Extending automotive certification processes to handle autonomous vehicles

Dr Zeyn Saigol
Principal Technologist | 14th November 2019

CATAPULT
Connected Places

Why is certifying AVs an important problem?

All major car manufacturers are now, somewhat to their surprise, actually multi-billion-dollar robotics startups.

"Startups", because they've been thrust into developing products that they have no history or knowledge of before around 2014.

This creates a safety challenge

- OEMs (car manufacturers) have limited experience of verifying complex robotics systems
- They do have a lot of experience of verifying complex mechanical systems, and this experience doesn't directly translate
- Verification of AVs is just a really hard problem

CATAPULT
Connected Places

Outline

- AV (autonomous vehicle) challenges
- Traditional automotive safety assurance
- Why AVs, and regulating AVs, are different
- Shape of the technical solution for certification
- CPC work: MUSICC and VeriCAV
- Remaining challenges, and the future

CATAPULT
Connected Places

Connected Places Catapult

CATAPULT
Connected Places

Our mission

To help British businesses address the grand challenges of today in order to create connected places, fit for the future.

Our vision

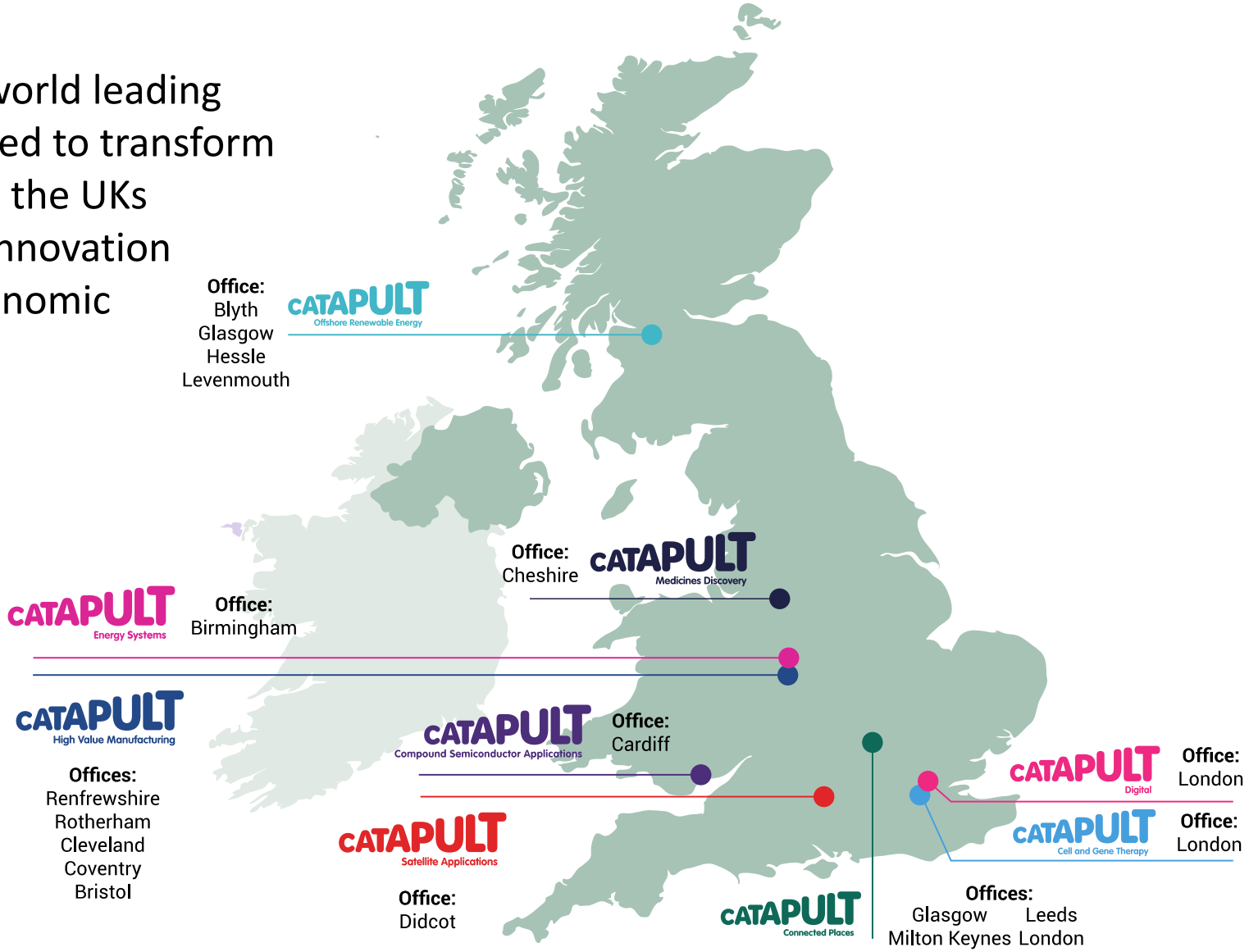
For the UK to lead the world in creating cities, towns and places which thrive on their ability to connect people to resources, opportunities, ideas and each other. Where the smooth flow of people, goods, transportation and services, drives economic success, productivity and wellbeing.

Delivering and growing

- New market opportunities for businesses
- Social and environmental benefits to places
- Robust transportation networks and mobility strategies fit for the next generation

Catapults – a force for innovation and growth

A network of world leading centres designed to transform and accelerate the UKs capability for innovation and future economic growth.



CATAPULT
Connected Places

9

Innovation Centres across the UK

AV challenges

CATAPULT
Connected Places

AV interest and investment

Uber's self-driving car unit valued at $7.3bn as it gears up for IPO

US firm received $1bn from consortium including Toyota and Saudi Arabia

▲ Uber is a loss-making firm but has received financial backing to develop autonomous cars. Photograph: Hollis Bennett/Uber

Uber's self-driving car unit has been valued at $7.3bn (£5.6bn), after receiving $1bn of investment by a consortium including Toyota and Saudi Arabia's sovereign wealth fund.

The Guardian, 19 April 2019

VW invests $2.6 billion in self-driving startup Argo AI as part of Ford alliance

Kirsten Korosec @kirstenkorosec / 12:59 pm BST • July 12, 2019 Comment

TechCrunch, 12 July 2019

AVs promise:

- Reduction in road casualties
- Better mobility for the elderly and disabled
- Freeing up unproductive time

