sensory™

Sensory Automotive

Emergency Vehicle Detection

Copyright 2024 Sensory, Inc. - Confidential & Proprietary

Meet the Speakers

Andi Hagen
Director of Machine Learning
Sensory, Inc.

Jeff Rogers
VP Sales & Marketing
Sensory, Inc.

Sensory Quick Overview

1) Technology development, Licensing, and Productization

- > 30 years focused on neutral nets and AI for voice
- Mostly PhDs, linguists and technologists

2) Broad suite of in vehicle technologies

- Technologies required for a complete automotive assistant

3) Market proven with deployment expertise

- Many hundreds of customers (see chart on the right)
- >3B units shipped

4) Tools for devt. & deployment

- Complete automotive SDK
- VoiceHub
- Custom Auto OEM Language Models
- On Device LLMs and synthetic data for model building

Copyright 2024 Sensory, Inc. - Confidential & Proprietary
Not for Distribution

Sensory Automotive Solution

Always on, always available
- No need to rely on network connections

Wake Word and Wake words
- Single or multiple wake words
- Hot words – these are software driven events

Speech to Text
- Full speech to text for controls, dictation and more
- Multiple model size options from 30MB to 220MB

Voice Biometrics
- Combined with wake word or STT

Automotive Language models, Micro Language models and Small Language models
- All things you would say to a car
- Micro language model to identify intents & entities
- Starting with an LLM, compress down and run in vehicle with a SLM

Emergency Vehicle Detection
- Designed to run using internal microphones

Copyright 2024 Sensory, Inc. - Confidential & Proprietary
Not for Distribution

EVD Systems

We need two ingredients: **Technology and Data**

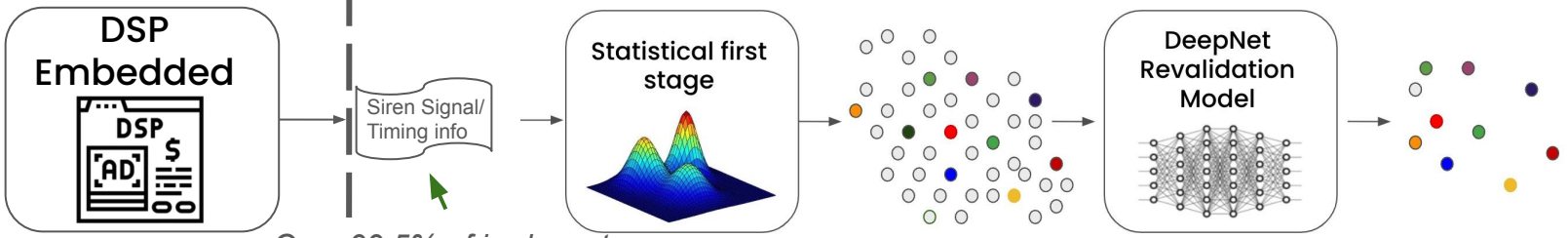
Variety of data is key!

