

- Camera intrinsics available
- Sensor extrinsics available
- Calibration check

Sync available

- Camera timestamps available
- Lidar timestamps available
- GNSS timestamps available
- Synchronization check

Legend: Complete / Checked | Warning / Non-ideal | Not available / info needed | To be done

EVALUATION TEMPLATE FOR

Auto Annotation & Neural Reconstruction

Camera intrinsics available
- Verify availability of intrinsic parameters (focal length, principal point, distortion)

Sensor extrinsics available
- Verify spatial relationships between cameras, LiDAR, IMU, and GNSS
- Ensure all sensors can be transformed into a common reference frame

Calibration check
- E.g., validate camera-LiDAR alignment using projected point clouds
- Verify consistency between observed vehicle motion and IMU measurements

Synchronization check:
- Verify that timestamps are available, interpretable, and referenced to a common time source

Recording available
- Minimum 1 minute
- Vehicle moving
- Vehicle standing still
- Vehicle turning
- Multiple pass
- *Drive coverage check (NR)*
- Sensor rigidity check
- Environment check

GNSS available
- Format converted
- Compatible device
- Unprocessed GNSS data
- Signal check (N/A for processed data)
- Correction signal coverage (N/A for processed data)
- Frame drop check (N/A for processed data)
- Egomotion precision check

Calibration available
- Camera intrinsics available
- Sensor extrinsics available
- Calibration check

Sync available
- Camera timestamps available
- Lidar timestamps available
- GNSS timestamps available
- Synchronization check

aiMotive target

0.1°	1 msec
calibration accuracy	synchronization accuracy

Legend: Complete / Checked | Warning / Non-ideal | Not available / info needed | To be done

EVALUATION TEMPLATE FOR

Auto Annotation & Neural Reconstruction

Camera intrinsic error - focal length error



GNSS to camera extrinsic yaw error – lane projection on camera image



Calibration available

- Camera intrinsics available
- Sensor extrinsics available
- Calibration check

Sync available

- Camera timestamps available
- Lidar timestamps available
- GNSS timestamps available
- Synchronization check

Recording available

Legend: Complete / Checked | Warning / Non-ideal | Not available / info needed | To be done



The end result

aiData by aiMotive

High-quality training and validation ready datasets with the aiData pipeline

aiData Auto Annotator





95%
precision & recall – superhuman performance

98%+
cost saving, compared to manual work

400x
faster compared to manual annotation

aiData™ by aiMotive

aiData by aiMotive

Scalable, photorealistic environments for simulation, based on neural reconstruction

3D Gaussian Splatting
of a parking lot in South Korea



US highway scene in aiSim
using our hybrid rendering



aiSim™ by aiMotive



Conclusions

Conclusions

- **A structured review** can reveal critical issues before development begins.
- **Many downstream algorithm failures originate from data quality** problems rather than model limitations.
- **Systematic dataset assessment helps reduce development risk**, improve reproducibility, and accelerate project onboarding.

Call for Standards

- We see a growing need for standardization in exchanging multi-sensor data.
- Calibration is a practical starting point, as formats and camera models vary significantly across datasets and organizations.
  - Standardizing intrinsic and extrinsic calibration descriptions would significantly improve interoperability and reduce integration effort.
- Together with **ASAM** and **Deepen AI**, we are exploring how to initiate this work and build the foundation for broader automotive data standards.

High-quality AI starts with high-quality data. Industry-wide standards for sensor data documentation and quality assessment are essential for scalable and trustworthy autonomous driving development.

Contact us

aiData by aiMotive

info@aimotive.com

+36 1 7707 201

Budapest, Hungary HQ
Szépvölgyi út. 18-22., 1025 Budapest, Hungary
+36 (1) 7707 201

Mountain View, USA
1907 Colony Street, Mountain View, CA, 94043
+1 (650) 282 5085

Yokohama, Japan
#501 Expert Office Shinyokohama, 3-7-18 Shinyokohama,
Kohoku-ku, Yokohama,
Kanagawa 222-0033
+81 (45) 620 6100



aimotive.com