These have prompted **$billions of R&D investment**

CATAPULT Connected Places

AVs are robots

=

Same technical challenges

- Perception
- Decision making
- Acting

Added safety concerns

- Bigger
- Faster
- Operate alongside the general public

CATAPULT
Connected Places

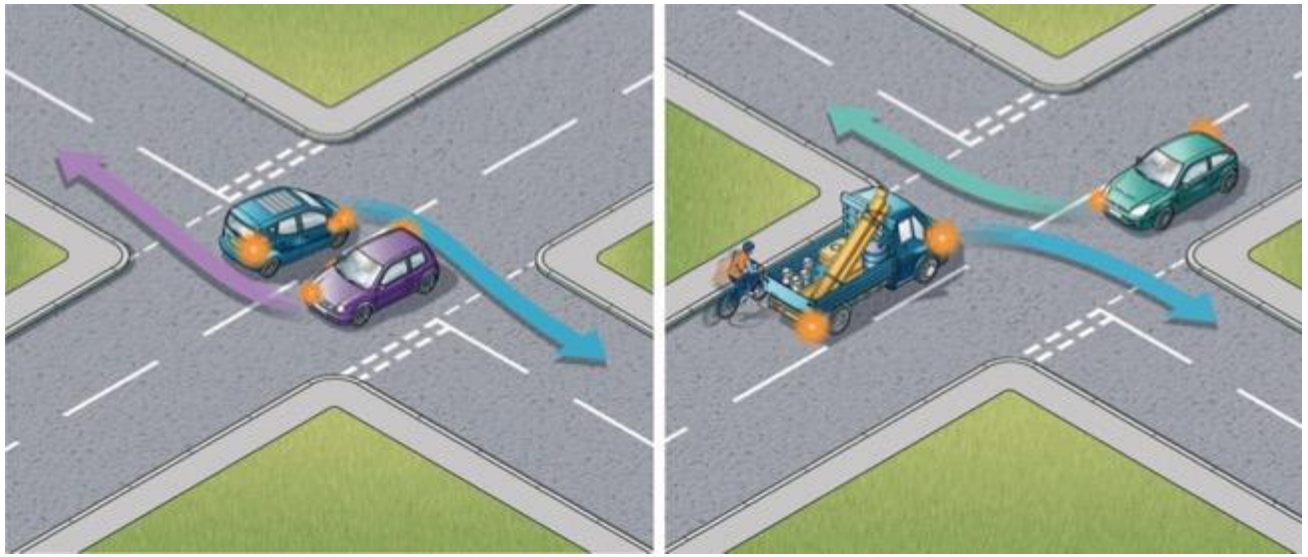
Why are AVs especially challenging?

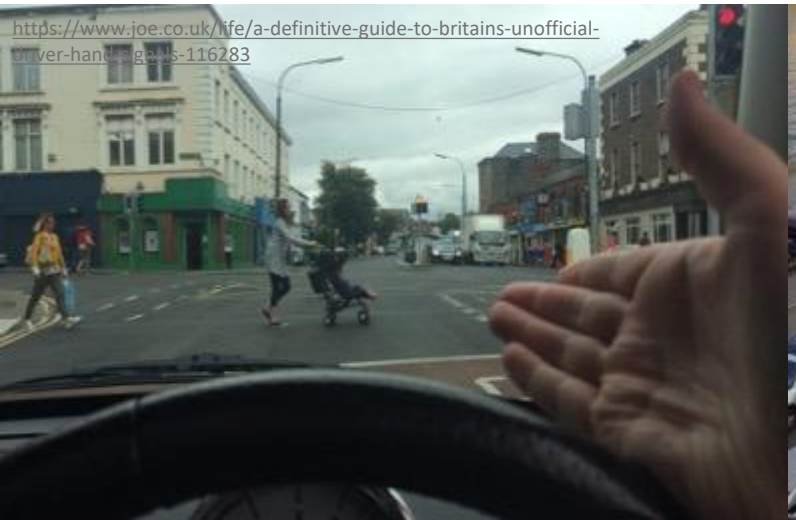
#1: Complex, diverse, and changeable environment

CATAPULT
Connected Places

Why are AVs especially challenging?

#2: Complex rules + human interaction





https://www.joe.co.uk/life/a-definitive-guide-to-britains-unofficial-driver-hand-signals-116283

Why are AVs especially challenging?