- Markets
- Situations
- Noises
- Distance

Optional DSP + AP Level Solution

Optional:
DSP - Embedded

CPU/Headunit - AP Level



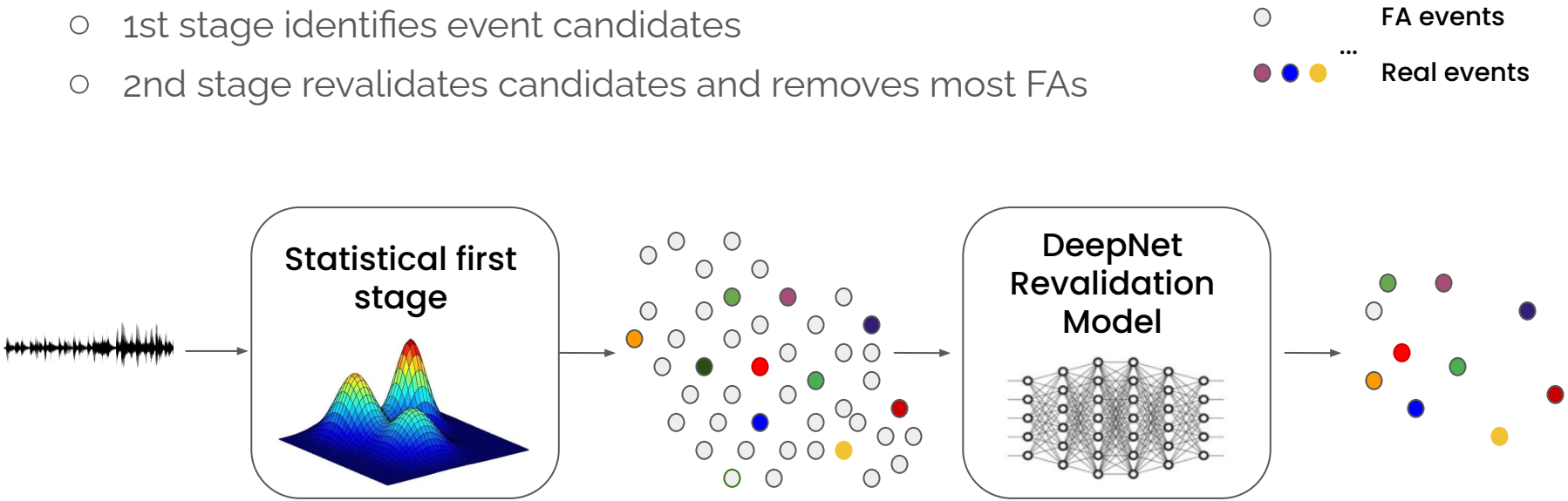
Over 99.5% of irrelevant audio is filtered out. Only potential siren signals are sent to the AP-level application

DSP solution runs a 80kB model - very efficient with 8 MIPS

DSP stage is optional - it reduces processing load on the Head Unit

System Architecture Overview

- **AP Level Two-stage System for Emergency Vehicle Sound Detection**
  - 1st stage identifies event candidates
  - 2nd stage revalidates candidates and removes most FAs



Detection of Sounds - FRR and FAR

How to measure Accuracy in real world conditions?

- **False Reject Rate (FRR)** = *# of missed events / # of events*

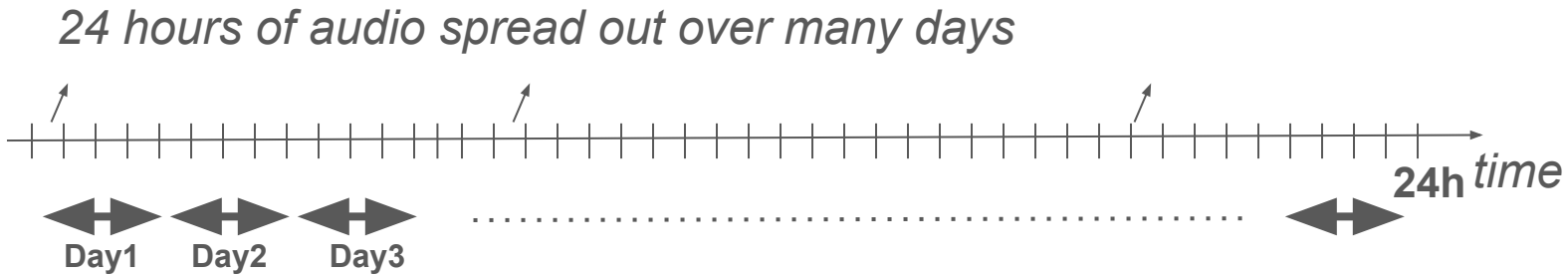
  For example: 8 out of 200 actual event cases were missed (4% FRR)

- **False Alarm Rate (FAR)** = *# of wrongly reported events / observation time window*

