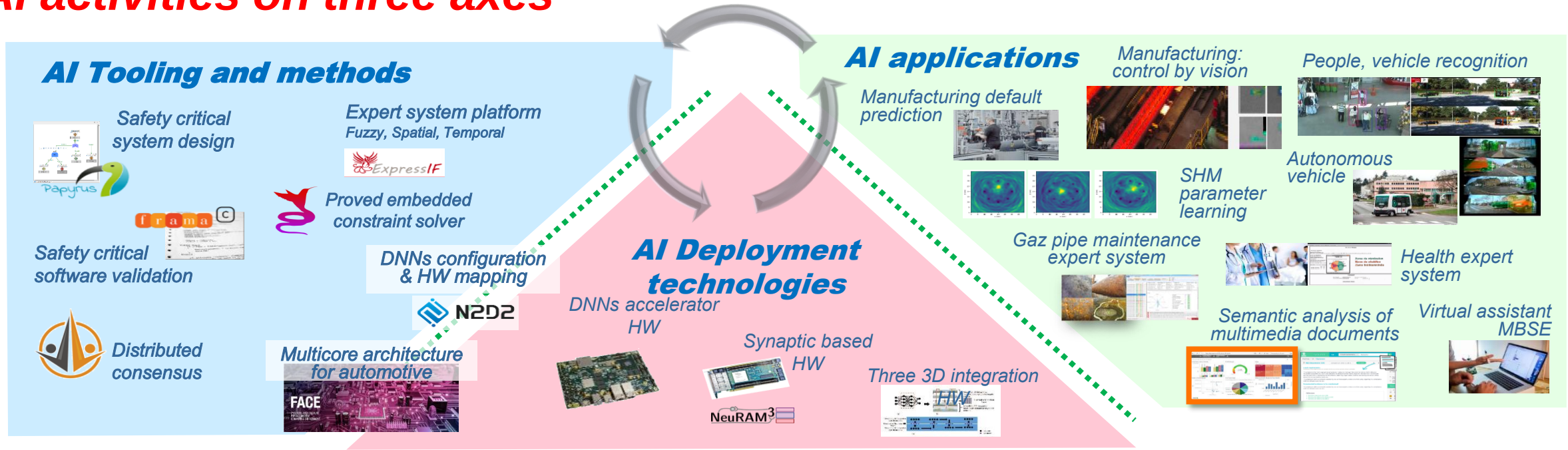
CONSIDERATIONS FOR EVOLUTIONARY QUALIFICATION OF SAFETY-CRITICAL SYSTEMS WITH AI-BASED COMPONENTS

Pr. François Terrier, Dr. Huascar Espinoza Ortiz, Dr. Morayo Adedjouma
Département d'Ingénierie des Logiciels et des Systèmes
Software and System Engineering Department

