



Introduction and History of Connected and Automated Vehicles

Dr. Gokulnath

CART, IIT Delhi

- Why is the automotive industry changing now?
- What exactly are “connected” and “automated” vehicles?
- How did we get here historically?
- And what barriers remain?

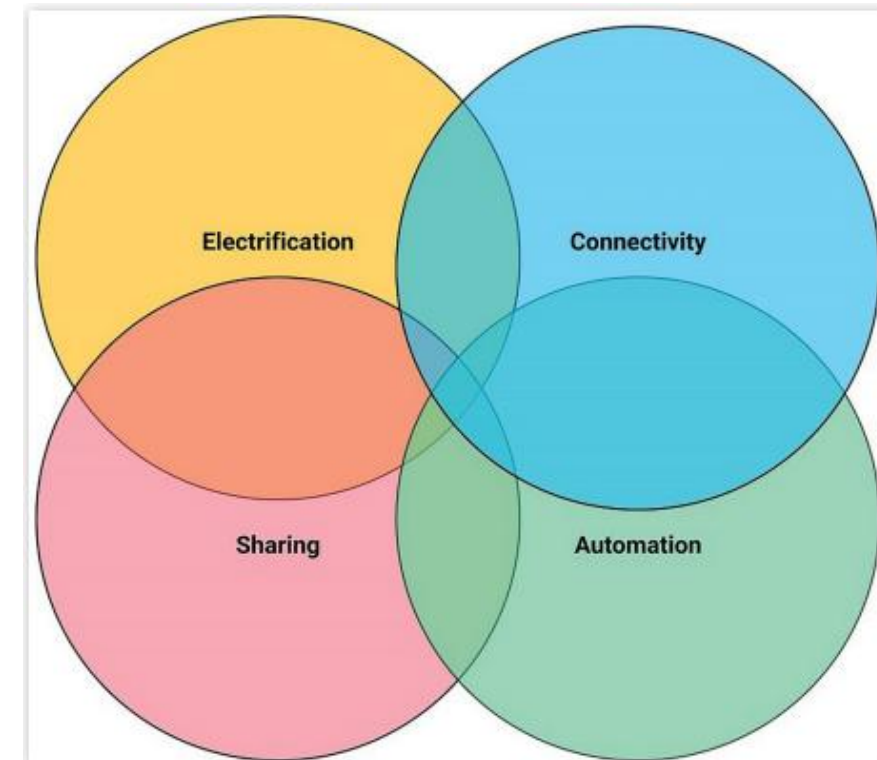


- Driver-centered paradigm stayed stable for ~100 years
 - Only incremental improvements
 - Human driver controls steering + pedals
 - Minimal communication (turn signals)
- Automotive industry
 - Improves safety, efficiency, onboard computing

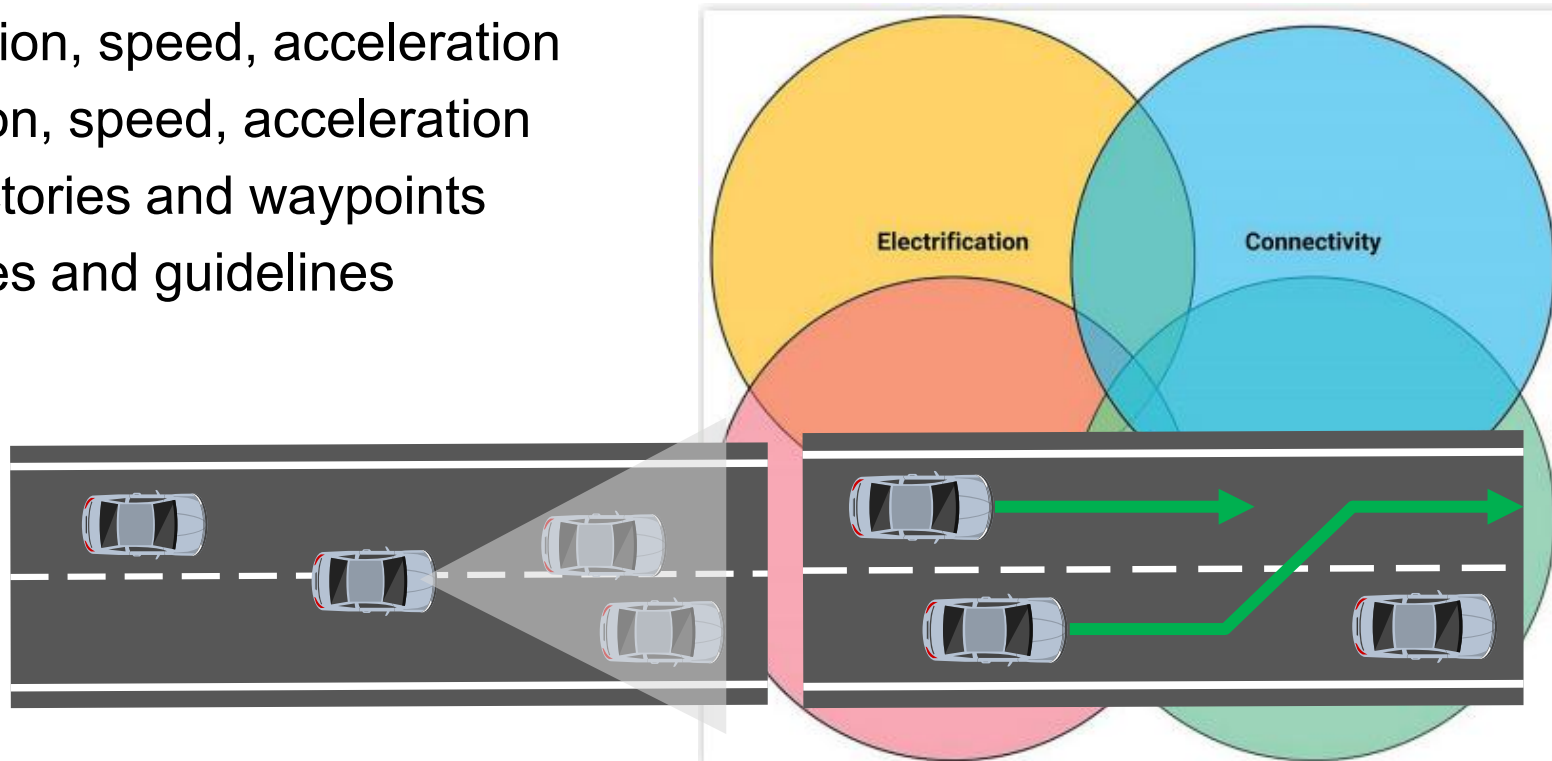
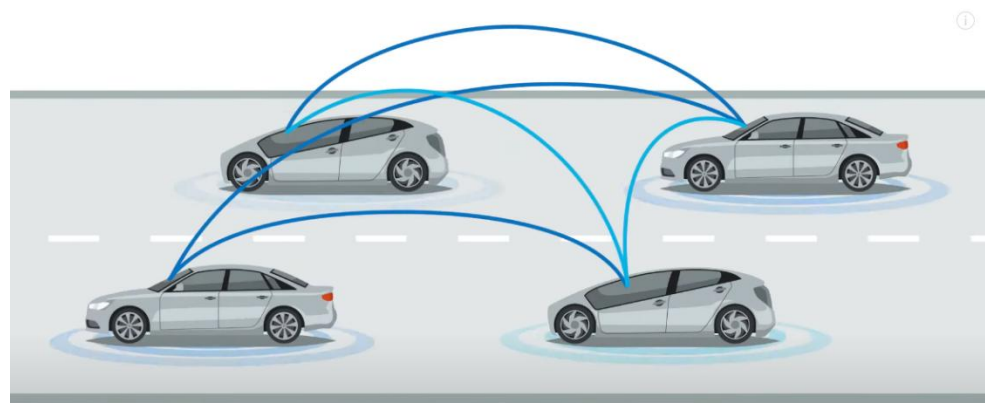