  For example: 1 false alarm event in 24 hours

Evaluation Methodology

- Measure False Rejects and Number of False Alarms in 24 hours
  - **Cars are used only about 1 - 2 hours per day**
  - Therefore 24 hours of driving time covers **more than a week in real-life situations**
  - We chose 1 FA in 24 hours of driving time as our operating point



Detection Accuracy

The table below shows the performance in **percent** of false rejects at **one false alarm in 24 hours of driving - more than a week in reality -** at different background noise levels on a realistic internal siren test set

False Reject Rates

Noise Level	20 dBA	10 dBA	5 dBA	0 dBA	Mean
1st stage (%)	1.3	0.9	0.3	0.5	0.8
2nd stage (%)	1.9	1.4	1.2	1.9	1.6

Sensory's EVD Models

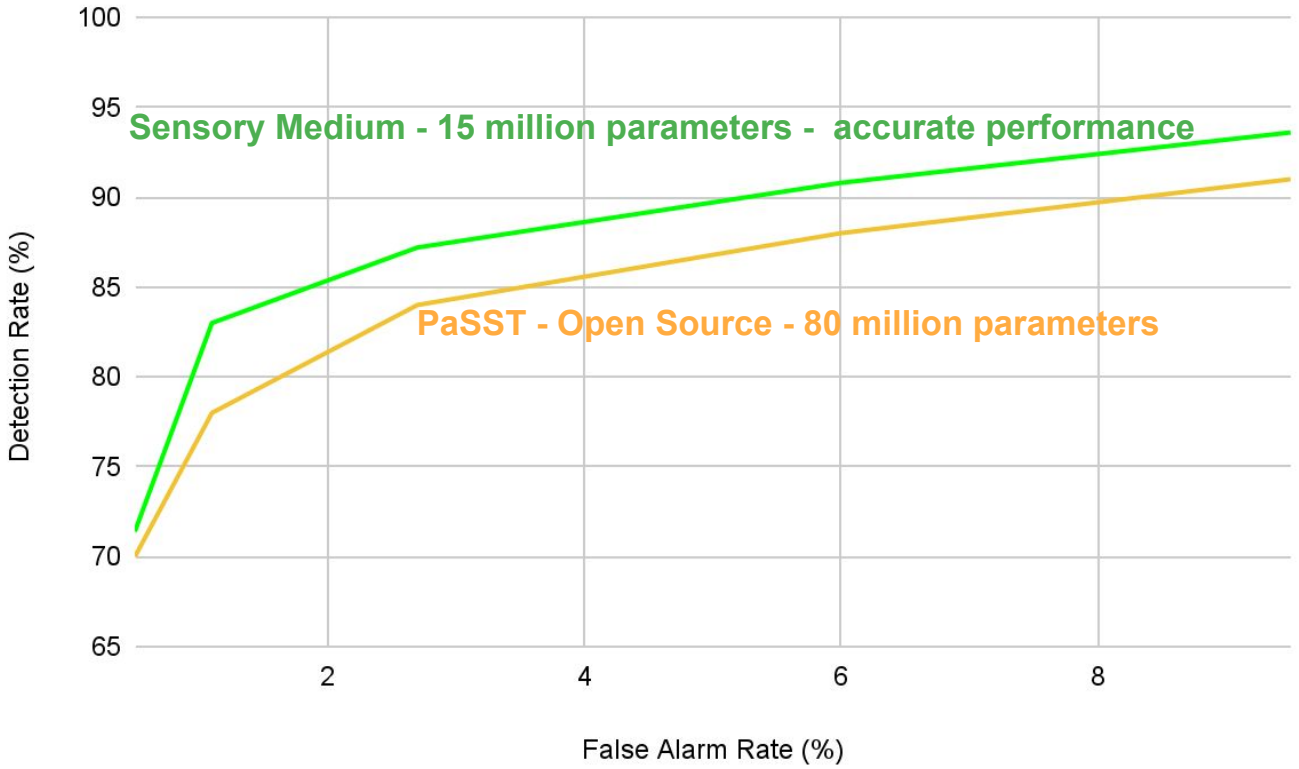
Sensory has two EVD models available at different sizes