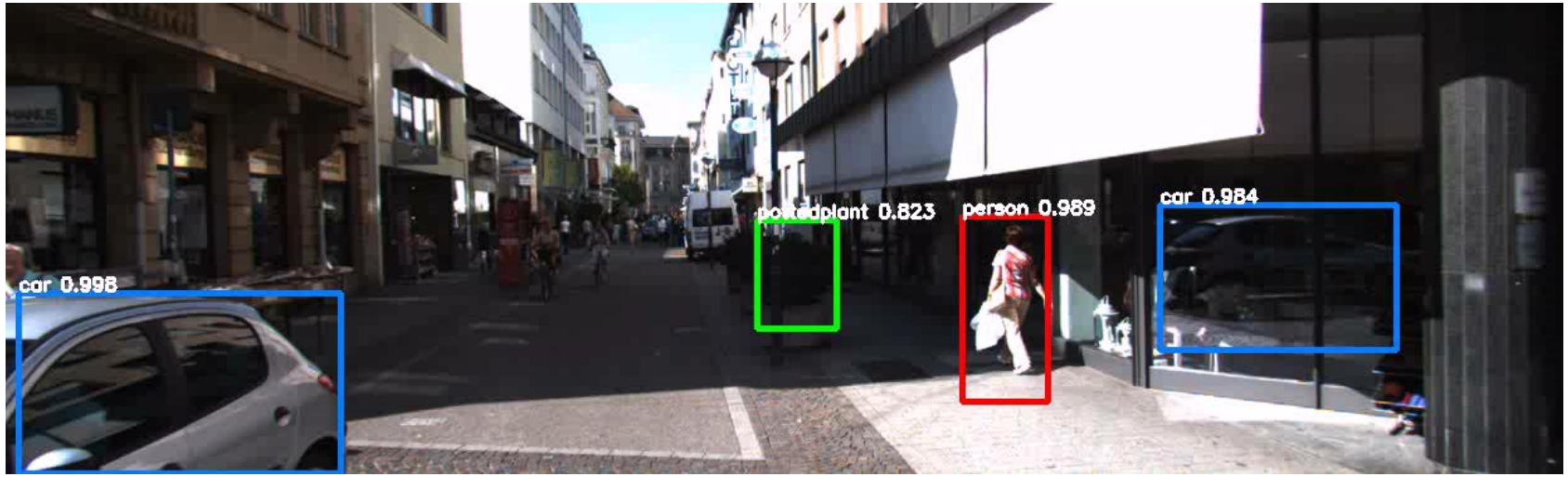
CEA's MAIN MISSIONS

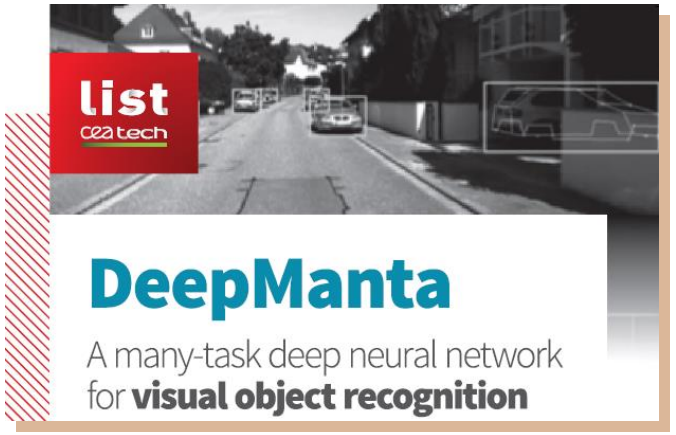
INSTITUTE FOR INTEGRATION OF SYSTEMS & TECHNOLOGIES

➔ AI activities on three axes



EXAMPLE OF DEEP LEARNING APPLICATION: REAL TIME VIDEO INTERPRETATION





Vehicle environment interactions understanding

Presented at CES 2018, 2019 on Drive4U stand of Valeo

The KITTI Vision Benchmark Suite

KIT Karlsruhe Institute of Technology

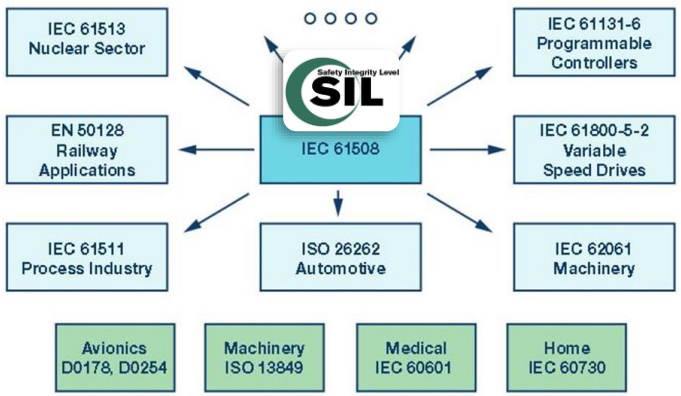
1st For detection and estimation of orientation and distance

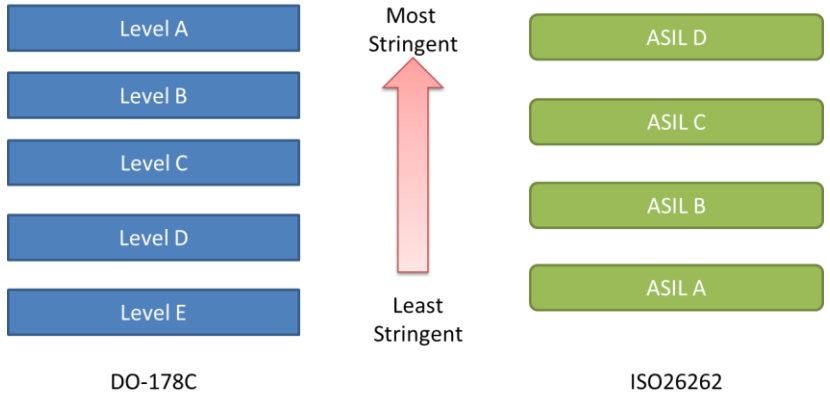
Performances:
+30% / State of the Art

Valeo

Certification/qualification of safety-critical systems with AI-based components

CURRENT QUALIFICATION PRACTICE FOR SAFETY-CRITICAL SYSTEMS





- **Based on (prescriptive) industrial Standards**
  - Conformance requirements
  - Prescribe a set of practices
  - Trace the decision, and assessment artefacts