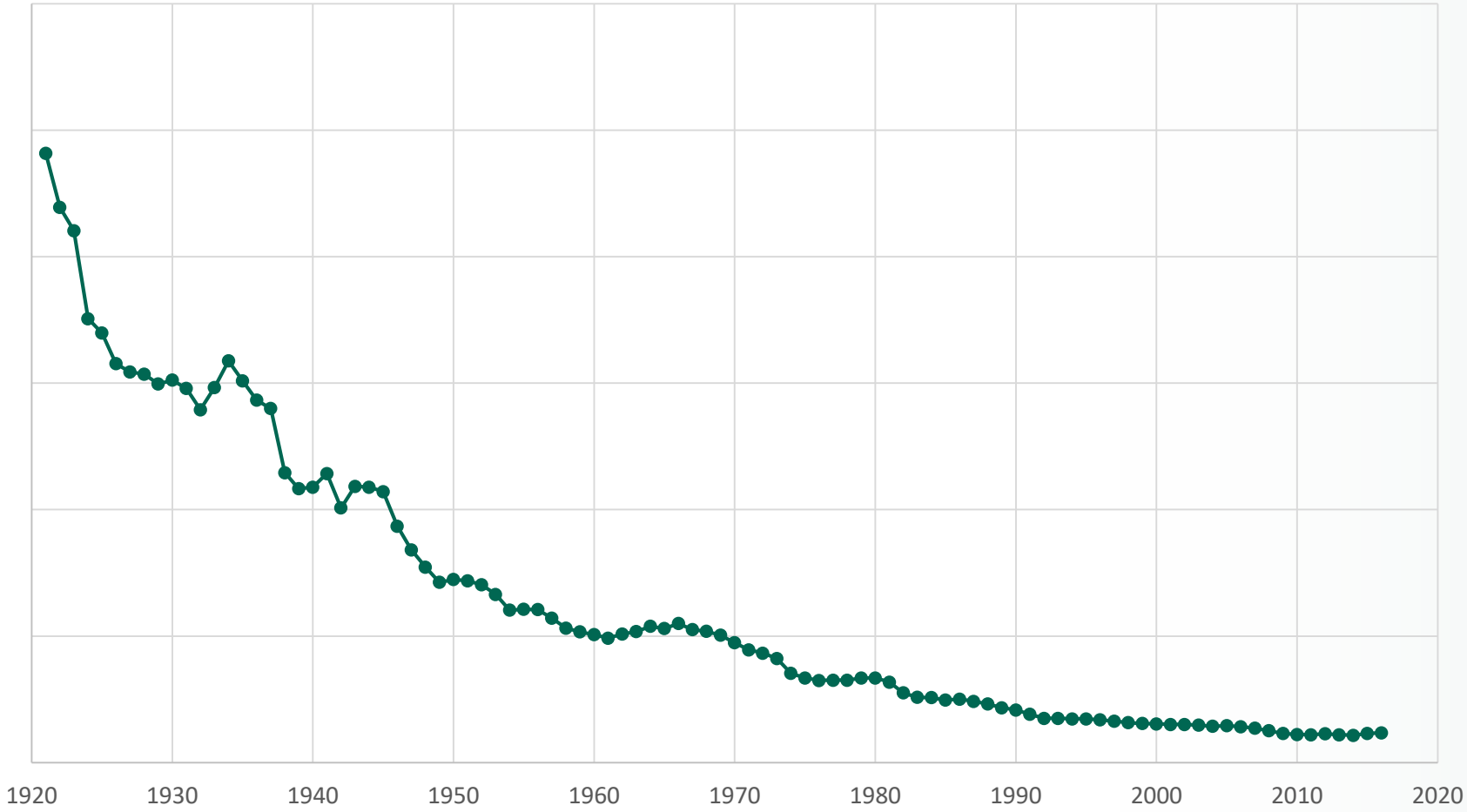
#3: Perception challenges



Traditional automotive safety assurance

CATAPULT
Connected Places

History of automotive safety



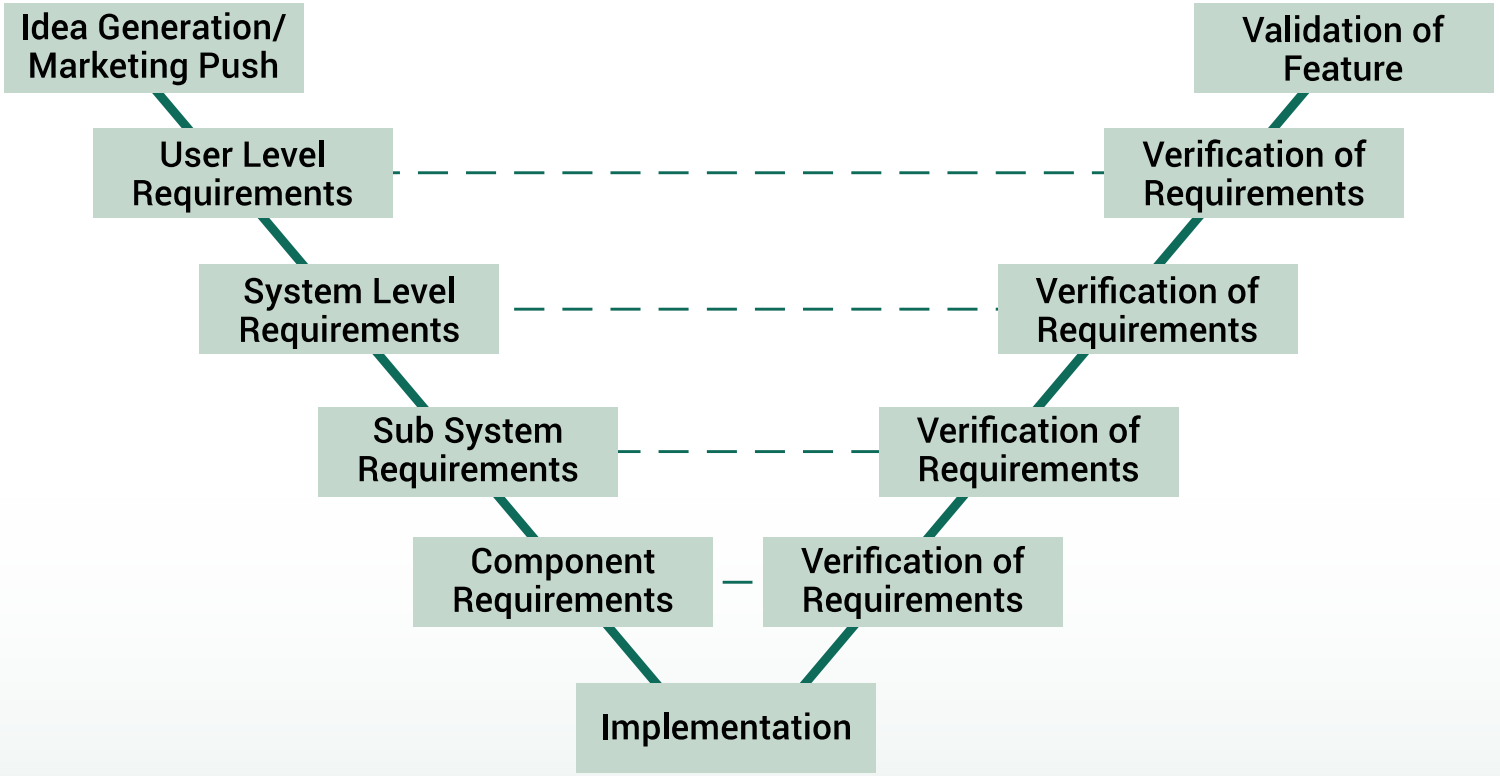
US fatality rate per 100 million vehicle miles travelled

Automotive industry safety processes are highly effective

They are also well established and very prescriptive

CATAPULT
Connected Places

Traditional automotive safety assurance: V-cycle



Systems engineering V-model

Standard process for verification and validation

- Designed to ensure nothing ‘slips through the gaps’

CATAPULT Connected Places

Traditional automotive safety: ISO 26262

Risk-based functional safety methodology

- Designed to apply to all electronic and software systems on a vehicle
- ADAS systems (e.g. lane-keep assist), but also electronic stability control, ABS, and even fuel injection systems
- Processes to be followed at all stages of V-cycle
- Functional safety based:
  - Consider all possible failures, and the likely severity of the consequences
  - Use these to assign an Automotive Safety Integrity Level (ASIL) to the failure
  - Higher ASILs require more robust processes for specification, development, and V&V
- Traceability of requirements, specification, and implementation, use of change control, use of safe coding standards such as MISRA C

According to industry insiders, verification and validation can absorb

40%

of the budget for developing a new model

CATAPULT
Connected Places

Traditional automotive safety: beyond failures

SOTIF (safety of the intended functionality, ISO/PAS 21448)

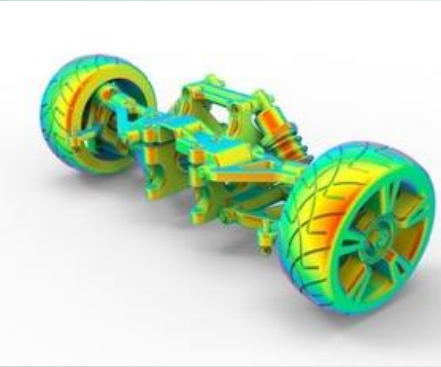
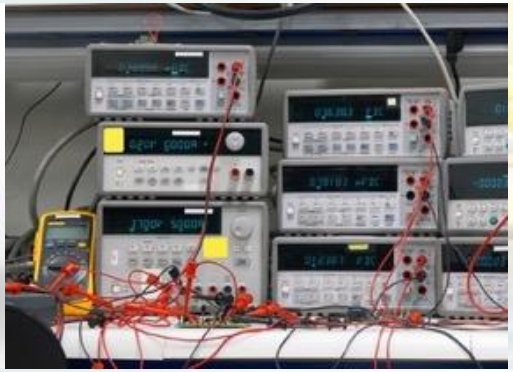
- ISO 26262 only considers failures of electrical/software systems
- SOTIF fills in some of the gaps – focus on complex systems that use sensors to build up a situational awareness
- “functional insufficiencies of the intended functionality”: spec bugs
- Still a hazard-focused, process-based standard

CATAPULT
Connected Places

Traditional automotive safety: testing

Testing is exhaustive and manual

- Proceeds through simulation, hardware-in-the-loop, VeHIL, private track tests and public road tests
- Final testing with multiple vehicles and continents (ensure cover all weather conditions)
- Test drivers working in shifts, and still takes many months