Small	**1 million parameters**	**1.4 MB**
Medium	**15 million parameters**	**17 MB**

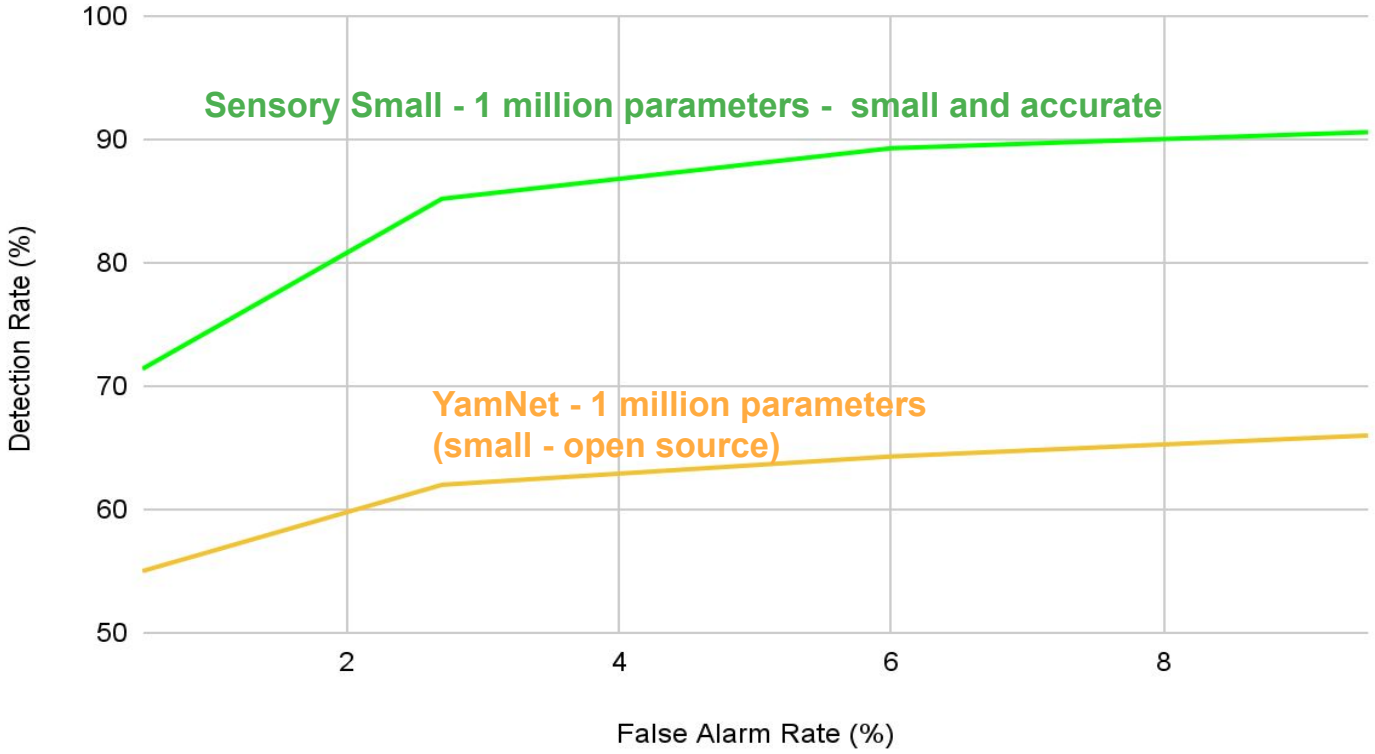
Open Source Models

YamNet	**1 million parameters**
PaSST	**80 million parameters**

Accuracy - Sensory Medium



Accuracy - Sensory Small

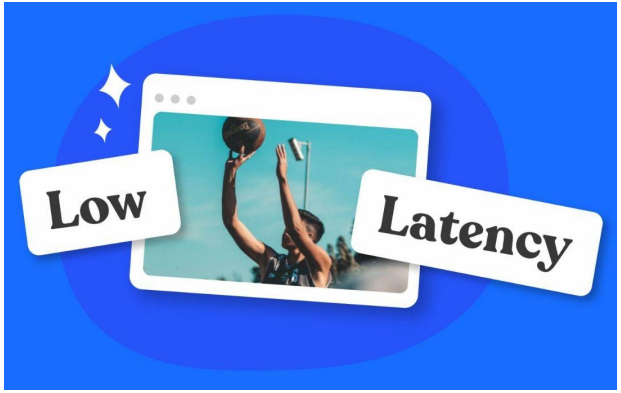


Road to Production

Key points for a production system for reliable EVD?

- Needs to be small and efficient
- Needs to provide precise timing
- Needs to be accurate under special conditions
- Needs to be deployable on diverse set of hardware/chips
- Needs to be well tested

Sensory SDK - Low Latency

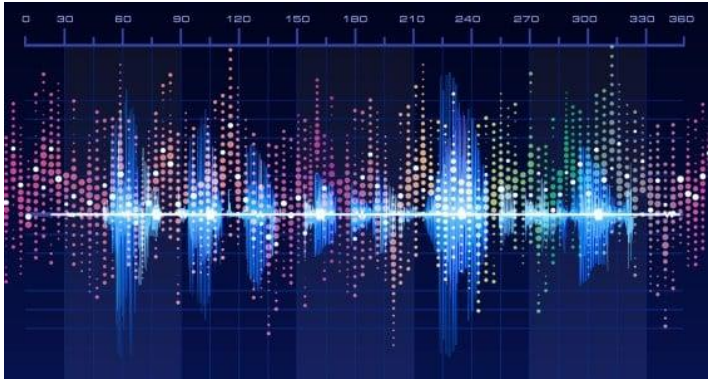


- Sensory EVD is fast
  - Overall response time within a few hundred milliseconds for first and second stage combined
  - Quick response time is essential
- System is optimized for footprint and latency
  - Technology maintains a sliding acoustic history of 1.5 seconds
  - Partially recognized sound events can trigger

Sensory SDK - Detection Characteristics

- We report detected events and provide
  - Class information (i.e. Siren)
  - The begin & end time
  - An event score

- Given a very long event, multiple event triggers may happen over the course of a single emergency vehicle passing by