  ➔ ***Safety integrity levels***

- **Assurance effort is proportional to function/system criticality/integrity levels**
  - Degree of risk wrt. criticality of failures
  - Process & techniques adapted to each level
  - Highest level applied to most severe failures

  ➔ ***Validation costs can add e.g. 30-150% (DO-178)***

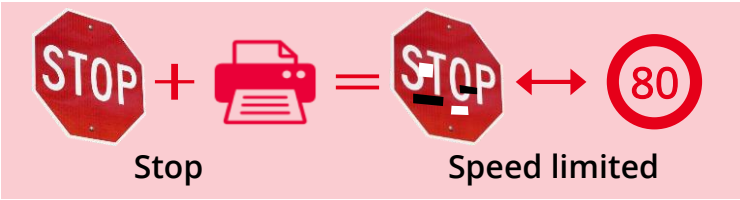
- **Incremental qualification is a core concern**
  - Integrated Modular Avionics (IMA) in DO-297
  - Safety Element out of Context (SEooC) in ISO 26262 (Automotive)
  - Generic Safety Case in EN 50129 (Railway)

  ➔ ***Modular architectures***

…SO, WHY WONDERING ABOUT AI TECHNOLOGY QUALIFICATION?

The technology comes with a short "maturity", with clear weakness on

Usage, specification, design, robustness, security…



… DEVELOPMENT PROCESSES ARE (still) NOT UNDER CONTROL!

Machine learning has become alchemy

Ali Rahimi
(Google)

Engineering artifacts have preceded the theoretical understanding

Yann LeCun
(facebook)

Formal, traced & rational approach
- Requirements
  - *Functions* ↳ *Sub-functions* ↳ *Actions*

Each instruction, each value is deduced and justified

vs

An empirical approach
- Informal requirements « by examples »
  - *trails and errors*

Poor justification, explanation of the result

Its true for both knowledge based AI and data based AI

With a very pregnant pressure on ML based AI

AI/ML qualification is still an open issue*

Unformal requirement with less / no structuration, dynamic evolution of system definition

Hard-to-scale-up operating conditions, Verification completion criteria: when are we done with testing?

Breaks all the conformity assessment principles and processes…?

… to a 3rd AI Winter?

**Some standardization bodies are working on AI, e.g. ISO/IEC JTC 1 Standards Committee on Artificial Intelligence (SC 42), EUROCAE WG-114, and the working groups of IEEE SA's AI standards series*

…SO, WHY TO CONSIDER EVOLUTIONARY AI QUALIFICATION SO EARLY!?

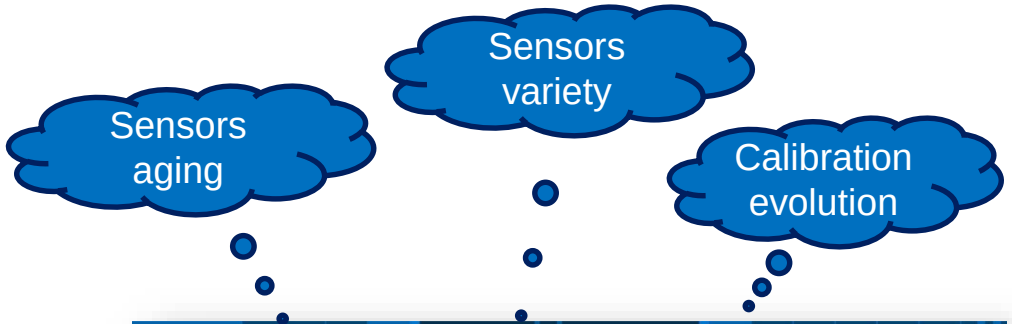
- **The frequency of changes is potentially large.**
  - AI-based systems are more influenced by obsolescence of data, system's operating environment, sensors…