- Electrification

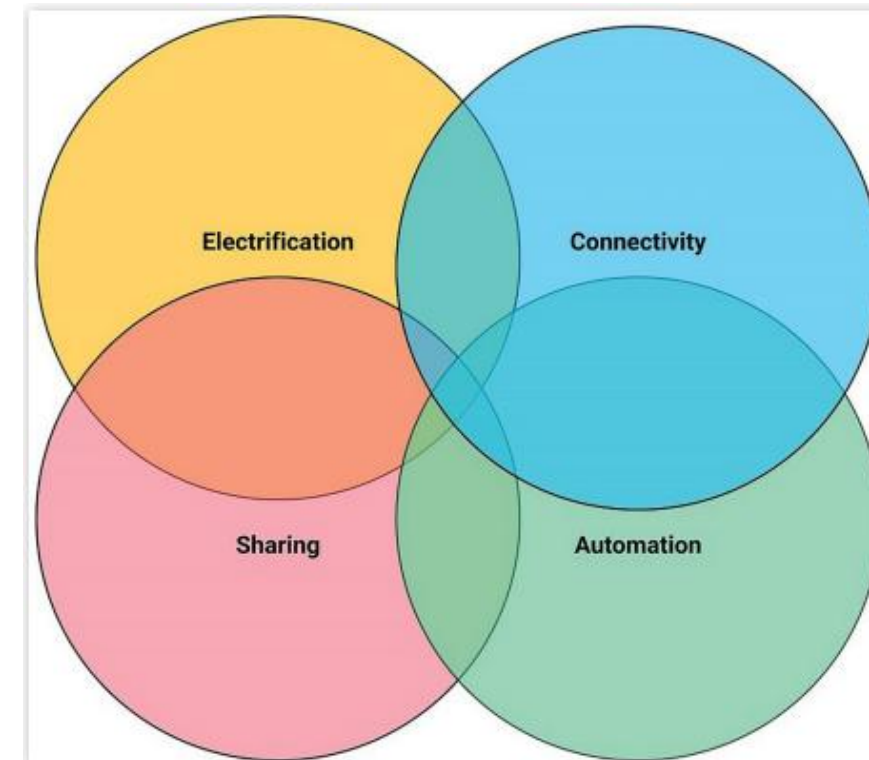
- Electrification is the most mature of the four trends
- Powertrain shift from Internal Combustion Engines (ICE) → hybrid/EV (electric motors)
- R&D work in vehicle electrification is only two decades old
- Norway, the percentage was 54% in 2020
- More efficient at converting energy to vehicle motion
- Fewer moving, mechanical parts → lower maintenance
- Drastically reduce emissions



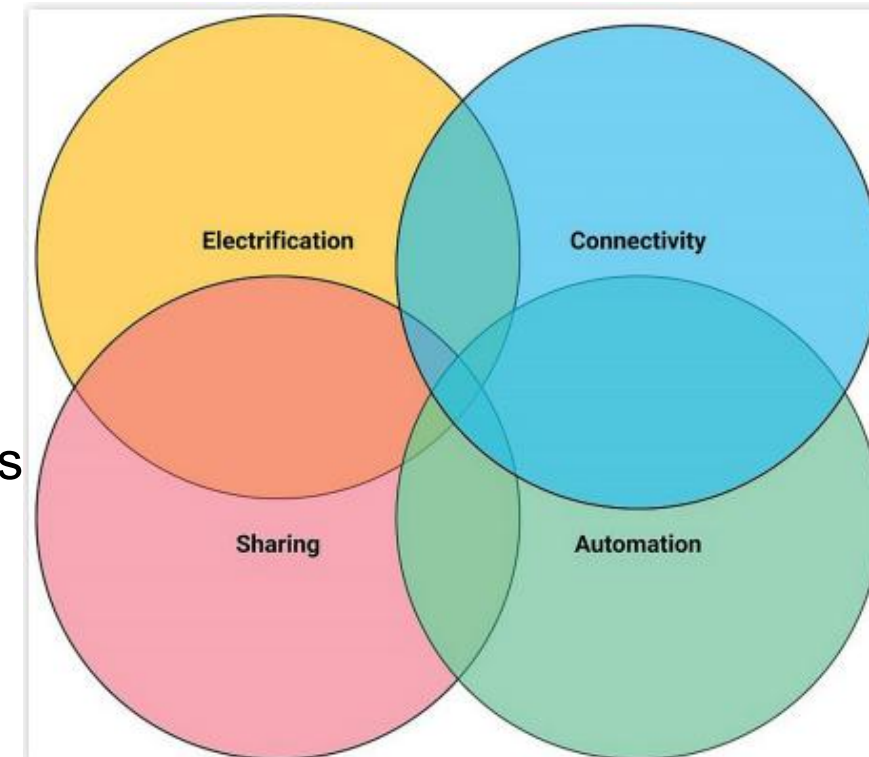
- Connectivity
 - Connectivity is the next mature trend
 - GPS-equipped vehicles were introduced by Oldsmobile Ciera 1994
 - Early vehicles can receive Signal Phase and Timing (SPaT) information from traffic signals
 - Provide information on the traffic light sequence and timing
- Possible information exchanged between vehicles
 - Own dynamic properties: position, speed, acceleration
 - Sensor detected object: position, speed, acceleration
 - Driving intentions: future trajectories and waypoints
 - Traffic management procedures and guidelines



- Sharing
 - Sharing is the second-least mature trend
 - Sharing includes ride hailing (Uber) and car sharing (AutoShare)
 - Sharing is also known as multi-modal transportation
 - From scooters to bicycles to cars to buses and trains
 - Sharing provide enhanced mobility for individuals including
 - disabled persons, seniors, children, etc.
 - Sharing allows for lower levels of vehicle ownership



- Automation
 - Automation = vehicles with **Automated Driving Systems (ADS)**
 - **Automation Spectrum (SAE J3016):**
 - Low Level: Driver Assistance (e.g., Adaptive Cruise Control)
 - High Level: Full Automation (No human supervision required)
 - Automation helps humans to focus on other tasks
 - Automation reduce accidents by removing human error
 - Automation helps for those who cannot drive themselves
 - Automation is the **biggest change** in the vehicle industry
 - Since the invention of the automobile itself
 - Automation is the "last in maturity" compared to other trends
 - Holds the highest potential for total sector disruption



- Airplane autopilot enabled flying + navigation simultaneously
 - Example: Sperry Gyroscope Autopilot (1930s)
 - Automatically maintain a desired compass heading and altitude
- Early torpedoes: maintain course + depth (1860s)
 - Later: added sonar targeting by WWII
- German V2 rocket: gyroscope-based guidance
 - Early human-made object into outer space



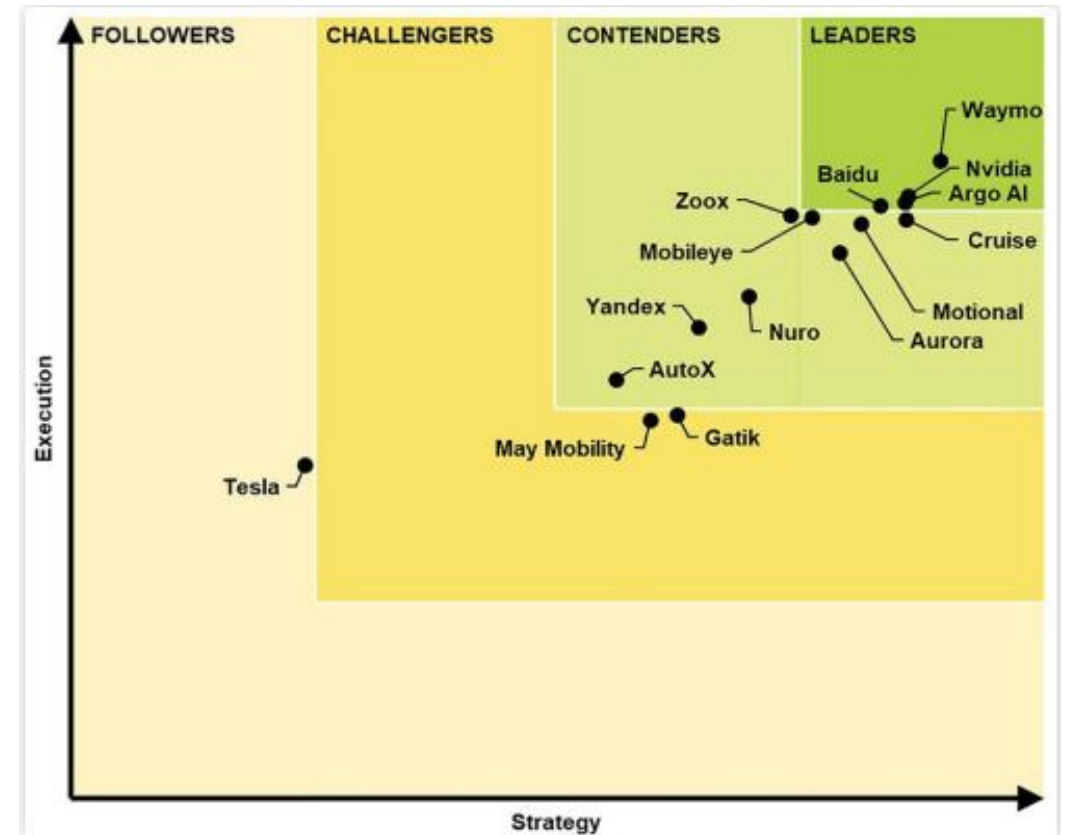
- Cockroach-like motion: Sensing → Processing → Reacting
 - Sensing and reacting were possible
 - Hardest part historically: **processing (machine intelligence)**
- 1980s–1990s: High-Speed Autonomy
 - The Mercedes Van (1980): Travel on highways using a primitive automated driving version
 - VaMoRs Van (1997): Third generation of advanced vision systems for autonomous navigation



