# A Behavioural Safety Centric Approach for E2E ADS

Gareth Price: SOTIF & Safety Specialist

The26262Club | March 2025

# Agenda

- Introduction to E2E AI
- Behavioural Competencies (BC)
- ML and BCs
- Influential ODD attributes
- Coverage strategy
- Measuring success

# Disclaimers and Caveats
## (my safety net😊 )

- Examples are illustrative, they do not target any particular autonomy level and are not expected to be complete

- Some illustrations assume end-to-end automated driving system

- The presented strategy targets ADS but is extendable to other applications
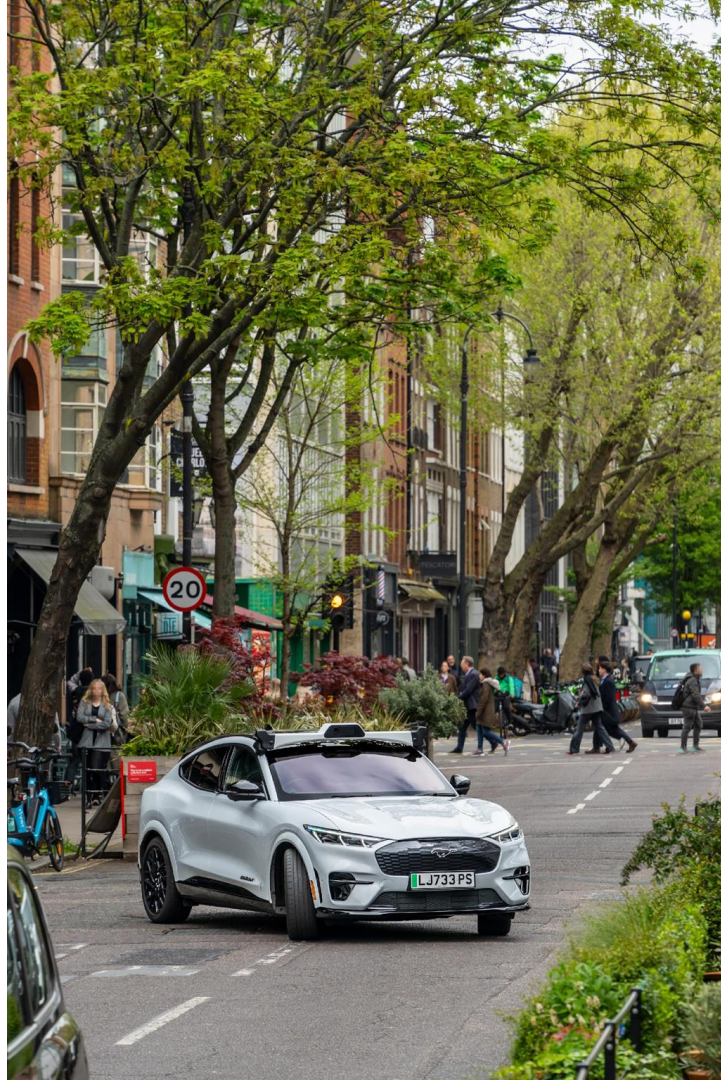
# Wayve Introduction

# Wayve Overview

Reimagining autonomous mobility with **Embodied Intelligence** Developing **end-to-end** deep learning for driving since 2017.

| | |
|---|---|
| **2018** | First to show an <u>end-to-end deep learning driving</u> policy. |
| **2020** | Expanded <u>operations to London</u>, demonstrating a safe deployment and testing framework for end-to-end AI. |
| **2022** | Demonstrated <u>generalisation</u> to new vehicles and cities. |
| **2023** | First <u>vision-language-action model</u> for driving and <u>GenAI for simulation</u> |
| **2024** | First to test same driving model in <u>US</u> & UK |
| **2025** | First to test same driving model in <u>US</u> & UK & <u>DE</u> |

Years of expertise have enabled Wayve to deliver industry leading performance, with efficient use of data and compute.