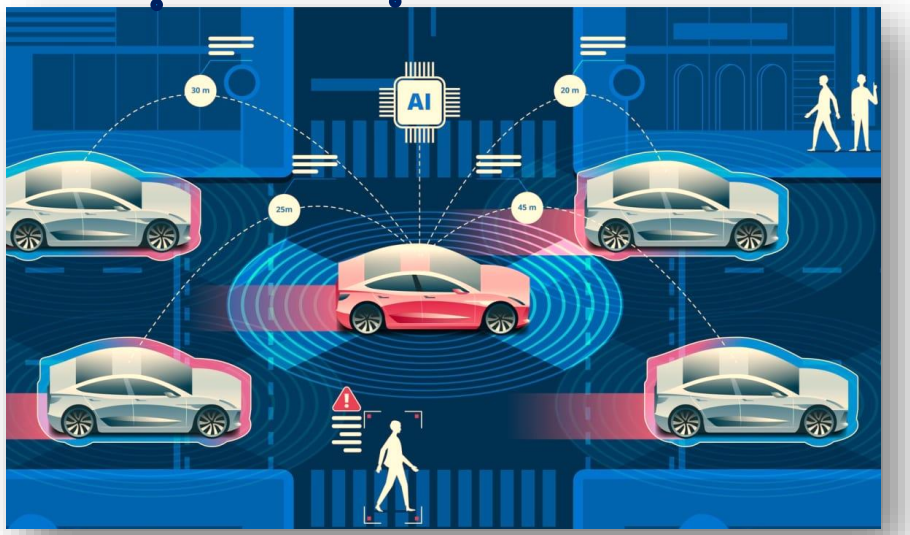
  ➔ …which leads to need of repetitive/continuous (re-)qualification processes.

- **The complexity of the validation process**
  - …the costs of revalidation, even for small changes are very high

    e.g. we could need re-training the system for slightest modification of a function

    (E.g. deep learning algorithms containing millions of parameters in close interaction)





Re-qualification is easier if the system has been designed with this objective…

EVOLUTIONARY AI* QUALIFICATION CHALLENGES

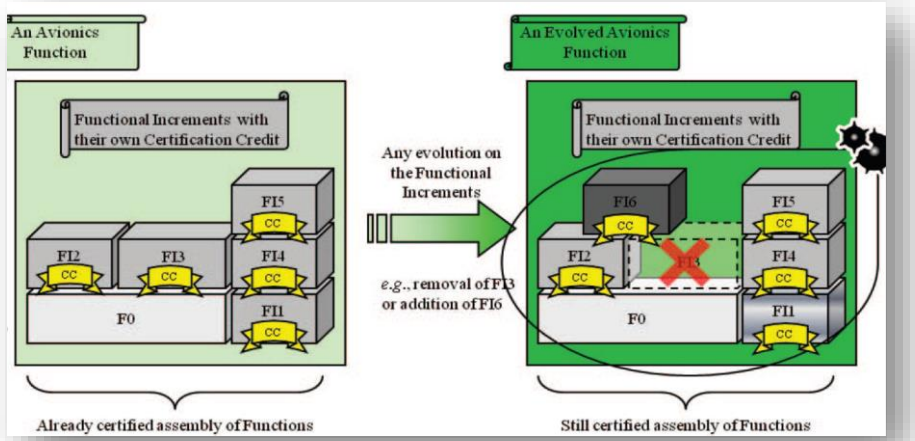
**Focus on ML*

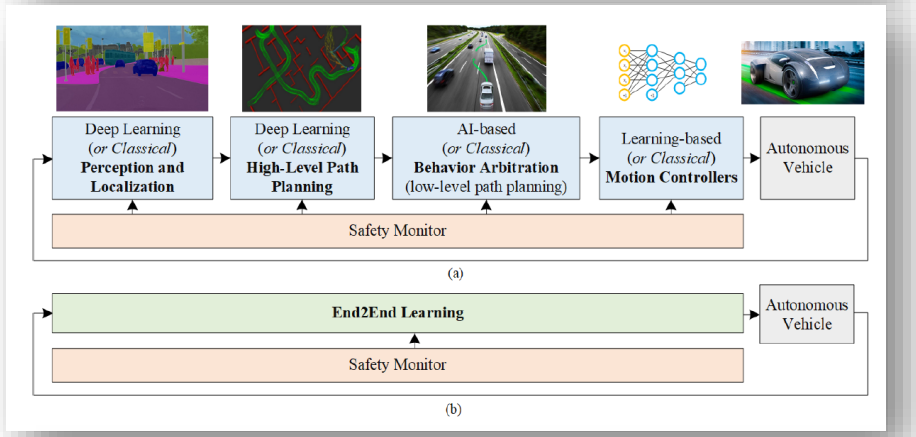
Need of Paradigm Change from Different Perspectives:

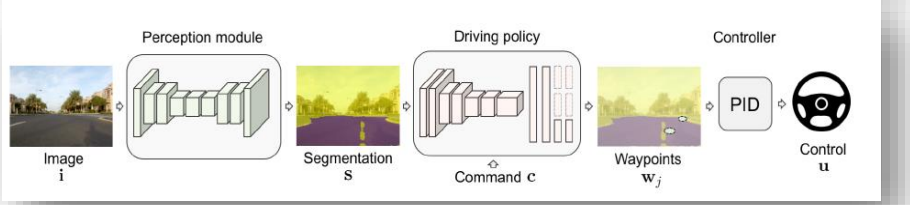
A- Modularity and Metrics for AI-based System Architectures

B- Evolution: a continuum between Development and Operational Phases

INCREMENTAL CERTIFICATION: BASED ON MODULARITY







- **Principle of Incremental Certification (e.g.: IMA)**
  - Functional component with a well defined perimeter of evolution and consistent with operational specs.
  - Functional Increment (addition or suppression)
  - Enable acquisition of qualification credits at component level

- **Current AI-Based Architectures**
  - Modular architecture with mixed AI/non-AI functions
    - Pros: easier to validate, optimize, deploy.
    - Cons: error accumulation, calibration is harder

    Self-Driving Cars: A Survey. C Badue, et al. Oct. 2019.

  - End-to-end ANN-based architectures

    End-to-end Driving via Conditional Imitation Learning". F Codevilla, et al. ICRA 2018.

    - Pros: optimal representation wrt to desired task
    - Cons: large amount of data, hard to scale, less explainability.

  - Combination Modular/End-to-end

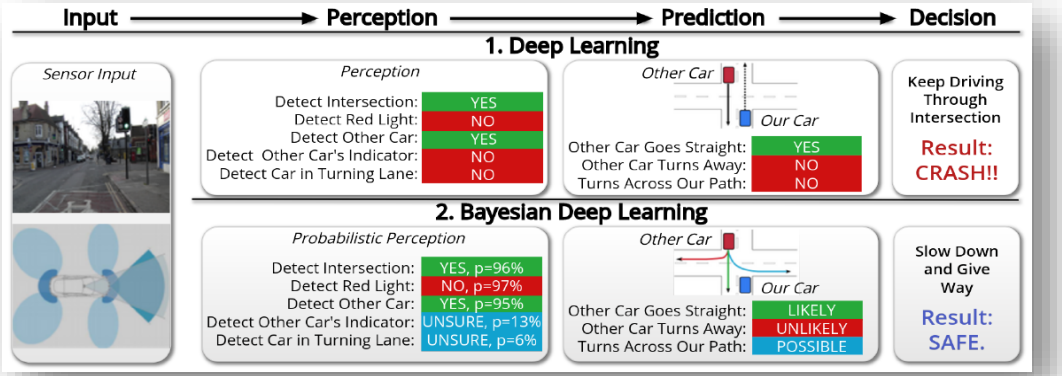
    Driving Policy Transfer via Modularity and Abstraction"., M Muller, A Dosovitsky, et al. 2018.

    - Encapsulate driving policy and transfer driving policies from simulation to reality via modularity and abstraction.

TRENDS TO HELP SAFETY MANAGEMENT THROUGH MODULARITY

Decomposition of safety-related properties

- Safety can be improved by quantifying component output uncertainties & propagating them forward through the pipeline.
- This improves interpretability by explaining what the different modules observes and why the whole system makes the decisions.

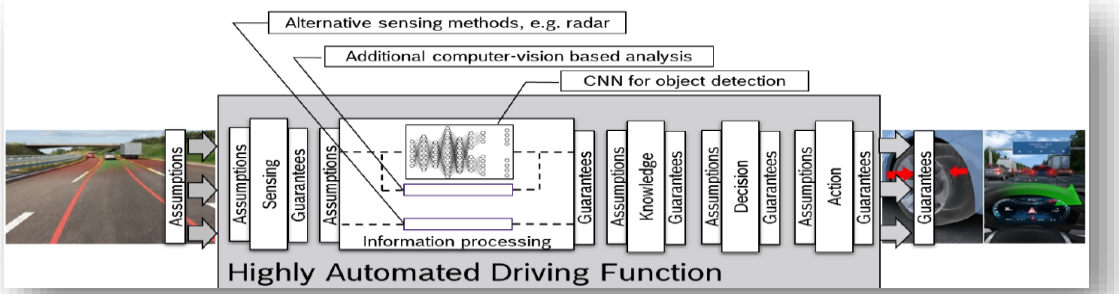


Concrete Problems for Autonomous Vehicle Safety-Advantages of Bayesian Deep Learning, McAllister,Ygal.. 2017

➔ ***Measure/Prediction of accuracy/uncertainties/probabilities is a key***

Contract based approaches

- Demonstrate each function/module meets its safety guarantees under all conditions where the assumptions hold.
- Particularly challenging is finding the boundaries (formal verification can help here!), measuring abnormal situations (including uncertainty) and managing global safety integrity.