- DARPA Challenge (2004–2007)
 - First long distance competition for driverless cars in the world
 - Stanley (2005) and Boss (2007) winners
 - Used machine learning and sensor fusion (LiDAR/Radar) for automation
 - Modern AV industry born
 - Creating companies like Waymo, Tesla, and Uber



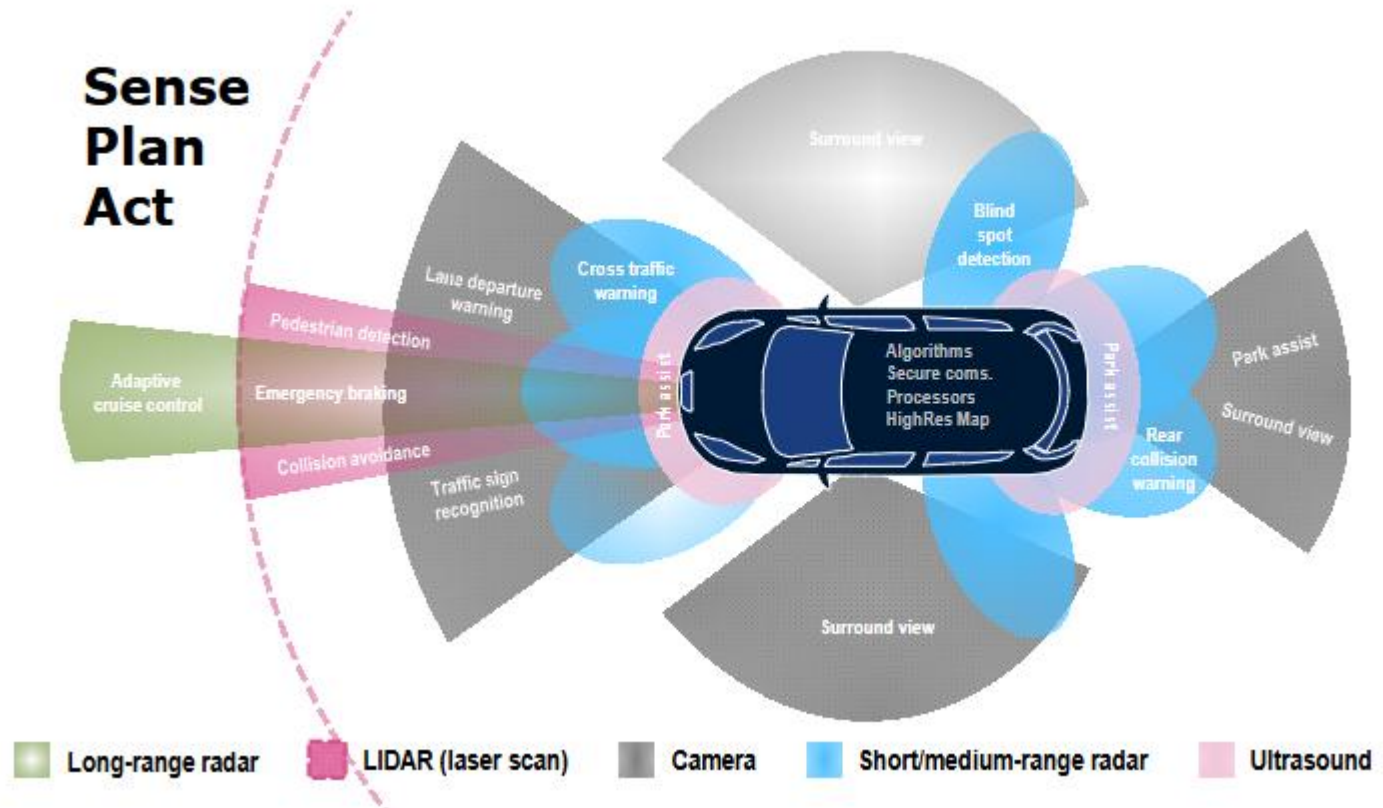
- CAV industry rankings criteria
 - Vision, Go-to market strategy, Partners, Production strategy, Technology, marketing, and distribution, Product capability, Product quality and reliability, Product portfolio, Staying power
- CAV industry leaders
 - Waymo, Nvidia, Argo AI, and Baidu
 - Tesla low in both strategy and execution
 - Public perception may differ



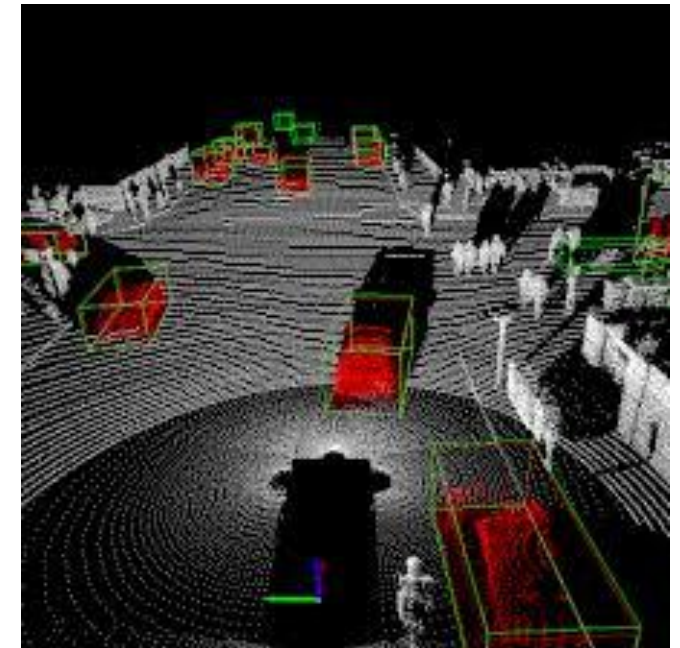
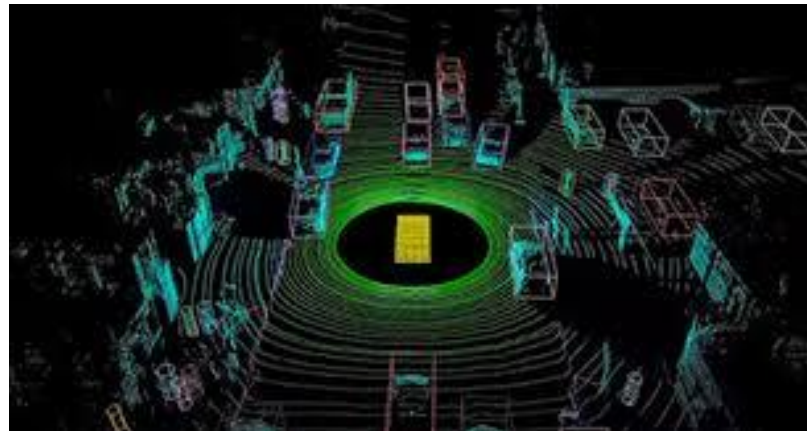
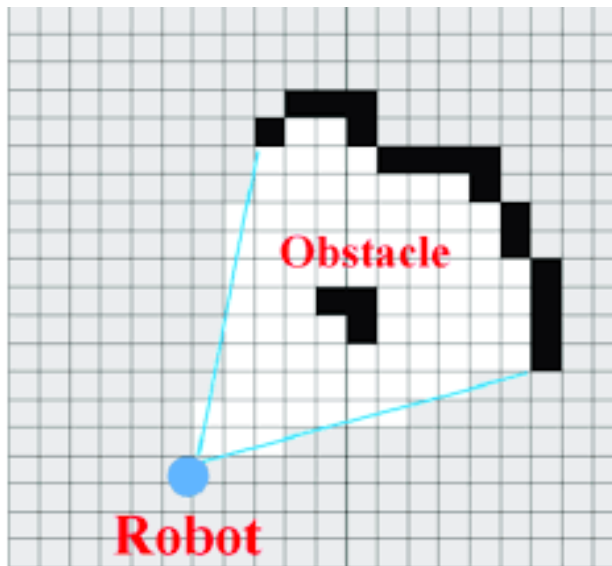
- Strong excitement from industry/government/public in 2020
 - But widespread deployment is not near-term
- **The 80/20 Rule:** The industry has mastered roughly 80–90% of driving tasks,
 - But the final 10–20% (corner cases) is proving exponentially more difficult
 - Perception must handle difficult scenarios:
 - Weather, temporary obstacles/restrictions, parking lots, heavy pedestrian/cyclist traffic, non-mapped areas
- **Timeline Debate:** Opinions are split
 - Some see mass commercialization by 2030, while others predict it is still decades away