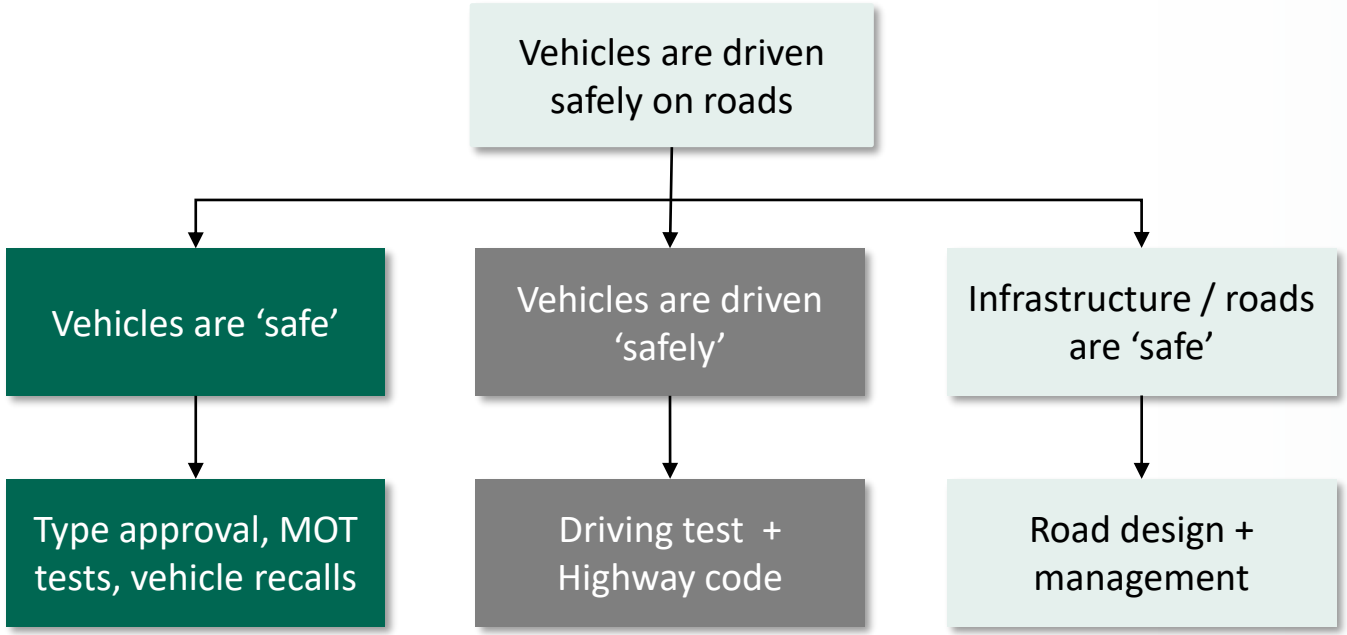

CATAPULT
Connected Places

Why AVs, and regulating AVs, are different

CATAPULT
Connected Places

Why doesn’t this map to AVs?

Certification of Automated Driving Systems

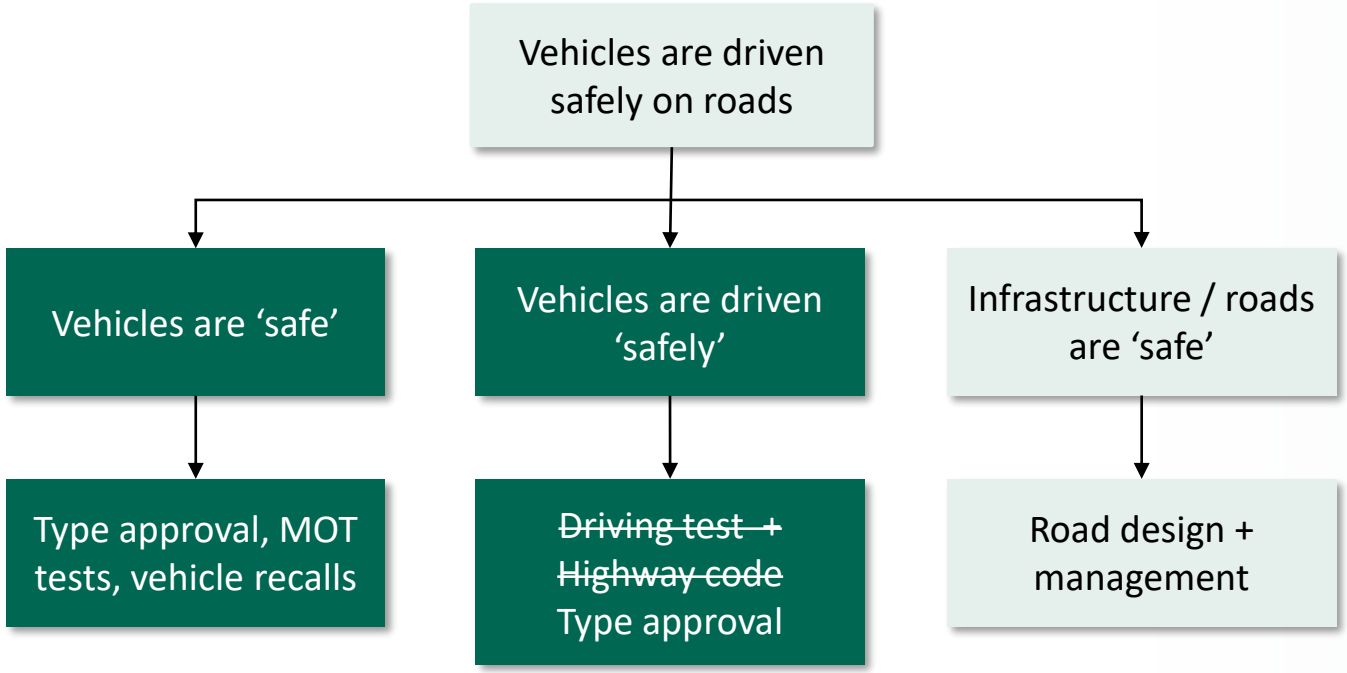


UK processes for assuring road safety

CATAPULT Connected Places

Why doesn't this map to AVs?

Certification of Automated Driving Systems



UK processes for assuring road safety

Fully autonomous vehicles require a completely new type of testing to be included in type approval

- Partial autonomy (e.g. Teslas) is different

CATAPULT Connected Places

Why doesn't this map to AVs?

Certification of Automated Driving Systems

> Can't achieve coverage needed by just testing on public roads: "To demonstrate that fully autonomous vehicles have a fatality rate of 1.09 fatalities per 100 million miles [...] with a fleet of 100 autonomous vehicles being test-driven 24 h a day, 365 days a year at an average speed of 25 miles per hour, this would take about 12.5 years."[1]

- **The dynamic driving task has an input space too large and complex to test using traditional methods**
  - Not possible to write a comprehensive specification for the task
- **26262 and the V-cycle apply to simpler systems**
  - Random hardware failures are a major consideration
  - ASIL categories assume a human driver is present to mitigate any failure
- **Spec errors and complex system interaction failures are key for AVs**

[1] "Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?" Nidhi Kalra & Susan M. Paddock, RAND Corporation 2016. https://www.rand.org/pubs/research_reports/RR1478.html

CATAPULT Connected Places

Requirements for regulation

Architecture and fairness

- Test process must work with any ADS architecture
- Must be seen as fair: cannot advantage any specific developer or technology
- Should not constrain innovation

Test rigour

- OEMs should not be able to design-to-test
- Randomisation of test cases would help prevent this
- At the same time, tests should be repeatable