Making the case for safety of machine learning in highly automated driving. In: International Conference on Computer Safety, Reliability, and Security. Burton, S., Gauerhof, L., Heinzemann, C. pp. 5–16. Springer (2017)

→ Modular approaches and more precise means to specify/measure safe behavior at component level, would help!

EVOLUTIONARY AI* QUALIFICATION CHALLENGES

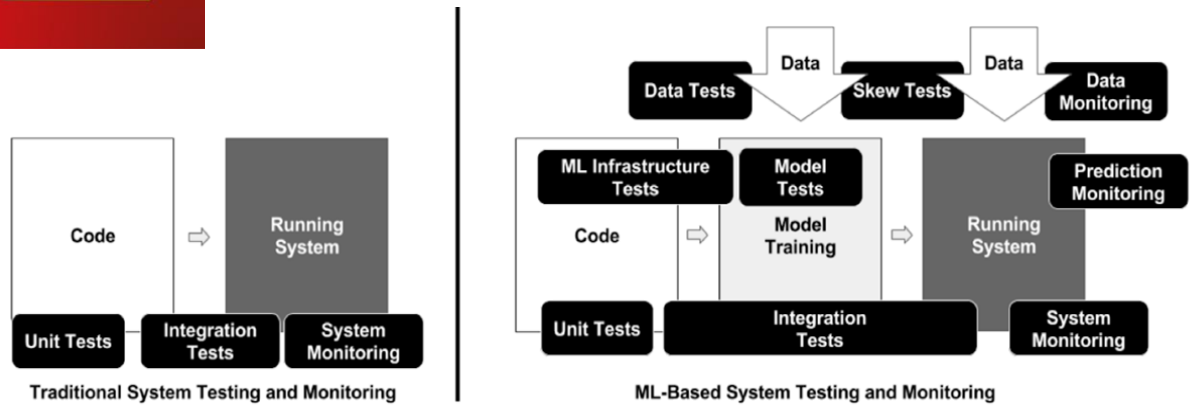
**Focus on ML*

Need of Paradigm Change from Different Perspectives:

A- Modularity and metrics for AI-based System Architecture

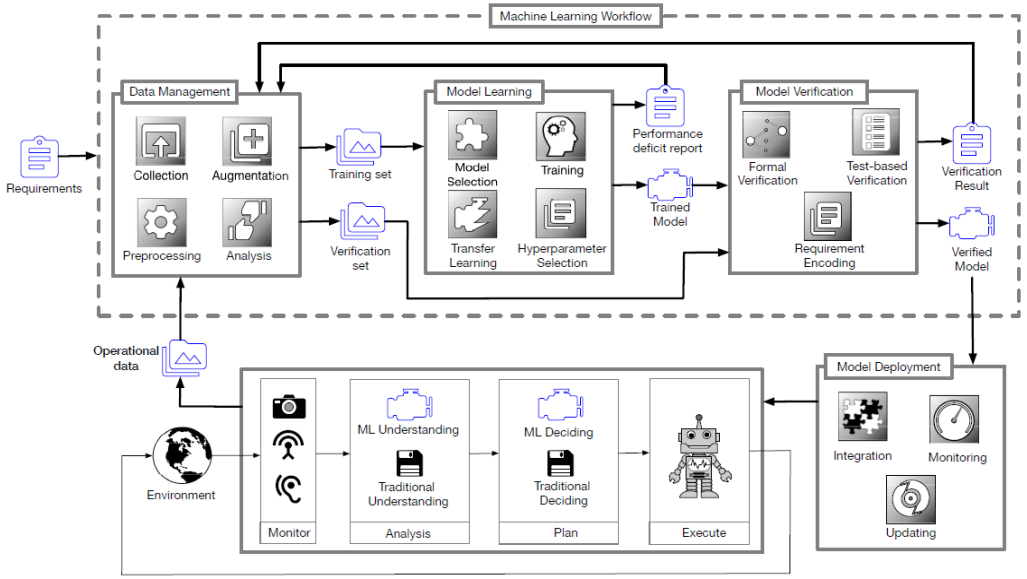
B- Evolution: a continuum between Development and Operational Phases

THE MACHINE LEARNING (DYNAMIC) LIFECYCLE



The ML Test Score: A Rubric for ML Production Readiness and Technical Debt Reduction. Eric Breck, Shanqing Cai, Eric Nielsen, Michael Salib, D. Sculley

- Traditional development processes carries out qualification activities as a separate stand-alone after-the-fact activity on final products.