- A CAV works
 - Perception: Understanding the world
 - Planning: Deciding what to do
 - Actuation/Control: Executing safely



- Perception must produce a machine-readable scene of the environment
 - Ego-state: position, velocity, yaw rate, acceleration (vehicle's own motion)
 - Objects: detection + classification (car, truck, pedestrian, cyclist...)
 - Tracking: position/velocity/direction over time of objects
 - Free space: drivable area, boundaries, curbs, road edges
 - Traffic controls: signals, signs, lane markings, right-of-way cues
- **Outputs:** occupancy grid, tracked objects list



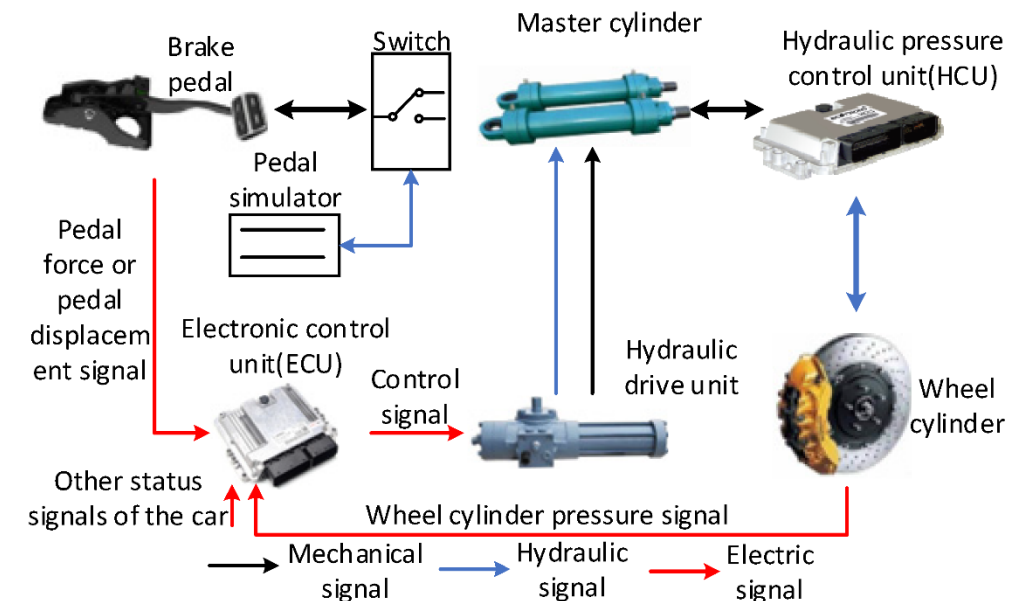
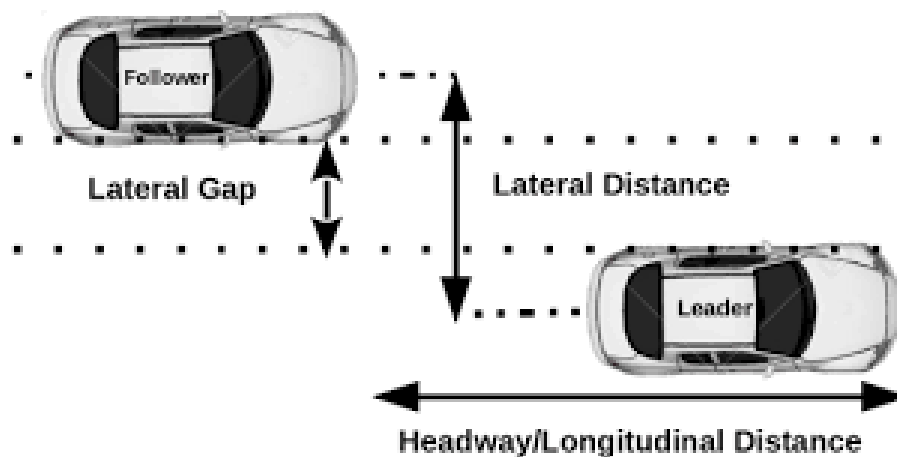
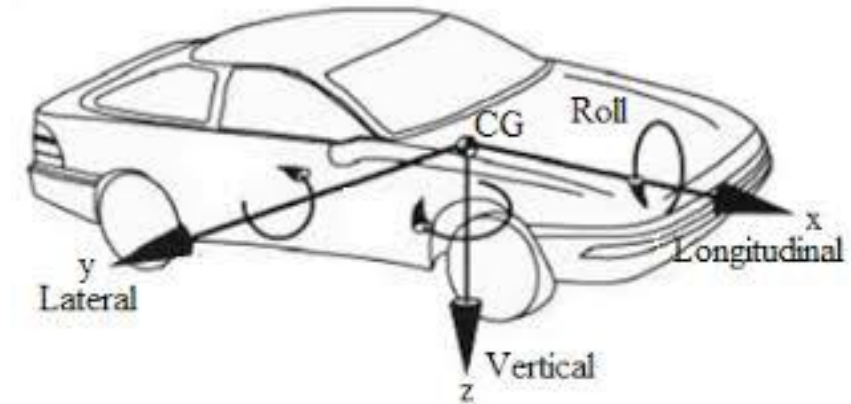
- Sensors (typical):
 - **Cameras:** rich semantics (lanes, signs), weak in glare/rain/night
 - **Radar:** strong range/velocity, weaker shape/class semantics
 - **LiDAR:** accurate geometry, cost/packaging concerns
 - **GNSS/IMU/Wheel odometry:** ego motion + localization backbone
- Fusion levels:
 - **Raw-data fusion** (hard, heavy compute)
 - **Feature-level fusion**
 - **Object-level fusion** (common, robust)



- Planning converts world model + mission into actions:
 - **Mission / Route planning:** where to go
 - **Behavior planning:** what maneuver (keep lane, yield, overtake, stop)
 - **Motion planning:** exact trajectory (path + speed profile) within constraints
- **Output:** a trajectory: $x(t), y(t), v(t)$ over next few seconds
- **Prediction**
 - Forecast trajectories of vehicles/pedestrians/cyclists
 - Estimate maneuver (will cut-in? will cross? will stop?)
- **Prediction Approaches:**
 - Physics-based (constant velocity/acceleration)
 - Rule-based (yielding, lane following)
 - Learning-based (data-driven intent + interaction)



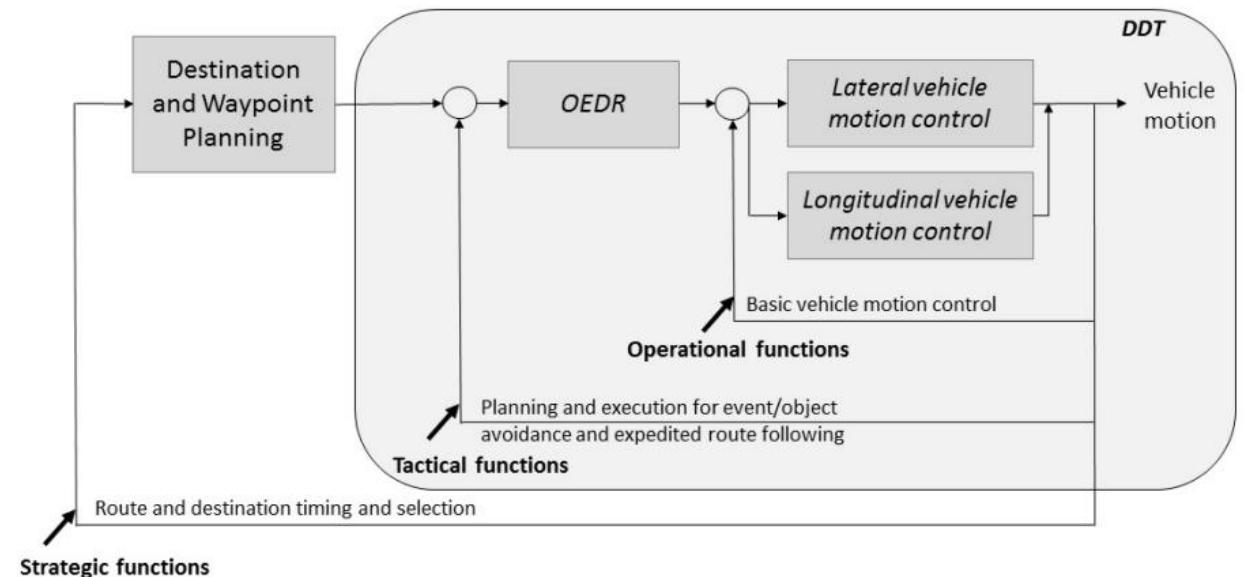
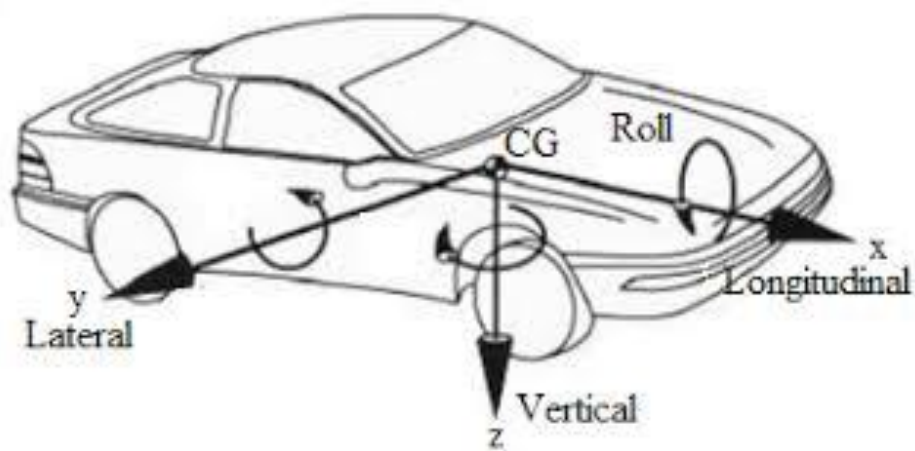
- Control turns planned trajectory into commands:
 - **Lateral control:** steering to follow path
 - **Longitudinal control:** throttle/brake to follow speed profile
- Vehicle actuators:
 - Steering actuator, brake-by-wire, throttle, gear, etc.
 - Safety monitoring: actuator faults, degraded modes
- **Output:** steering angle + throttle/brake commands at high rate