# Our mission is to reimagine autonomous mobility with *embodied intelligence*

We call this next-generation approach, AV2.0

BUILT FOR ANY VEHICLE

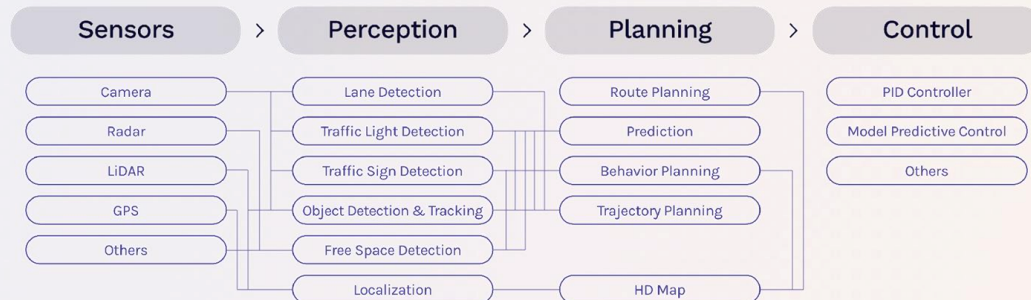BUILT FOR ANY ENVIRONMENT

BUILT FOR GLOBAL SCALE

# AV1.0
## Conventional Autonomous Driving System

Robotics-inspired architecture

Each module is developed for a specific task and tied through rule-based designs; errors can compound across modules

Localization with centimeter-accurate HD maps in real-time is the key to this approach

**AV1.0**

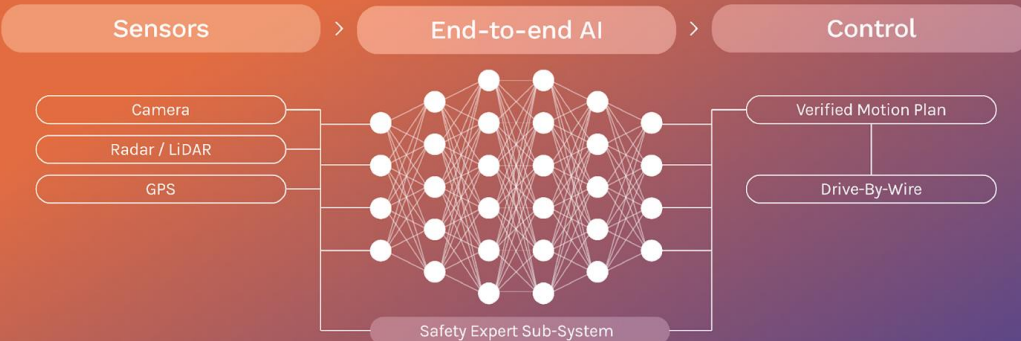| Sensors | > | Perception | > | Planning | > | Control |
|---|---|---|---|---|---|---|
| Camera | | Lane Detection | | Route Planning | | PID Controller |
| Radar | | Traffic Light Detection | | Prediction | | Model Predictive Control |
| LiDAR | | Traffic Sign Detection | | Behavior Planning | | Others |
| GPS | | Object Detection & Tracking | | Trajectory Planning | | |
| Others | | Free Space Detection | | | | |
| | | Localization | | HD Map | | |

# AV2.0
## End-to-end (e2e) AI System

An AI-driven, streamlined solution that combines perception, prediction, and planning into a single end-to-end neural network.

Transforms raw sensor data into safe driving actions

Data-driven approach enables efficient improvements through scaling data

**AV2.0**

| Sensors | > | End-to-end AI | > | Control |
|---|---|---|---|---|
| Camera | | | | Verified Motion Plan |
| Radar / LiDAR | | | | Drive-By-Wire |
| GPS | | | | |

Safety Expert Sub-System

Right turn on red light (prohibited in the UK)

Navigating through crowds and temporary roadworks
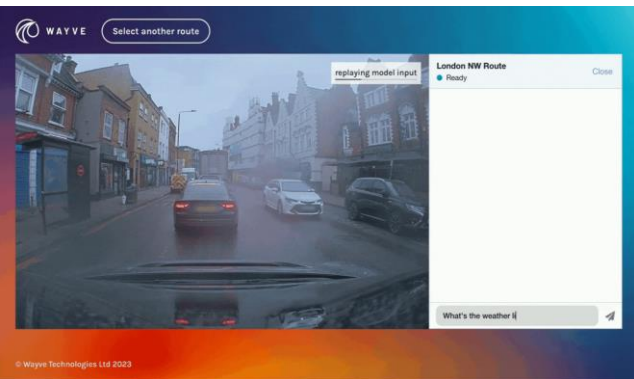
First drive in snow (never seen during training)