Assuring the Machine Learning Lifecycle: Desiderata, Methods, and Challenges, Rob Ashmore, Radu Calinescu, Colin Paterson

→ Even in the conservative case where we **disable learning-based adaptation** before deployment we need to:

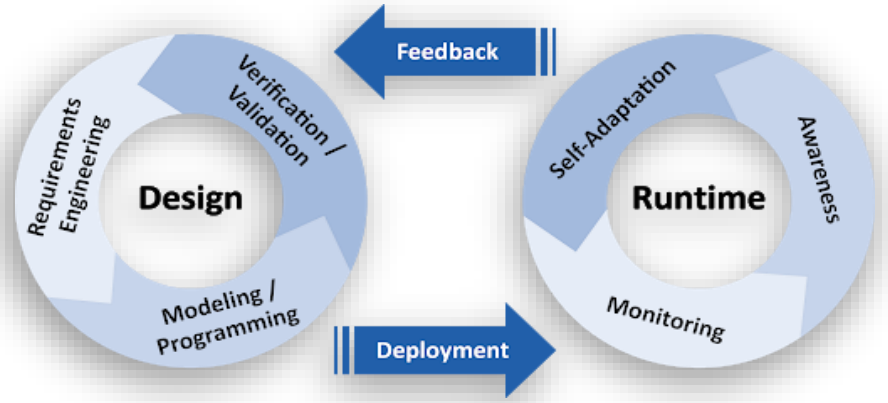
- Track breaking changes that trigger re-qualification process
- Monitor unacceptable operational conditions (unpredicted)
- Observe inconsistencies between training and operation ML statistical quality.

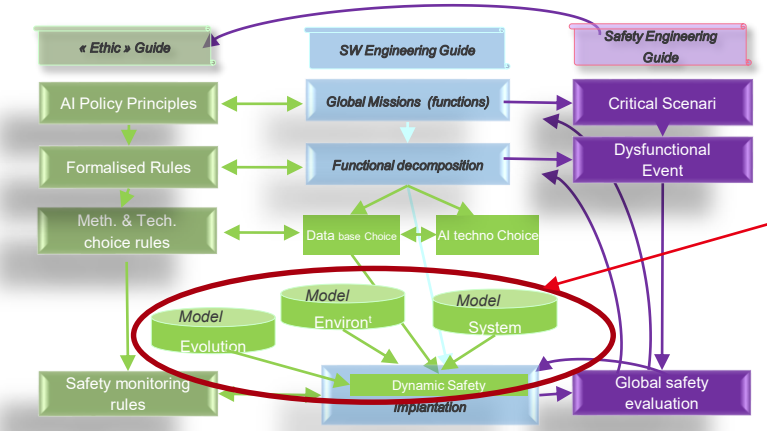
→ ML-based systems require pervasive monitoring!

NEED FOR EVOLUTIONARY PROCESSES FOR AI ASSURANCE

Towards an evolutionary AI-oriented lifecycle:

- Allow **each stage** of development and assurance **preserve** the **evidence chain** in a disciplined way and leaving **auditable records**.
- This needs an explicit specification of the assurance and qualification process, as well as the management of **specific metrics**.
- **Incremental construction, systematic reviews**.





Safe-by-Design Development Method for Artificial Intelligent Based Systems, G. Pedroza, M. Adedjouma, SEKE 2019, June Portugal

Need to embed evolution models in the architecture

- Containing specific principles triggering re-qualification needs.
- Metrics to assess and filter new stimuli and situations (e.g. unpredicted environment conditions).
- Mechanisms to integrate new filtered knowledge so as to grow up system and environment models.

→ We need more integrated development-operation processes and system evolution records to warn qualification stakeholders!