Fit

- Must work internationally
- Must work within existing regulatory regime

CATAPULT
Connected Places

Regulatory challenge

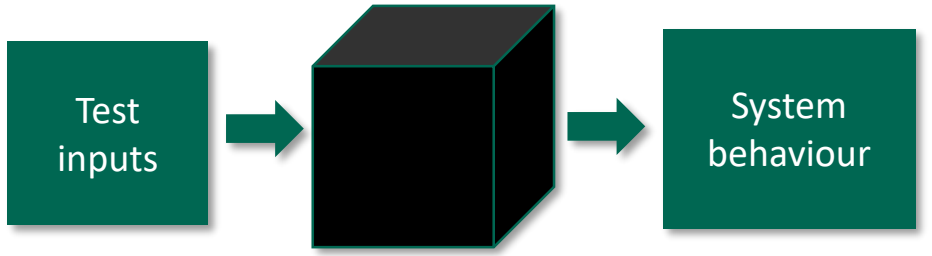
Independent certification testing

Context for type approval

- In Europe, regulators strive to provide independent assurance of the safety of products
- This implies certification tests should be conducted by an impartial organisation

Black box testing

- Likely to be necessary, given independent testing, architecture neutrality, and (current) reluctance of OEMs to provide access within their systems
- Prevents testing individual components – in particular, unable to test perception separately
- Prevents application of code and model checking methods



System-under-test is a black box

Novelty

- Very different to existing regulations
- Even concept of regulations that apply to software is fraught

Shape of the technical solution for certification

CATAPULT
Connected Places

Solution 1: Simulation automotive safety: testing

CARLA simulator http://carla.org/

Real-world testing can't provide the coverage

Simulation means you can:

- Cheaply run many tests in parallel
- Potentially run tests faster than real-time
- Avoid danger to participants
- Control test parameters precisely

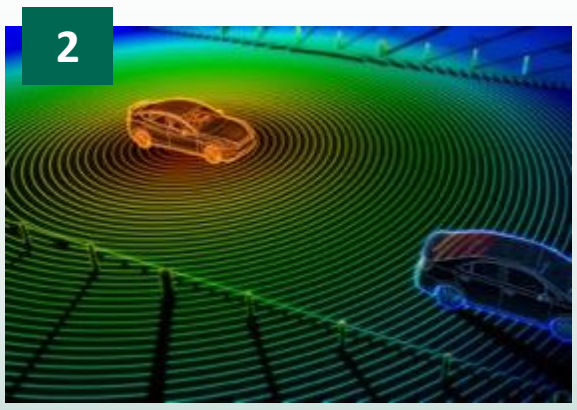
Simulation – what's the challenge?

Need to simulate the whole environment

- This is much harder than previous simulations used in automotive

Modelling challenges include:

1. The physical environment, ideally in sub-mm detail
2. Sensors, corresponding exactly to sensor models used on the AV
3. Weather
4. Actions of other road users



Solution 2: Scenarios

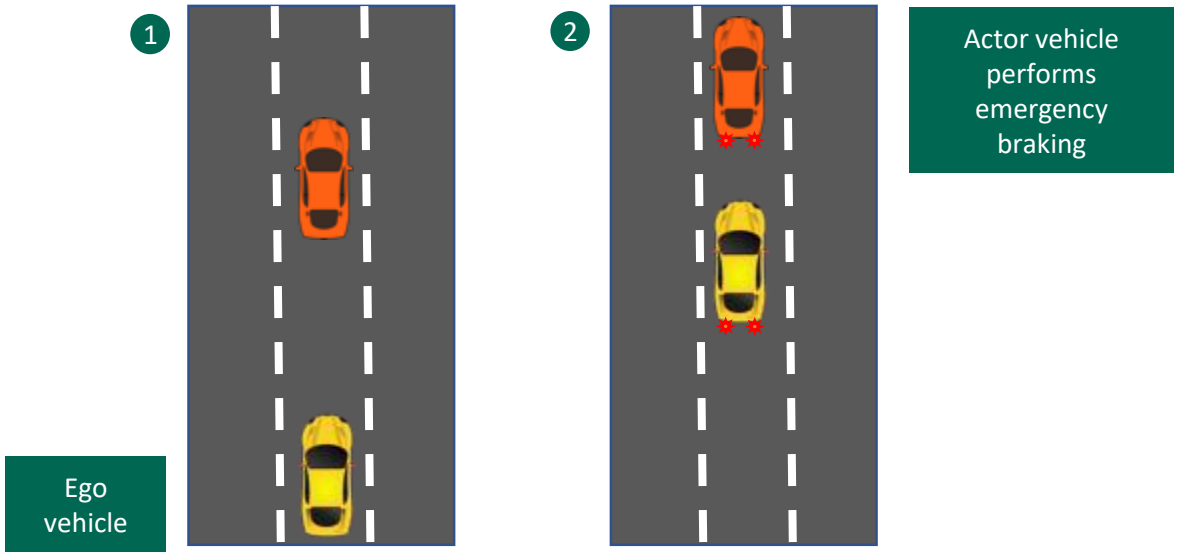
Simulation alone doesn't boost coverage enough

A lot of testing is uninformative

- Unlikely to find failure cases

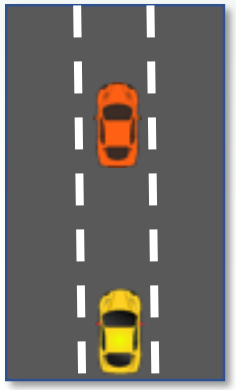
Instead, test against defined scenarios

- Test far more edge cases than would be encountered in everyday driving

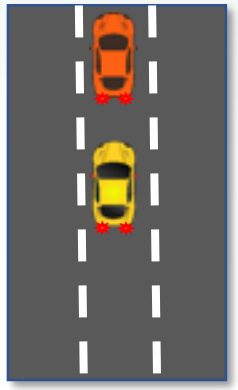


What does this mean for regulation?

1



2



- Traditionally regulations specify concise, explicit performance standards
- Most stakeholders are clear that scenarios represent the most effective way of specifying the test cases for certification
  - Given enough scenarios at the right level of abstraction, almost all cases can be captured
- The certification process would then be driven from a shared international database of scenarios

MUSICC
Multi User Scenario Catalogue for Connected Autonomous Vehicles

Department for Transport

- CPC, working with the DfT, have created an open, online scenario database for certification – MUSICC

CATAPULT
Connected Places

CPC work: MUSICC and VeriCAV

CATAPULT
Connected Places

MUSICC: Multi-User Scenario Catalogue for CAVs

Department for Transport

MUSICC
Multi User Scenario Catalogue for Connected Autonomous Vehicles

CATAPULT
Connected Places

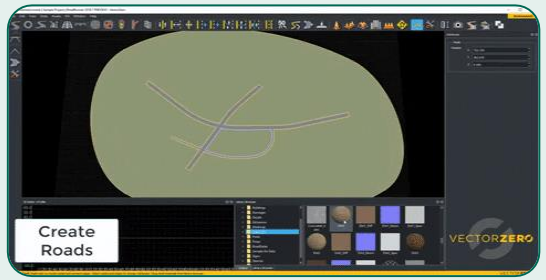
Objectives:
- Implement a language to describe scenarios, aligned with industry standards
- Build an open and expandable library for CAV certification scenarios

Approach:
- Proof-of-concept project, Apr 2018 – Mar 2020
- Close collaboration with vehicle manufacturers, ADS developers, organisations with expertise in CAV validation, and regulators
- Focus on simulation testing environments