Zero shot generalization on one-lane country road

# World-leading Embodied AI Science

**AI EXPLAINABILITY**

**PHOTOREALISTIC SYNTHETIC DATA**

**FUTURE PREDICTION**



LINGO-1
First language model for self-driving

PRISM-1
End-to-end data driven simulation

GAIA-1
First generative world model

# What makes an ADS Safe?

# What makes an ADS Safe?

*confidence* that the ADS exhibits safe *driving skills* within the *ODD*

1) Check that:
   - The ADS understands the main driving rules and associated decisions (*i.e. skills*)
     - e.g. stopping at a signalised intersection, yielding at a roundabout
   - The ADS can safely apply those skills in various real-world situations (*i.e. ODD*)
     - e.g. stopping at a signalised intersection in bad weather
     - e.g yielding at a roundabout under heavy traffic

1) Gain a certain level of *confidence* that the above is true

| Learned skills | + | Over ODD | + | Confidence | ≈ | Safe ADS |

↳ Behavioural Competencies (BCs)

# Defining Driving Skills

Intended behaviour proven across ODD = competency acquired

A Behavioural Competency (BC) is:

"*The ability to handle __segments and nodes__ of the road topology to achieve a strategic __goal__ whilst respecting a __set of rules__*".

- e.g. The ability to turn left *[goal]* at a roundabout *[node]*.
- e.g. The ability to stop *[goal]* at a zebra crossing *[node]*
- e.g. The ability to go straight *[goal]* at an intersection *[node]*.

*Note 1  BCs are mostly ODD agnostic except at the road topology level*
*Note 2  The applicable rules vary depending on the goal*
*            therefore the goal is part of the BC definition*



[DriveSafeAI blog @ drive-safe.ai]

# DriveSafe AI: developing the evidence base for advanced safety assurance methods for assisted and automated vehicles



A Research Project in Partnership With

WAYVE

WMG
THE UNIVERSITY OF WARWICK

Exploring Four Open Innovations

Generalisable safety assurance methods

Guidance for simulation correlation methods and validation

Safety Pool™ scenario datasets

AI safety architecture design rationale

Informing evidence-based policy-making and industry guidelines, standards and regulations

# What does it take to learn driving?

# What does it take to learn driving?

## Theoretical Learning

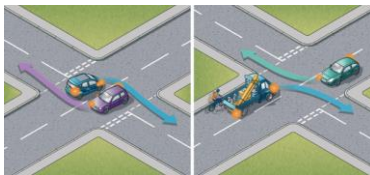- Learn the rules of the road that dictate driving decisions
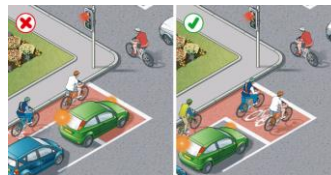
*Examples used in this presentation*

### Ex.1: Zebra crossings

### Ex.2: Unprotected turn

### Ex.3: Traffic lights

# What does it take to learn driving?

## Practical Learning

- Practice vehicle controls

- Practice application of the driving rules in the real-world



[source: https://gonpass.com/]

# What does it take *for an ADS* to learn driving?

## Theoretical Learning

- Learn the rules of the road that dictate driving decision

Training datasets f(BCs, ODD)

Driving model learning

# What does it take *for an ADS* to learn driving?

## Practical Learning

- Practice vehicle controls

- Practice application of the driving rules (BCs) in the real-world (ODD)