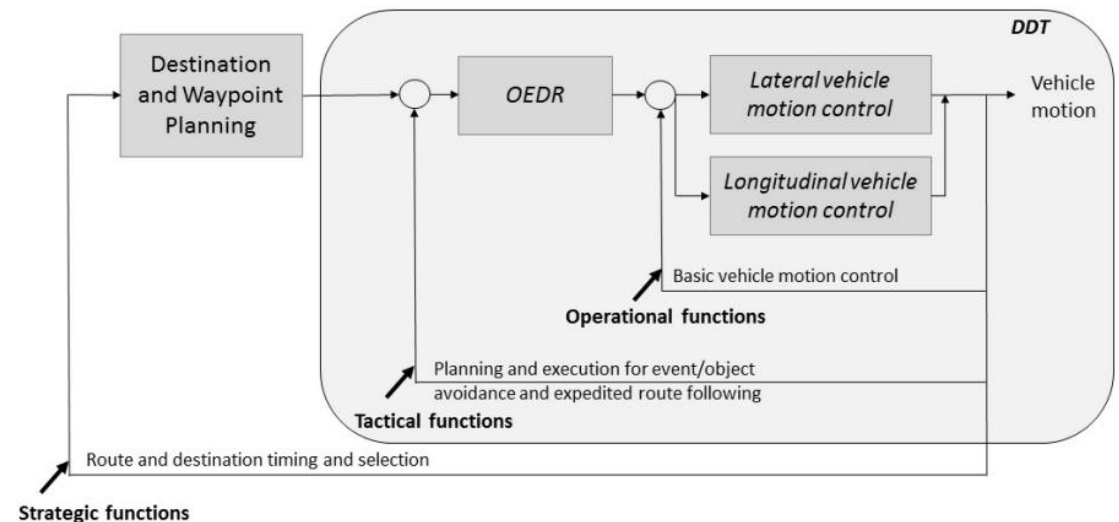
- What is SAE J3016?
 - Defines a **taxonomy** with **six levels of driving automation (0–5)** for on-road vehicles
 - Levels apply to **features**, not necessarily the whole vehicle
 - Car might have a Level 2 highway feature and a Level 4 parking feature
- Three actors model involved
 - **The (Human) User**: Can be a driver, passenger, or remote assistant.
 - **The Driving Automation System**: The hardware/software performing the task.
 - **Other Vehicle Systems**: Brakes, steering, lights (the "conventional" parts of the car)
- Active Safety is **not** driving automation because it is momentary, not sustained
 - Anti-lock Braking System (**ABS**), Electronic Stability Control (**ESC**), Automatic Emergency Braking (**AEB**)



- Dynamic Driving Task (**DDT**)
 - System performs **real-time operational + tactical** driving functions, excluding **strategic** trip planning
- **Strategic**: Trip planning (Destination, Waypoints)
 - This is outside the scope of DDT
- **Tactical**: Maneuver planning (changing lanes, signaling)
- **Operational**: Split-second micro-adjustments (steering, braking)



- **ODD (Operational Design Domain)**
 - Set of operating conditions under which the driving automation feature is designed to function
 - e.g., road type, speed range, geography, weather/visibility, traffic conditions, time-of-day
- **OEDR (Object and Event Detection and Response)**
 - Monitoring the driving environment (detect/recognize/classify objects and events)
 - Responding appropriately to complete the DDT and/or support fallback
- **Key:**
 - ODD = What conditions is this feature supposed to operate
 - OEDR = What it must perceive and how it responds within those conditions



- DDT fallback triggered by either
 - **DDT performance-relevant system failure**
 - Malfunction that prevents the automation system to perform DDT
 - Eg: sensor outage (camera/radar), steering/brake actuator fault, compute failure
 - **ODD (Operational Design Domain) exit**
 - Vehicle/feature is no longer operating within ODD
 - Eg: road type changes (highway → city), speed out of range, heavy fog/rain, construction zone
- DDT fallback outcome
 - **Driver takeover:** Human performs the DDT (lateral + longitudinal control)
 - **Minimal Risk Maneuver:** Automated Driving System (ADS) move to a safe minimal-risk state
 - Move to a safe location and stop
 - Reduces crash risk



- **Definition:** Driver performs the entire **Dynamic Driving Task (DDT)** at all times
- **Key points:**
 - Warnings / momentary interventions do **not** count as automation (not sustained DDT control)
 - Driver is always responsible for steering, speed control, and environment monitoring
- **Examples:** Anti-lock Braking System (ABS), Electronic Stability Control (ESC), Automatic Emergency Braking (AEB), Forward Collision Warning (FCW), Lane Departure Warning (LDW)
- **Roles at Level 0:**
 - Dynamic Driving Task (**DDT**): Driver
 - Object and Event Detection and Response (**OEDR**): Driver
 - **DDT** fallback: Driver
 - Operational Design Domain (**ODD**): Not applicable (no driving automation feature)



- **Definition:** System performs sustained **either lateral OR longitudinal** motion control (not both)
- **Key points:**
 - Driver performs OEDR (continuous supervision and environment monitoring)
 - Driver controls the other axis not controlled by the system
- **Examples:**
 - Adaptive Cruise Control (ACC): controls speed + gap (longitudinal only); driver steers
 - Cruise Control (CC): controls speed (longitudinal only); driver steers
 - Lane Keeping Assist (LKA): controls steering (lateral only); driver controls speed/brake
- **Roles at Level 1:**
 - Dynamic Driving Task (DDT): Driver + System (system controls one axis)
 - Object and Event Detection and Response (OEDR): Driver
 - DDT fallback: Driver
 - Operational Design Domain (ODD): Feature-specific



- **Definition:** System performs sustained **both** lateral **and** longitudinal motion control simultaneous.
- **Key points:**
 - Driver performs OEDR and must supervise continuously.
 - Driver must be ready to intervene immediately at any time.
- **Examples:**
 - Highway Assist / Pilot Assist: controls steering + speed at the same time; driver still monitors
 - Traffic Jam Assist: low-speed stop-and-go and lane centering simultaneously; driver supervises.
 - Traffic-Aware Cruise Control: steering + speed control; driver performs OEDR and fallback.
- **Roles at Level 2:**
 - Dynamic Driving Task (DDT): Driver + System (system controls lateral + longitudinal)
 - Object and Event Detection and Response (OEDR): Driver
 - DDT fallback: Driver
 - Operational Design Domain (ODD): Feature-specific



- **Definition:**
 - **Automated Driving System (ADS)** performs the entire **DDT including OEDR** within its **ODD**; human is **fallback-ready** and must respond to a **request to intervene**
- **Key points:**
 - User is not required to supervise continuously while ADS is engaged
 - When ADS requests takeover, the user performs fallback (takes over DDT)
- **Examples:** Traffic jam ADS within defined conditions (limited-access roads, speed range, etc.)
- **Roles at Level 3:**
 - Dynamic Driving Task (DDT): ADS
 - Object and Event Detection and Response (OEDR): ADS
 - DDT fallback: Fallback-ready user (upon request)
 - Operational Design Domain (ODD): Required and must be stated