MUSICC scope and context

MUSICC
Multi User Scenario Catalogue for Connected Autonomous Vehicles

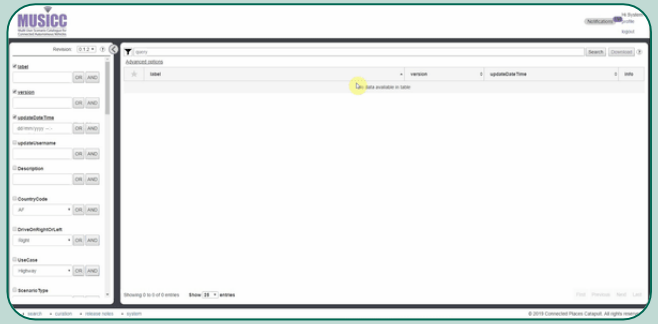
Scenario generation (external tools)



Central regulatory database

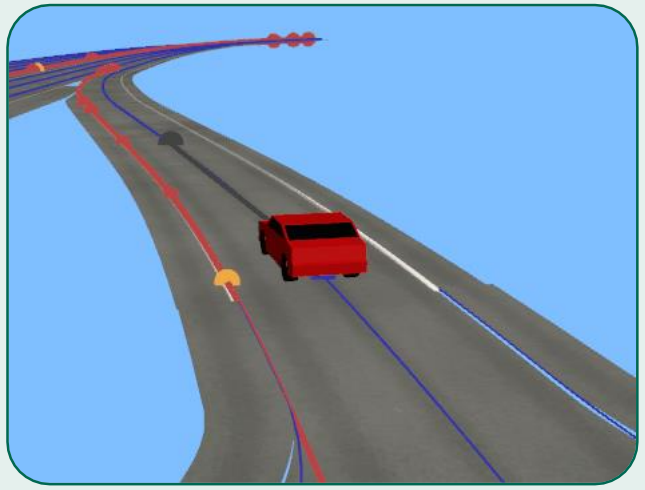
Scenario library

Export API



Web Interface
- Import
- Search
- Basic editing

Regulatory testing (external tools)



Made possible by use of open standards, including ASAM OpenDRIVE and OpenSCENARIO

CATAPULT Connected Places

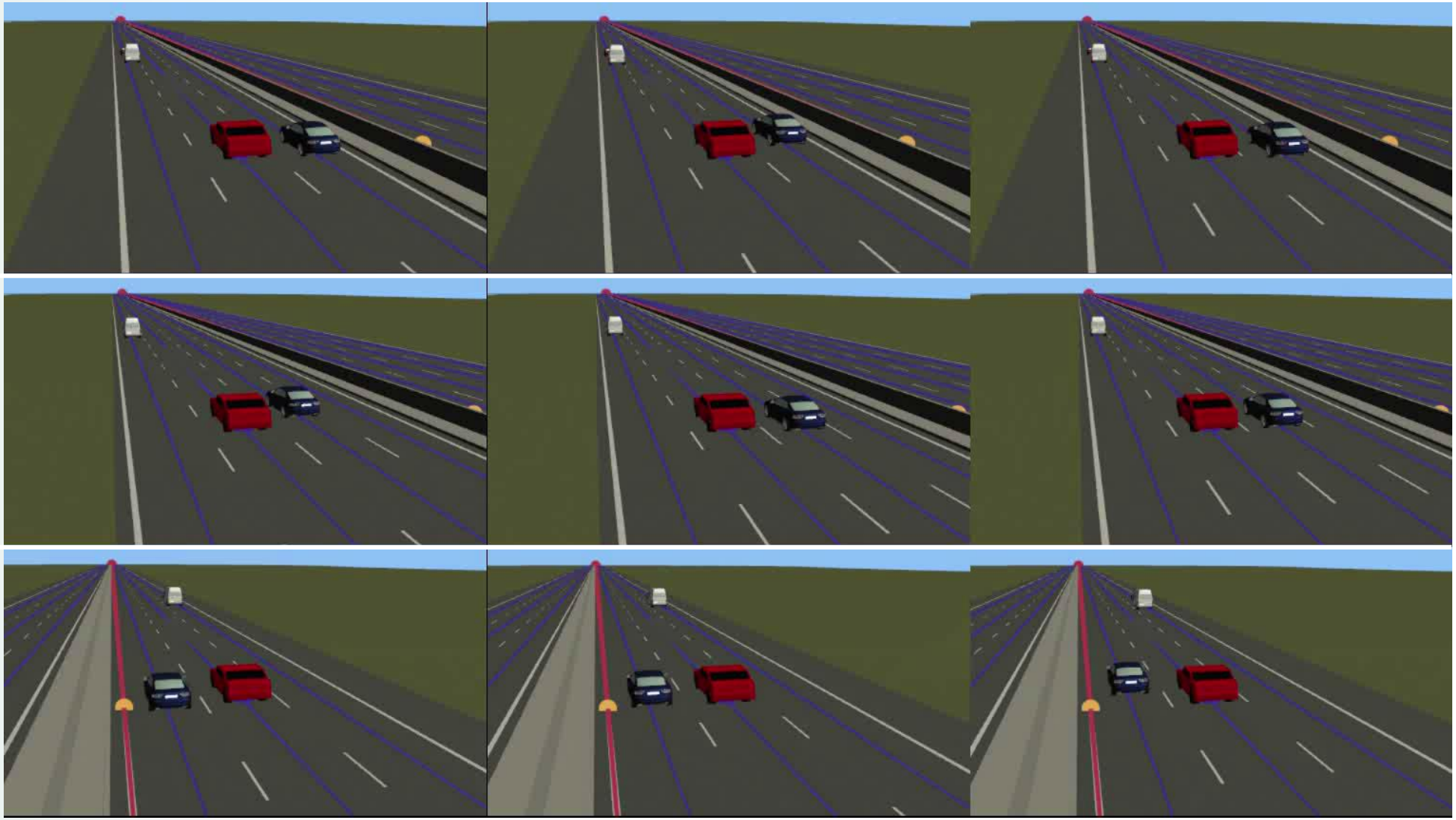
Randomisation

Generates multiple concrete scenarios from each abstract scenario

3 lane - GB

4 lane - GB

3 lane - FR



Research topic #1:

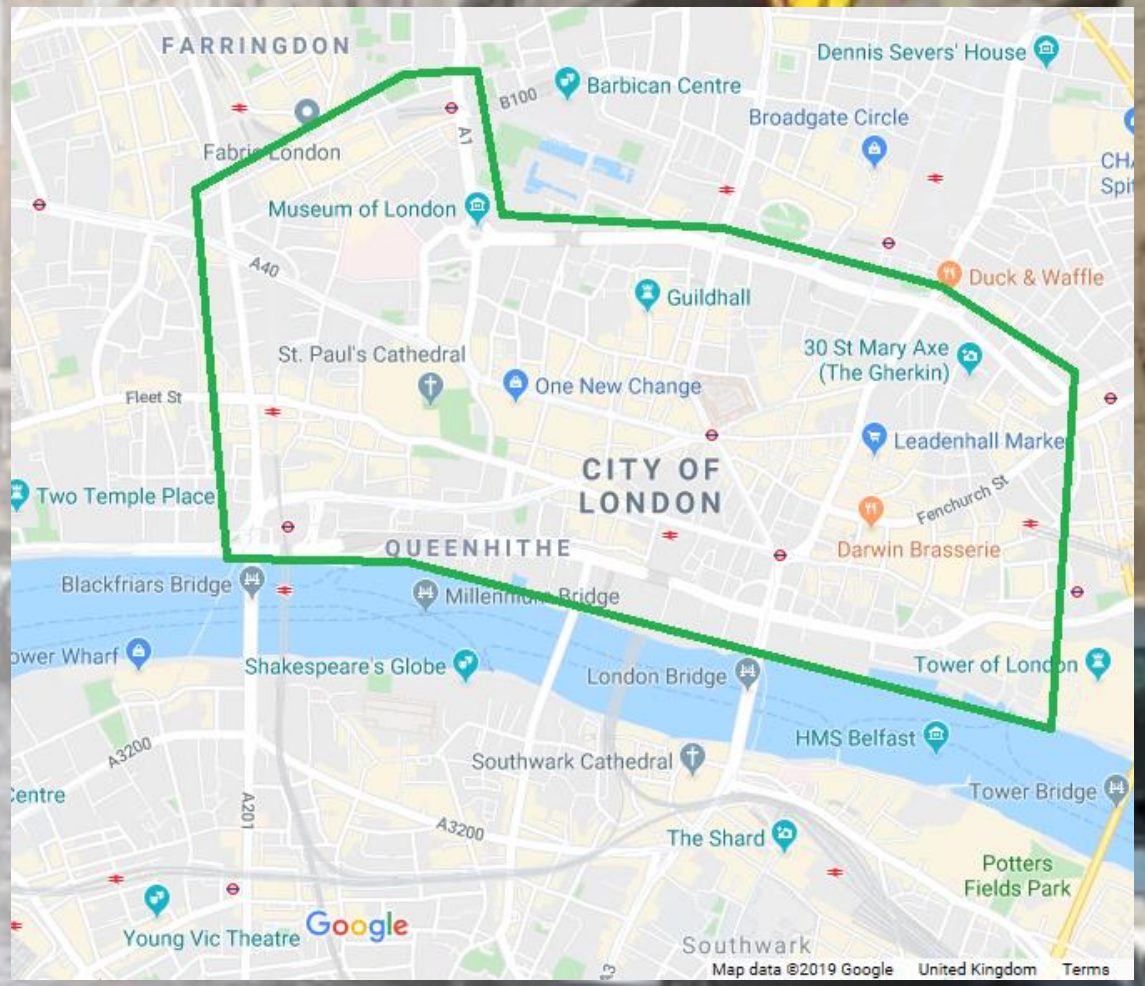
Representing the ODD

Operational Design Domain is critical, given technical challenges of ADS

ODD defines conditions under which ADS will operate.

Can cover:

- Weather
- Time-of-day
- If in cities only, motorways only, etc
- Road type restrictions
- Explicit geofence
- Traffic levels
- Possible manoeuvres



MUSICC
Multi User Scenario Catalogue for Connected Autonomous Vehicles

CATAPULT
Connected Places

Research topic #1:

Representing the ODD

MUSICC
Multi User Scenario Catalogue for Connected Autonomous Vehicles

Critical to test all the applicable scenarios for an ODD

- MUSICC supports certification tests by allowing ODD-aligned queries

ODD representation needs an ontology plus a definition language

Ontology

Physical infrastructure
  Road type
    Arterial
      Urban
        ……
      Rural
    ……
Environmental conditions
  Weather
    ……
  Road surface conditions
  ……

Language [WIP]

- Basic approach to list permissible items within each top-level category
- Complications with dependencies within and between categories

For example:
– Work on motorways when precipitation one of (none, light rain, medium rain)
– Work on trunk roads, so long as there are no roundabouts

- Not currently any standard for ODDs
- We will work with ASAM and BSI on standardising the ODD representation language

CATAPULT
Connected Places

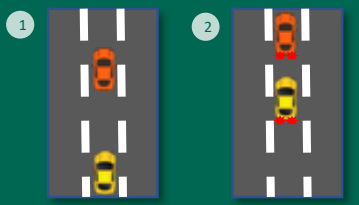
Research topic #2:

Representing the required performance standard

MUSICC
Multi User Scenario Catalogue for Connected Autonomous Vehicles

- Not just collisions – they may be unavoidable in certain scenarios
- Also consider rule compliance, safety margins, confusing behaviour, and making progress

Digital Highway Code



Per-scenario pass/fail criteria

MUSICC's scenario-specific language [WIP]

- Scoring depends on inputs such as position within the lane, lane/road departures, speeds and accelerations, and minimum distances to other actors
- These kinds of criteria require a powerful language to express
- Framework consists of Python core, plus:
  - Set of parameterised variables that can be used in pass/fail criteria
  - A standard way of reporting failures or scores
  - A library of common functions (e.g. assert-vehicle-did-not-collide)

CATAPULT
Connected Places

VeriCAV project

- 24-month CR&D project
- Started January 2019
- Supported by:

Centre for Connected & Autonomous Vehicles

Innovate UK

Consortium

With HORIBA MIRA the industrial lead

www.vericav-project.co.uk

VeriCAV aims

VeriCAV is addressing three important challenges for testing in simulation:

1) Level of **human effort** in test management is considerable and slows testing

2) **Behaviour of other actors** in simulation is insufficiently realistic

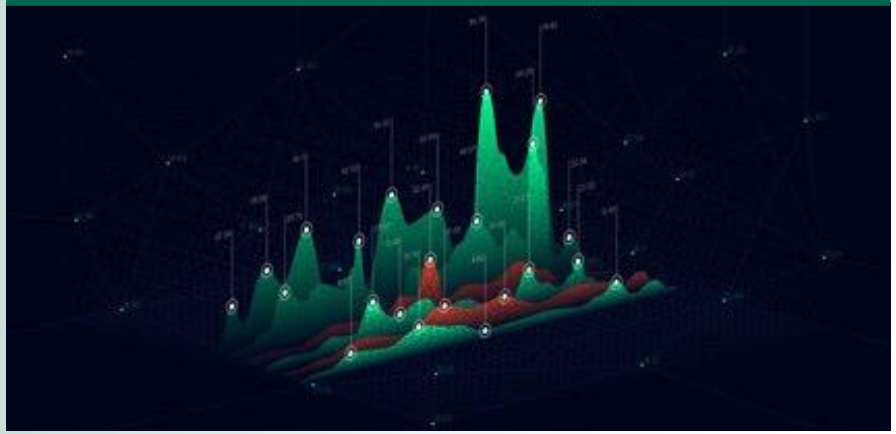
3) **Interfaces** between ADS and test framework tools are not mature

Coverage and manual effort

VeriCAV
VERIFICATION OF CONNECTED AND AUTONOMOUS VEHICLES

CATAPULT
Connected Places

Ensure coverage of test space



Focus on critical areas for the ADS under test

- Enabled by the Test Oracle: automated analysis of the performance of the ADS under test
- Feeds test results back into the randomisation engine, which uses these to focus tests on the most informative areas of the test space