  - SDK generally will report events every 500-1000 msecs



Sensory SDK - DSP and Embedded

- Supported DSP Platforms
  - ARM Cortex M-Series, ESP32, etc.
  - Optimized version for the HIFI5 DSP
  - Potentially portable to many DSPs
- Supported Embedded SDK Hardware Platforms
  - Linux: x86_64, Arm, Arm64
  - Windows: x86_64, MacOS: x86_64, Arm64
  - Android, iOS
- Supported Programming Languages
  - Java, C++, Python, Objective-C, Swift, C#

Low Memory and Power Consumption

- DSP solution runs at as low as **5 MIPS** (depending on dsp model size)
  - 10 MIPS for High Accuracy DSP model
  - 5 MIPS for small model

- AP level solution runs at about **30 MIPS**

Alarm-Sound-Trigger	1st stage model	1st and 2nd stage model
Model size	~88 kB	~1.4MB
MIPS / Memory	27 MIPS / 1.1 MB	33 MIPS / 4.6 MB

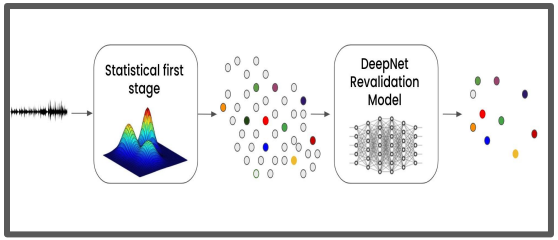
Software Dependencies

- EVD/SoundID is part of our Embedded SDK
  - Embedded SDK provided by Sensory
  - LiteRT (and LiteRT Micro) as inference engine
  - Using latest Neural Network runtime solutions for best performance
  - Tuned for small footprint and efficiency

New Improved EVD Tech Since Q1 2025

Scalable technology and valuable data lead to improved accuracy!