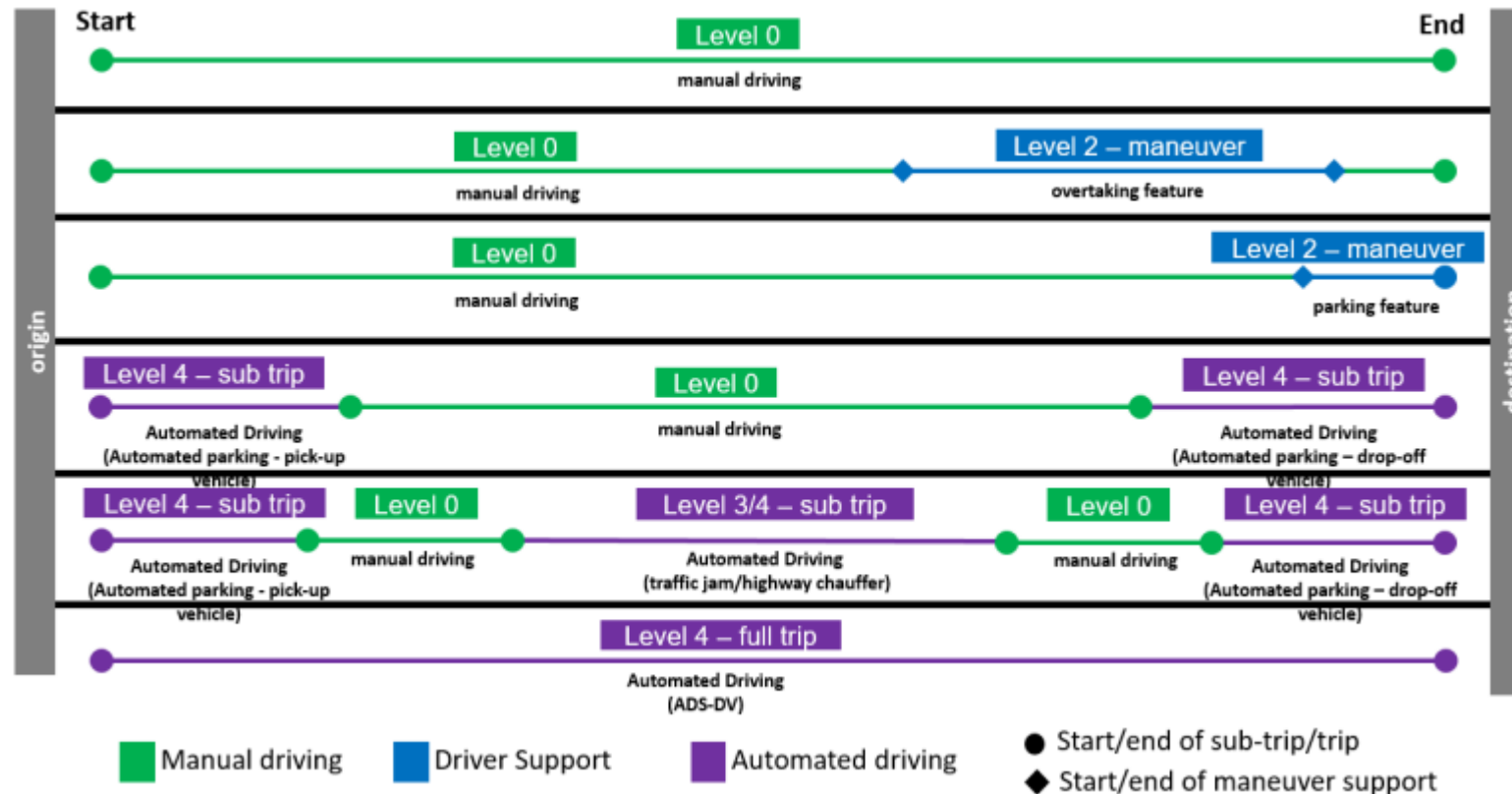
- **Definition:** ADS performs **DDT + OEDR + DDT fallback** within its **ODD** (no expectation of human takeover).
- **Key points:**
 - If a failure occurs or ODD is reached/exceeded, ADS handles fallback (e.g., achieves minimal risk condition).
 - Operation is limited to its ODD (geo-fenced area, specific roads, conditions, etc.).
- **Examples:**
 - Geo-fenced robotaxi, automated shuttle on fixed routes, driverless valet parking in controlled areas
- **Roles at Level 4:**
 - Dynamic Driving Task (DDT): ADS
 - Object and Event Detection and Response (OEDR): ADS
 - DDT fallback: ADS
 - Operational Design Domain (ODD): Required and must be stated



- **Definition:** ADS performs **DDT + OEDR + fallback** under **all conditions a human driver could manage**.
- **Key points:**
 - No operational limitation by ODD (conceptually “anywhere a human can drive”).
 - No human driver needed at any time.
- **Examples:**
 - Concept level; no broadly deployed unrestricted Level 5 systems today
- **Roles at Level 5:**
 - Dynamic Driving Task (DDT): ADS
 - Object and Event Detection and Response (OEDR): ADS
 - DDT fallback: ADS
 - Operational Design Domain (ODD): Not limited (no defined constraints)



- Sub-trip feature:
 - Requires a human driver for at least part of every trip
- Full-trip feature:
 - Automated Driving System (ADS) features that can operate throughout complete trips



- Step 1: Sustained automation of the Dynamic Driving Task (DDT)?
 - Ask: Does the system perform any part of the DDT on a sustained basis (not just warnings or brief interventions)?
 - **If No → Level 0**
- Step 2: How many motion-control axes are automated (when engaged)?
 - One axis only (either lateral steering OR longitudinal accel/brake) → Level 1
 - **Both axes simultaneously (lateral + longitudinal together) → Level 2**
- Step 3: Who performs Object and Event Detection and Response (OEDR)?
 - If the driver must continuously monitor the environment and respond → Level 1 / Level 2
 - **If the system performs OEDR as part of driving → it is an Automated Driving System (ADS) → Level 3+**



- **Step 4: Who performs DDT fallback when needed (system failure or ODD exit)?**
 - If the feature expects a **human takeover** after a request to intervene → **Level 3**
 - If the **ADS performs fallback** and can reach a **minimal risk condition** without expecting human takeover → **Level 4 / Level 5**
- **Step 5: Is the ADS limited to a defined Operational Design Domain (ODD)?**
 - **Yes (ODD-limited: specific roads/area/speeds/weather conditions, etc.)** → **Level 4**
 - **No ODD limitation** (can drive under all conditions a human driver could manage) → **Level 5**
- **Important:**
 - Assign levels using **design intent + specified user role + stated ODD**, not “what it looked like” in a demo.

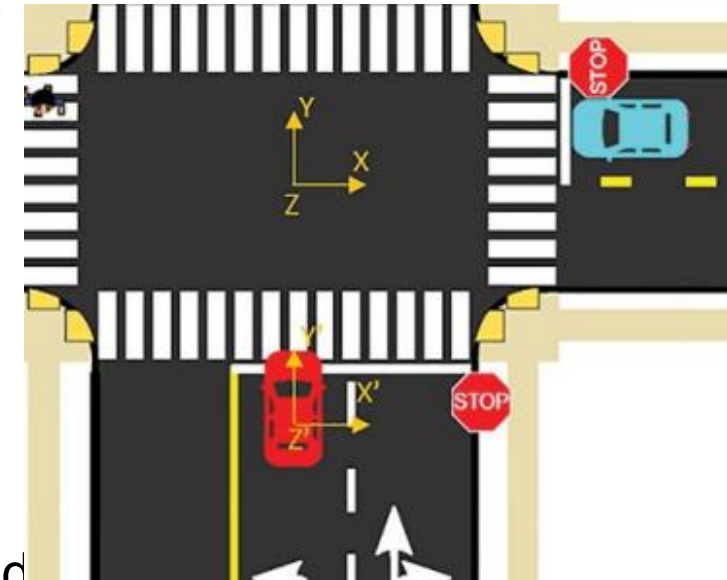
- ACC — Adaptive cruise control
- ADAS — Advanced driver assistance system
- ADS — Automated driving system
- ADS-DV — Automated driving system-dedicated vehicle
- AEB — Automatic emergency braking
- DDT — Dynamic driving task
- ESC — Electronic stability control
- LKA — Lane keeping assistance
- ODD — Operational design domain
- OEDR — Object and event detection and response