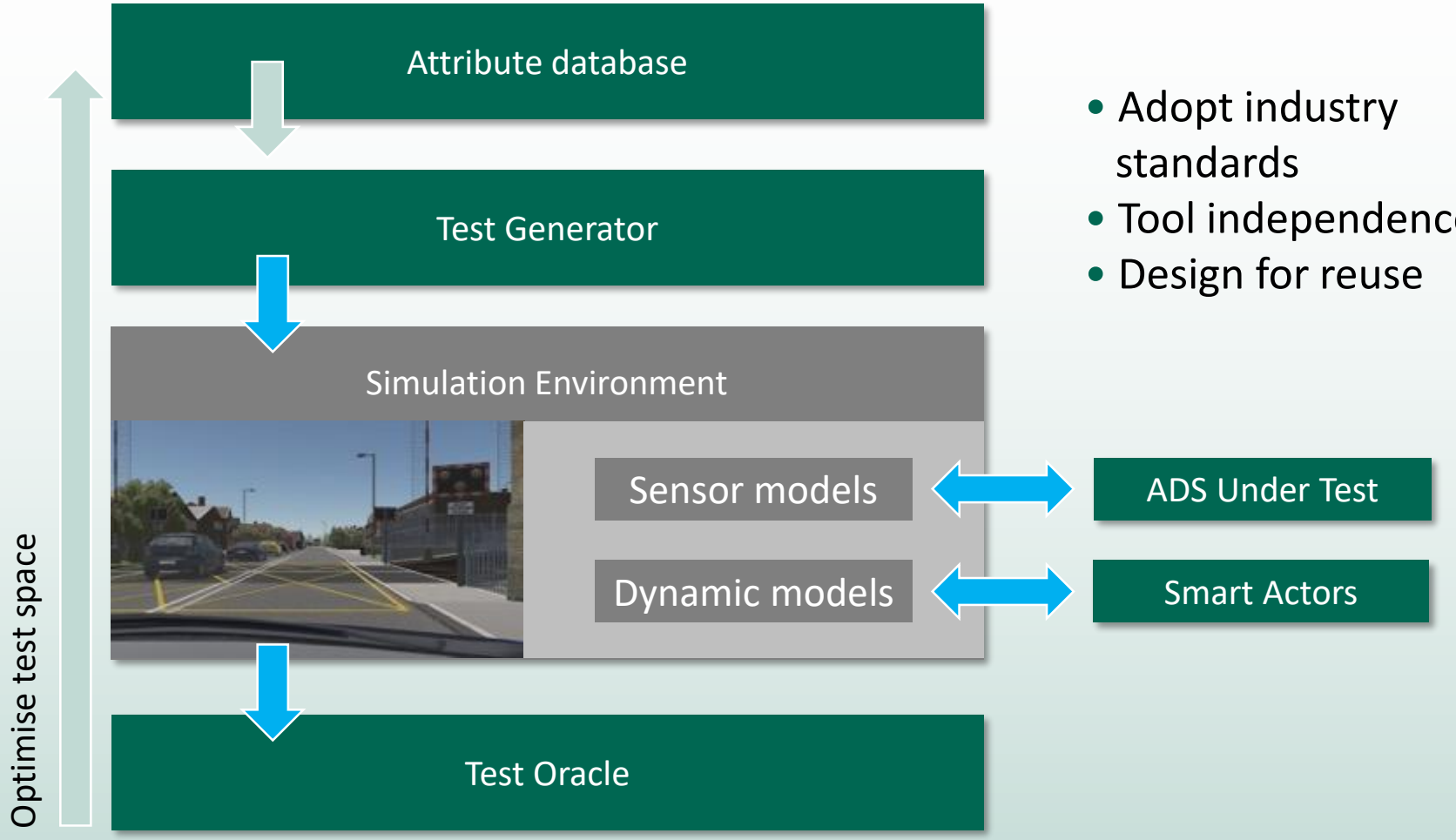
Smart actors

- Humans (drivers, pedestrians, and cyclists) make test cases challenging
- Vital to replicate real-world conditions in simulation tests

- Use computer vision to extract real-world behaviours from data e.g. traffic cameras
- Train “virtual humans” to imitate these behaviours using deep learning

- Create actor models from cognitive models of human decision-making
- Potential to combine AI and cognitive models

Modularisation + interfaces



- Adopt industry standards
- Tool independence
- Design for reuse

ADS Under Test

Smart Actors

Remaining challenges, and the future

CATAPULT
Connected Places

Remaining challenges

Technical – testing

- Finding “good” scenarios, ensuring test space coverage, fault injection
- Fidelity of AV simulation environments (sensor models, maps and 3D world models)

Technical – advanced methods

- Explainable AI and AI verification, modelling and formal methods

Certification

- What level of safety-relevant performance is acceptable for AVs
- Scenario-based test process, simulation vs. physical, and interfacing to the ADS
- Safety of semi-autonomous vehicles and handover
- Assuring safety by following systems engineering processes, and how to verify this
- Non-functional requirements such as component redundancy

Lifecycle

- Verifying software installed as an over-the-air update
- MOT-type testing, handling damage and dirt
- Continuous in-service performance monitoring, data recording
- Accident investigation, sharing of new safety-relevant scenarios

Collaboration is likely to be critical to making progress on these – between regulators, industry, and academia

International ecosystem of projects and initiatives is building to address aspects of these

CATAPULT Connected Places

Regulatory apparatus

UNECE WP.29

- WP.29 has a Working Party called GRVA, focused on AVs
- Key sub-group: Verification Methods for Automated Driving (VMAD)
- Addressing:
  - Closed-road tests
  - Real-world test drive
  - Audit and simulation

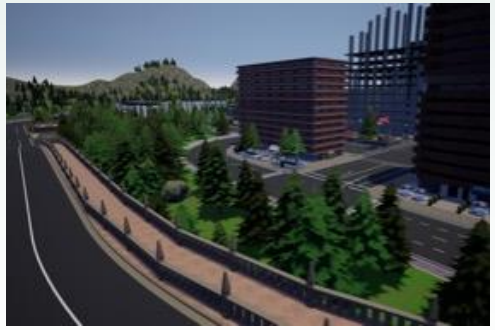
Other international work

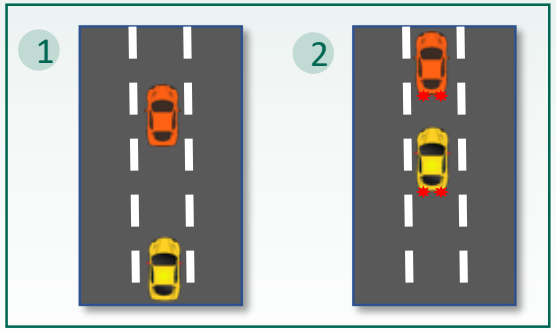
- United States – NHTSA, SAE, and UL 4600
- EU Commission – Joint Research Council (JRC)
- Singapore – CETRAN programme

CATAPULT
Connected Places

Path to a workable certification methodology

Near term





Longer term

- Improve coverage with existing test methods
  - Improved speed and fidelity in simulation tools, improved search optimisation, …
- Scale static verification methods to real-world systems
  - Requires cultural shift towards openness on the part of the OEMs
  - Advances in verification of AI systems, formal methods, model-based checking

CATAPULT
Connected Places

Next steps

Outlook

- Growing amount of activity in verification and certification aspects
  – UK is well represented
- Good potential for progress

Responsibility

- Ultimately with regulators
- However they will rely on advice and research from the whole community – industry, consultancies, and academia

Input

- CPC would be keen to work with academics doing relevant research

CATAPULT
Connected Places

Thank you for your attention

CATAPULT Connected Places

VeriCAV
VERIFICATION OF CONNECTED AND AUTONOMOUS VEHICLES

is supported by

Centre for Connected & Autonomous Vehicles

Innovate UK

MUSICC
Multi User Scenario Catalogue for Connected Autonomous Vehicles

is supported by

Department for Transport

zeyn.saigol@cp.catapult.org.uk