New models and inference engine



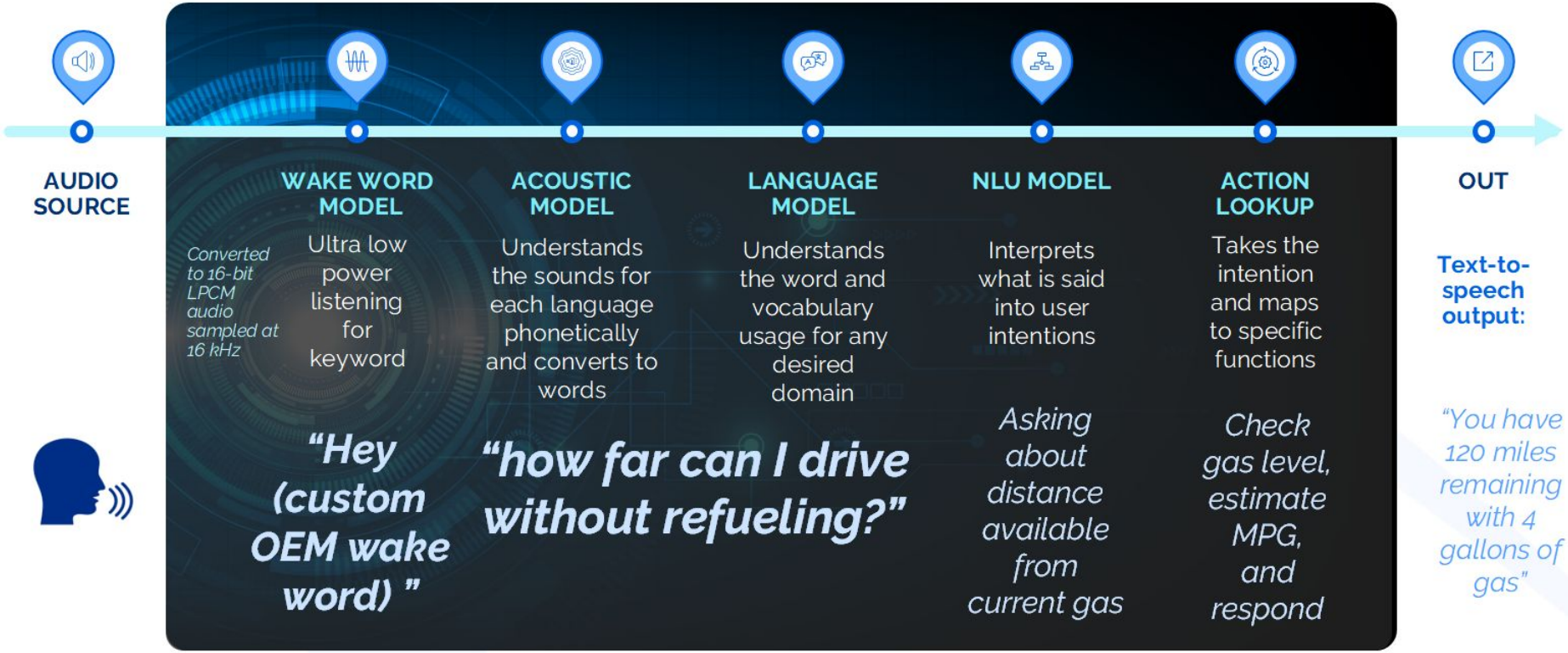
Improved accuracy

More data from real-life scenarios

sensory™

Q&A

Example: Automotive On Device Assistant



Copyright 2024 Sensory, Inc. - Confidential & Proprietary
Not for Distribution