**1.  Test all BCs**



**Scen. 1(7):** Negotiating zebra crossing

**Scen. 2(3):** unprotected right turn at T-junction

**Scen. 2(6):** traffic light w/ cycle box

**2.  Test BC's sensitivity to ODD variations**



**3.  Repeat to gain confidence**

**Takeaway: Testing starts with BCs**

# Influential ODD Attributes

# Influential ODD attributes

## Defining driving skills

…the rules of the road that *dictate* driving *decisions*…

…so *what element(s)* in the scene *dictate* our *decision*?
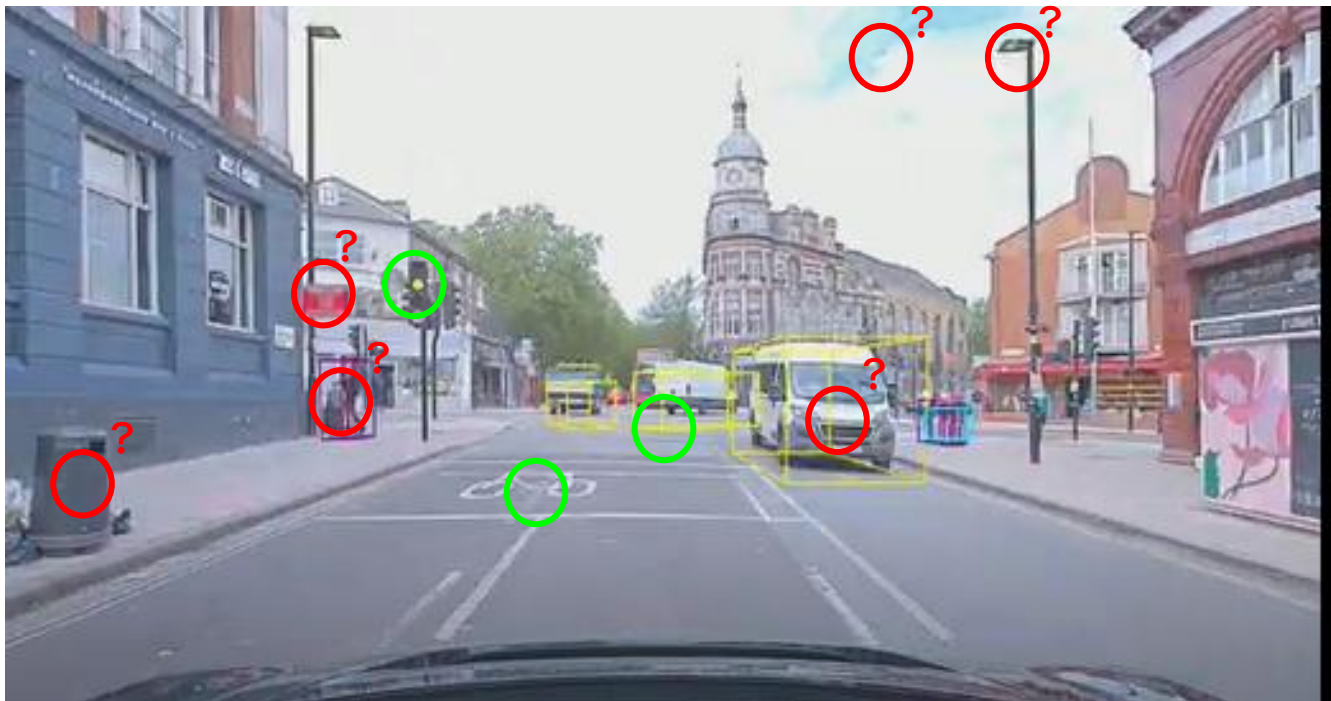
# Influential ODD attributes

## Defining driving skills

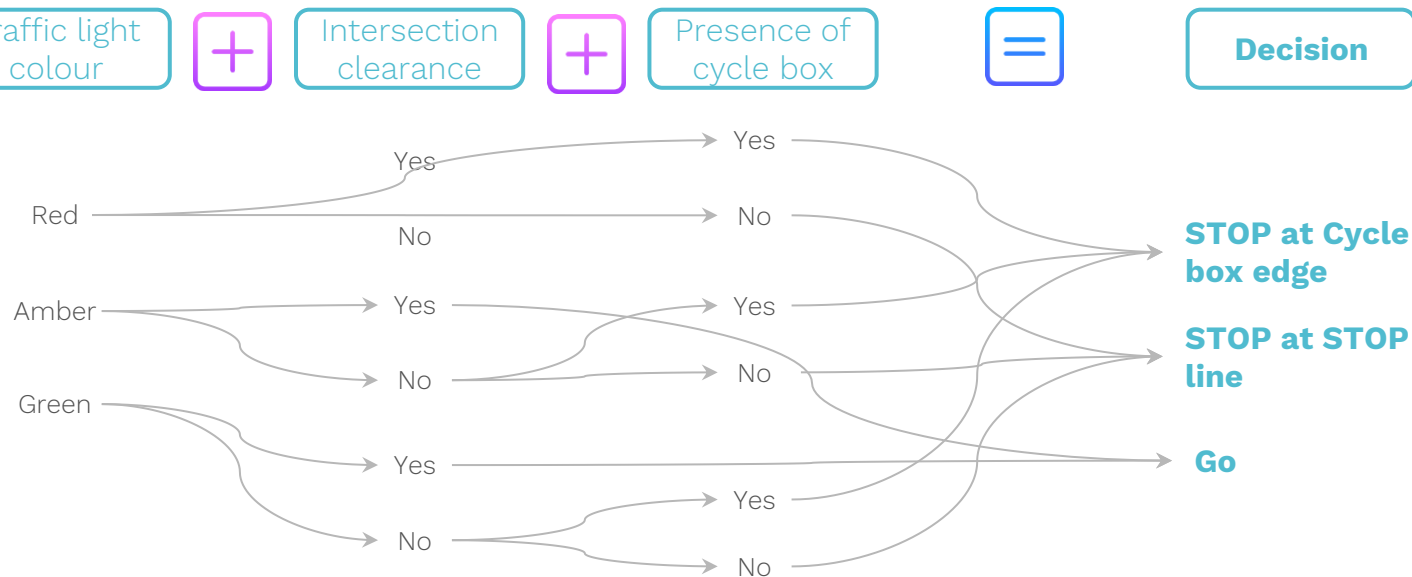Only a *subset of ODD attributes dictate* our *decisions*