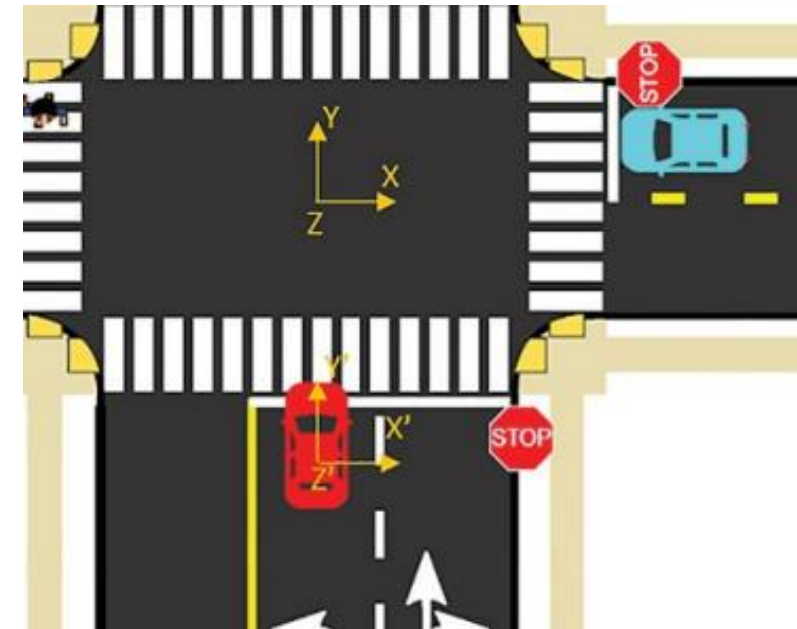
- **Definition:** Localization is the process by which a CAV determines its pose
 - 3D orientation, 3D position, speed, and acceleration
- **Core Question:** It answers the question?
 - Where am I exactly, and how am I moving relative to the known world?
- **Context:** It is the first step in the **Perception sub-system**.
 - Without knowing where it is, the vehicle cannot plan a path



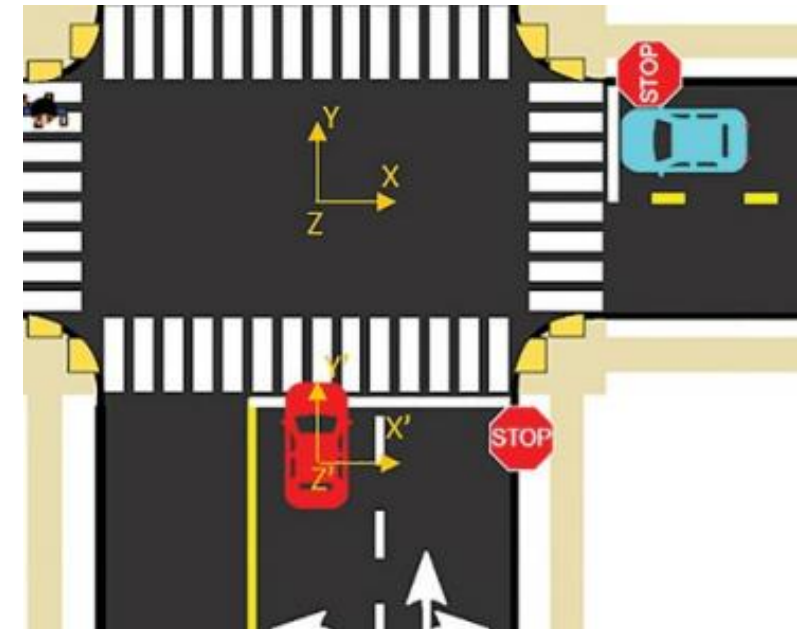
- **World Coordinate System** (X, Y, Z)
 - An arbitrary origin point (X, Y, Z) fixed on the map (e.g., the center of an intersection)
 - The choice of origin is arbitrary and determined by the developer
- **Local Coordinate System** (X', Y', Z')
 - The vehicle's own internal axes
 - Described relative to the World Coordinate System
- Localization algorithms calculate
 - Relationship between the vehicle's local frame and the fixed world frame
 - Vehicle's position (distance along axes) and orientation (Pitch, Roll, Yaw)



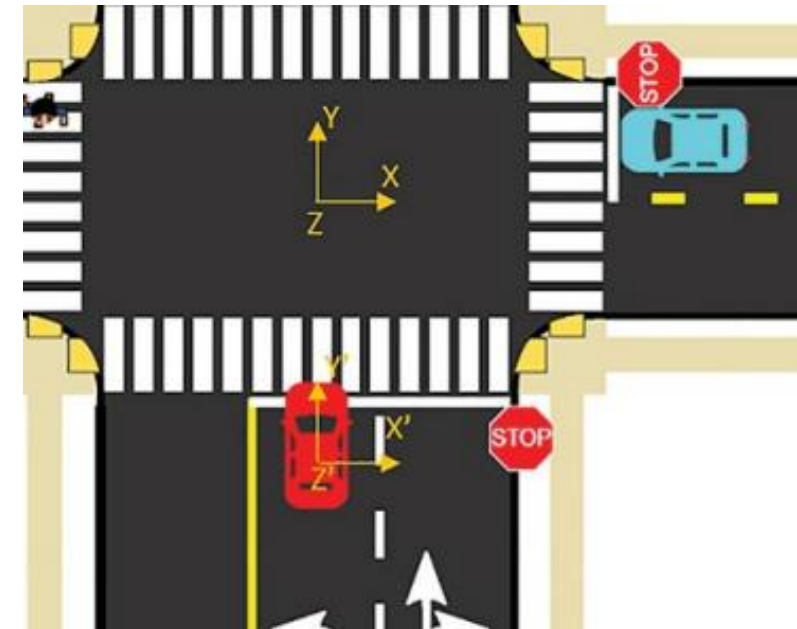
- **Sensors Used:** Cameras, Inertial Measurement Unit (IMU)
- Case Study: Vehicle Stopped Straight in Lane
 - Yaw (Rotation about Z'): 0° (Vehicle is straight)
 - Roll (Rotation about Y'): 0° (Assuming road is flat)
 - Pitch (Rotation about X'): 0° (Assuming road is flat)
 - **Note:** In this alignment, the pose is identical with respect to both local and world systems



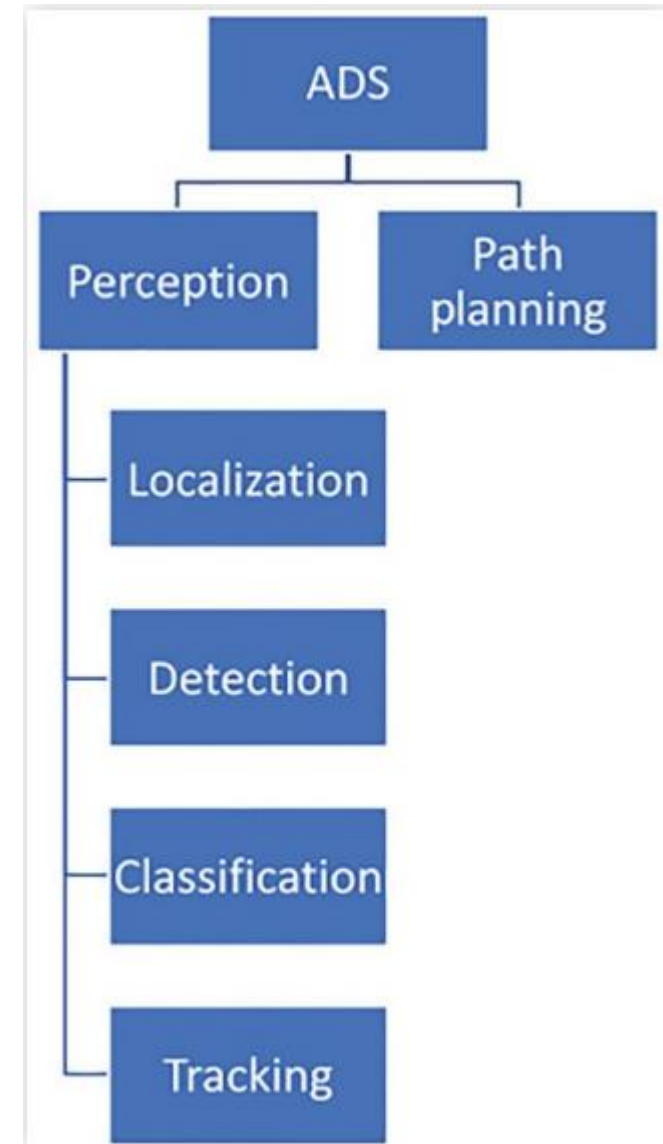
- **Sensors Used:** GPS Unit
- Case Study: Vehicle Stopped Straight in Lane
 - **X-Axis:** Distance is zero (assuming Y and Y' axes align)
 - **Z-Axis:** Distance is zero (assuming flat road/same elevation)
 - **Y-Axis:** Distance is a positive, nonzero value
 - Representing the distance from the intersection center to the vehicle



