Data Labeling and Standards
Santa Clara, CA – Hyderabad, India

Data Curation, Labeling and Validation for Autonomous systems

Daimler, Samsung, Pony AI, DeepScale, Starsky, Nuro, Marble, Uber, AutoX...

Nicola Croce
Technical Program Manager
Project Lead OpenLABEL, ASAM TSC
nico@deepen.ai
Data Labeling

3D Bounding Boxes
Individual frames or unlimited-length sequences with consistent size and ID.

Sensor Fused Sequence Annotation
2D and 3D labels in multi-sensor sequences with consistent ID.

Point Cloud Sequence Segmentation
State of the art instance or semantic segmentation of sequences.

Polylines
Accurately defined lane lines with rigorous quality checks.

2D Bounding Boxes
Precise object detection and localization in images and videos.

2D Semantic Segmentation
Pixel-perfect semantic segmentation tasks at scale.

2D Landmarks and Key Points
Transformative ground-truth dataset with a sequence of points.

Data Labeling = Ground truth production
Problem

Source: https://medium.com “The very first standard we need for Autonomous Vehicles”
Data Labeling

Source: https://medium.com “The very first standard we need for Autonomous Vehicles”
OpenLABEL: Intro

OpenLABEL: An open standard for data labeling
OpenLABEL: Use Cases

Technical Use Cases
- ML Model Training
- ML Model Validation
- Semantic Ground Truth
- Query-able Ground Truth
- Standardize Tooling

Business Use cases
- Dataset sharing
- Tooling development cost saving
- Enable scenario sharing/sorting/filtering
- Increase labeling quality
- Facilitate data marketplaces

End-user use cases
- Standardize HIL Processes
- Streamline vendor selection
- Reduce tooling implementation effort
- ODD testing (through search and filtering)
OpenLABEL: Work Packages

WP 1: Annotation Format and metadata
- concept for an annotation format including a proposed schema
- Metadata labeling specs
- The WP will consider existing labeling specs and formats.

WP 2: Labeling methods
- description of identified labeling methods for objects
- Labeling specs (instructions) according to labeling method
- draft description of how to use specific methods for different use cases

WP 3: Taxonomy & structure
- usage of ontologys for OpenLABEL
- list of requirements for the ASAM ontology project
- example list of required objects (entities+attributes) and labels

WP 4: Scene/ scenario labeling
- Scoping of labeling methods for objects for activities, events and scenes
- Labeling specs (instructions) according to labeling method
- Semantic concepts definitions and their structure
Semantic labeling

- Objects taxonomy
- Actions, Events taxonomy
- Scene attributes taxonomy
- Relations
- Constraints

Algorithms

- Automatic Annotation
- Manual Annotation (Objects)
- Manual Annotation (Actions, relations)

User Interface

- OpenXOntology
  - ./automotive.owl
- OpenLABEL
  - .json
- Structure file
  - .json

Data recording

- Inspection of sequences
- Annotation task definition

Recording

Data preparation

Annotation

- Label 2D-3D boxes, polygons, etc
- Label Object attributes
- Label scene attributes
- Label actions, events
- Label relations between elements

Annotation validation

Consumption

- Rule-based inferencing
- Deep Understanding

Metadata repositories

Applications

- Querying
- Training
- Testing
OpenX: Inter-relations
Timeline for OpenLABEL

- **Dec 2020**: 03.03 Proposal Workshop
- **Jan**: Labeling Ideation
- **Feb**: Input from Ontology Project
  - Align the OpenLABEL project and the ontology project
- **Mar**: TSC
  - Project will be presented for approval
- **Apr**: ASAM OpenLABEL Concept
- **May**: WP1.: Annotation Format
- **Jun**: WP2.: Labeling Methods
- **Jul**: WP3.: Taxonomy and Structure
- **Aug**: New standard development
  - Use OpenLABEL concept paper for alignment and kickoff
- **Sep**: ASAM OpenLABEL new Standard development
- **Oct**: Release Concept Paper
- **Nov**: OpenLABEL
- **Dec**:
Questions?

Nicola Croce
Technical Program Manager
nico@deepen.ai