↓

Influential ODD attributes:
- Light colour
- Cycle box
- Intersection clearance

# Influential ODD Attributes

| Traffic light colour | + | Intersection clearance | + | Presence of cycle box | | = | | **Decision** |

Red
  Yes → Yes → **STOP at Cycle box edge**
  No → No

Amber
  Yes → Yes
  No → No → **STOP at STOP line**

Green
  Yes → **Go**
  No → Yes
     → No

**Takeaways:**
1. Influential attributes characterise the driving skill,
    i.e. making the right decision = Behavioural Competency
2. Influential attributes are specific to each Behavioural Competency

# Coverage strategy

# Coverage Strategy

## Putting everything together...
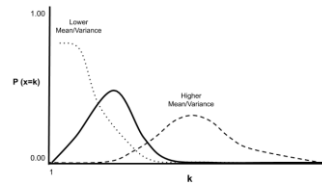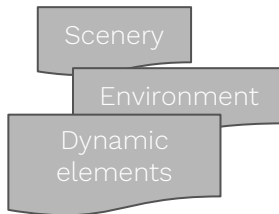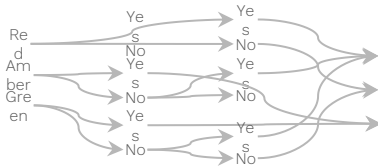
*Purpose*

| Learned skills | $+$ | Over ODD | $+$ | Confidence | $\approx$ | Safe ADS |

*Artefacts*



*Strategy*

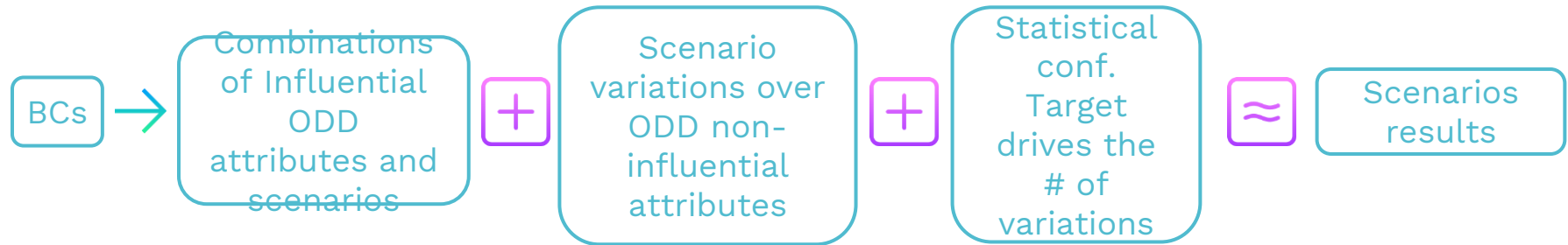BCs → Combinations of Influential ODD attributes and scenarios  $+$  Scenario variations over ODD non-influential attributes  $+$  Statistical confidence drives the # of variations  $\approx$  Scenarios results