- **Sensors Used:** Wheel Speed Sensors
- Case Study: Vehicle Stopped Straight in Lane
 - Speed and acceleration are zero
- Case Study: Turning Left (Dynamic)
 - Pitch, Roll, and Yaw values change due to steering input and suspension movement



- Localization is considered a part of the perception system
 - Input to the path planning system
- **Redundancy:**
 - Localization acts as a perception redundancy check
 - It compares what the sensors "see" against
 - what the map says *should* be there
- **Ground Truth:**
 - A highly accurate map establishes "Ground Truth."
 - If sensors fail (blind spots)
 - The map provides safety-critical location data for static objects
- Localization consists of two primary steps:
 - **Mapping and Sensing**



- Mapping
 - Mapping is a fundamental pillar of localization.
- CAVs obtain environmental maps through three primary methods
 - **Connectivity**: Received from an outside source (V2X or Cloud)
 - **On-board Storage**: Pre-loaded and stored on the vehicle's computer
 - **Real-time Creation**: Generated dynamically by the vehicle's sensors as it moves
- Current Industry Standard
 - Most ADS developers rely on **a priori** mapping (pre-existing maps) within **geo-fenced** areas
 - In geo-fenced operations, the vehicle accesses the map through:
 - Connectivity, or On-board stored maps
 - **Real-time mapping alone** is generally considered **insufficient** by most ADS developers
 - **Tesla** is a major exception (discussed later in the chapter)



- Conventional (SD) Maps
 - SD map providers: **Google Maps, Apple Maps, Waze**, etc
 - SD maps are useful for **human navigation** but **insufficient** for CAV localization
 - Missing lane widths, curb heights, and specific signage data
 - SD map **positional accuracy** is often **off by meters**, which is unacceptable for CAVs
 - Study reported **Google Earth** accuracy ranging from **0.4 m to 171.6 m**
- 3D High-Definition (HD) Maps
 - Provide centimeter-level resolution/accuracy
 - Establishes the **"Ground Truth"** for the vehicle
 - If sensors fail or have blind spots, the HD map provides the location of static objects



- **Reduced Computational Load**
 - Highly detailed maps allow the ADS to focus processing power on **change detection**
 - System doesn't need to "re-discover" the road layout
 - It only needs to identify dynamic elements (pedestrians, other cars)
- **Static vs. Dynamic Split**
 - **Static Environment:** Provided by the HD map
 - **Dynamic Perception:** Handled by real-time sensors (LiDAR, Radar, Cameras)
- **Increase reliability:**
 - Better HD maps reduce the depend on real-time sensor capture during complex scenarios



To organize complex data, HD maps are structured into five distinct layers:

Base Map	SD map that includes road curvatures, elevation, and GPS coordinates
Geometric	3D point cloud created from the mapping sensors (LIDAR, camera, GPS, etc.) that includes the stationary objects in the surroundings
Semantic	2D or 3D semantic objects (such as lane markings, street signs, and other details)
Map Priors	Prior information and entity behavioral data, such as average cycle/pedestrian speeds, and SPaT information
Real-time	Real-time traffic information that a CAV can receive (V2X) to assist with trip planning and navigation

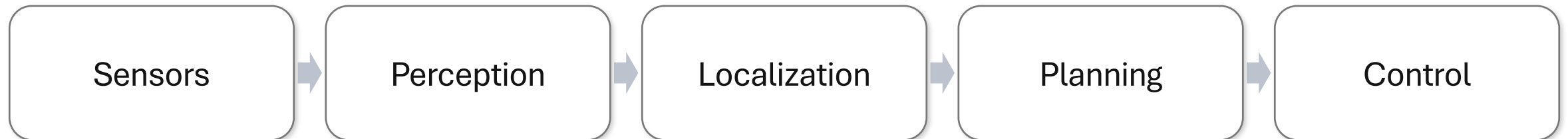
- Where to Store the Data?
 - On-board vehicle, Cloud, Network Operations Center (NOC), Edge



- Sensing
 - Obtain precise information regarding the **position** and **heading** of the CAV
 - Relative to a 3D HD map
 - Different from “perception for object detection” (covered later)
- Why sensing is hard (operating conditions)
 - Sensing must remain functional and accurate across
 - Low lighting, Tunnels, Urban canyons (tall buildings), Varying weather conditions (rain, fog, etc.)
- Multi-sensor data → Sensor Fusion requirement
 - Localization uses **multiple sensors**, each producing **different formats** of data
 - Examples: point clouds, images, ranges, velocities, inertial signals, wheel ticks
 - **Sensor Fusion**: Combining data from multiple formats into a single, cohesive world model
 - Sensors can degrade differently → need redundancy + robust fusion



- Sensor technologies used in localization are either:
 - **Exteroceptive**: sense the external surroundings (outside the vehicle)
 - **Proprioceptive**: sense the vehicle's internal motion/state (inside the vehicle)
- Key distinction
 - Exteroceptive sensors are used in mapping + sensing stages
 - Proprioceptive sensors are used only in sensing (for motion/odometry support)
- **Sensor: Passive or Active**
 - **Passive**: camera
 - **Active**: radar/LiDAR/ultrasonic



- **Sensor hardware: the 6 questions**
 - To predict **performance, failure modes, and cost/compute needs**
- **What physical quantity is measured?**
 - **Camera:** photons / intensity (RGB/IR)
 - **Radar:** RF reflections (phase/frequency)
 - **LiDAR:** time-of-flight of light
 - **IMU:** acceleration + angular rate
 - **Note:** the physics determines what the sensor can “see” and what it cannot
- **Passive or active?**
 - **Passive:** camera (depends on lighting).
 - **Active:** radar/LiDAR/ultrasonic (emits energy).
Note: active helps in darkness, but brings interference, eye-safety/regulation, and power considerations

Terminology

- FOV = field of view
- TOF = time-of-flight
- SNR = signal-to-noise ratio
- RCS = radar cross section



- **Sensor hardware: the 6 questions**
 - To predict **performance, failure modes, and cost/compute needs**
- **What is the measurement model?**
 - Example outputs:
 - **Camera:** bearing / pixels / intensity (2D)
 - **Radar:** range + Doppler + angle (often sparse, noisy)
 - **LiDAR:** 3D point cloud: range + angle (geometry-rich)
 - **Note: Perception must match the data type** (pixels vs points vs range–Doppler)
- **What limits SNR (and therefore accuracy)?**
 - **Camera:** low light, glare, fog/rain, motion blur
 - **Radar:** multipath, clutter, low RCS objects, interference
 - **LiDAR:** rain/fog, sun/retroreflectors, surface reflectivity
 - **Note:** SNR drives range/precision and failure probability.

Terminology

- FOV = field of view
- TOF = time-of-flight
- SNR = signal-to-noise ratio
- RCS = radar cross section



- **Sensor hardware: the 6 questions**
 - To predict **performance, failure modes, and cost/compute needs**
- **Sampling limits (frame/chirp rate, latency, sync)?**
 - Even perfect sensors fail if too slow or unsynchronized
 - Frame rate / chirp rate affects tracking of fast objects
 - Latency affects control and safety distances
 - Time synchronization matters for sensor fusion (camera–radar–LiDAR alignment)
 - Quick example: At 20 m/s, a 100 ms delay means the car moves 2 meters before you act
- **How does it fail, and how do we detect failure?**
 - **Failure modes:** occlusion, saturation, dropped frames, miscalibration, spoofing/interference, hardware degradation.
 - **Detection ideas:** sensor health checks, consistency checks across sensors, confidence metrics, redundancy, plausibility filters.

Terminology

- FOV = field of view
- TOF = time-of-flight
- SNR = signal-to-noise ratio
- RCS = radar cross section



- **Sensor hardware: the 6 questions**
 - To predict **performance, failure modes, and cost/compute needs**
- **Key metrics**
 - **Range:** How far the sensor can reliably detect an object
 - **Resolution:** How small a difference it can distinguish (distance/angle/speed)
 - **Accuracy:** How close the measurement is to the true value (systematic error)
 - **Precision:** How consistent repeated measurements are (random noise)
 - **Latency:** How long it takes from sensing to producing an output you can use
 - **Coverage:** What area the sensor can see (FOV + mounting + occlusions)
 - **Robustness:** How well it works in difficult conditions (weather, dirt, vibration, interference)
- Range \uparrow usually needs power / better optics / better antenna / larger aperture
- Resolution \uparrow increases compute and bandwidth
- Robustness \uparrow often needs redundancy + fusion + better hardware