# Coverage Strategy

BCs → Influential attr. → combinations → variations f(ODD) → SPIs results

BCs → Combinations of Influential ODD attributes and scenarios + Scenario variations over ODD non-influential attributes + Statistical conf. Target drives the # of variations ≈ Scenarios results

*Note 1* The application defines the ODD and the required BCs (e.g. Robotaxi, highwaypilot etc)
*Note 2* each BC's ODD is equal to or a subset of the application's ODD
*Note 3* not all combinations of influential attributes are valid

**Takeaways:**
1) The ODD exploration is driven by the influential attr. ⇒ we test what matters
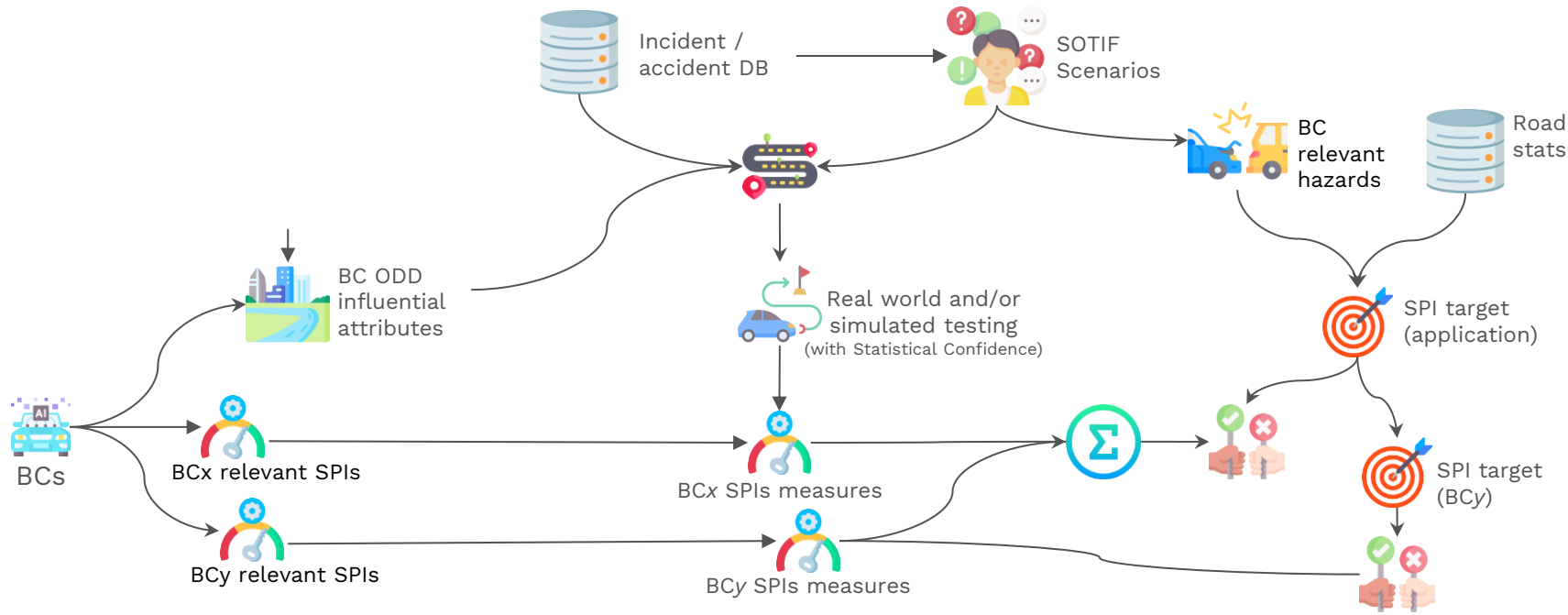2) The ODD exploration is contained by the influential attr. ⇒ no explosion of scenarios

# Measuring success

# Measuring Success

BCs → Influential attr. → combinations → variations f(ODD) → SPIs results



Incident / accident DB

SOTIF Scenarios

BC relevant hazards

Road stats

BC ODD influential attributes

Real world and/or simulated testing (with Statistical Confidence)

SPI target (application)

BCs

BCx relevant SPIs

BCy relevant SPIs

BCx SPIs measures

BCy SPIs measures

Σ

SPI target (BCy)

**Takeaway: Success measured by Application's SPIs and individual BC's SPIs**

# Thank you! 😊