EVOLUTIONARY AI* QUALIFICATION CHALLENGES

**Focus on ML*

Need of Paradigm Change from Different Perspectives:

A- Modularity and metrics for AI-based System Architecture

B- Evolution: a continuum between Development and Operational Phases

C- Dynamic assurance case

→ metrics definition and process support to build the arguments to justify confidence to the stakeholders (authorities, regulators…) that the system is enough safe, dependable, performant, etc. for the purpose it has been built.

EVOLUTIONARY AI* QUALIFICATION CHALLENGES

**Focus on ML*

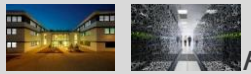
And now?

A NEW TOP LEVEL CEA' STRATEGIC PROGRAM for AI TRUST

TOOL & TECHNOLOGY platform for R&D on CERTIFIED AI

« Happy » AI developers

Building



Factory - IA

Understand

Deploy

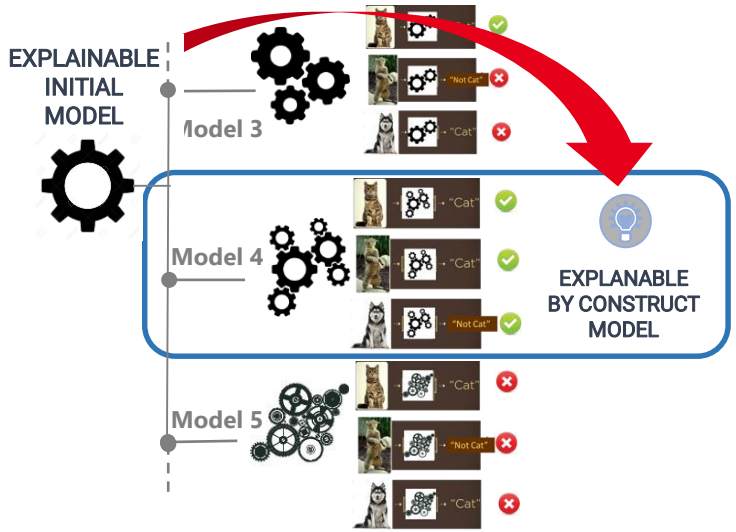
Certify

Usages

« Happy » AI users

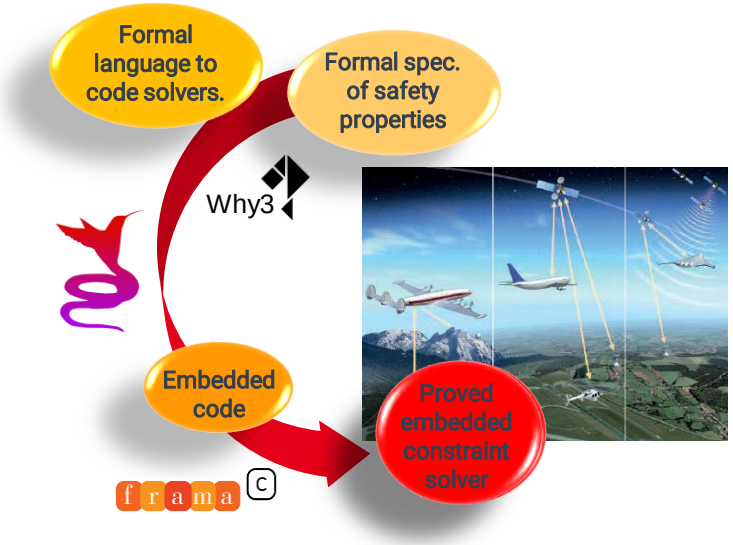
Ongoing R&D on certification, formal validation, uncertainty measurement, monitoring…

Reinvent learning: Machine discovering

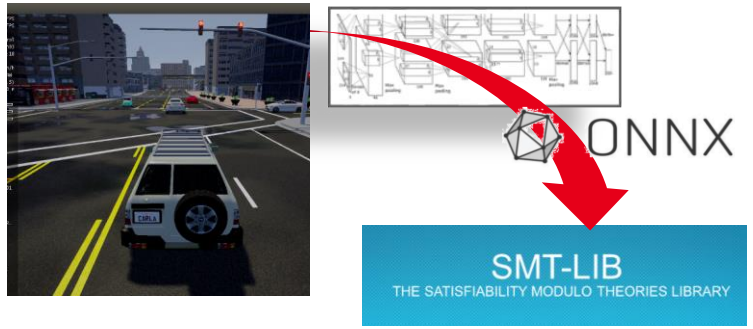


M. E. A. Seddik, M. Tamaazousti and R. Couillet "Kernel Random Matrices of Large Concentrated Data: The Example of GAN-generated Images." ICASSP'19

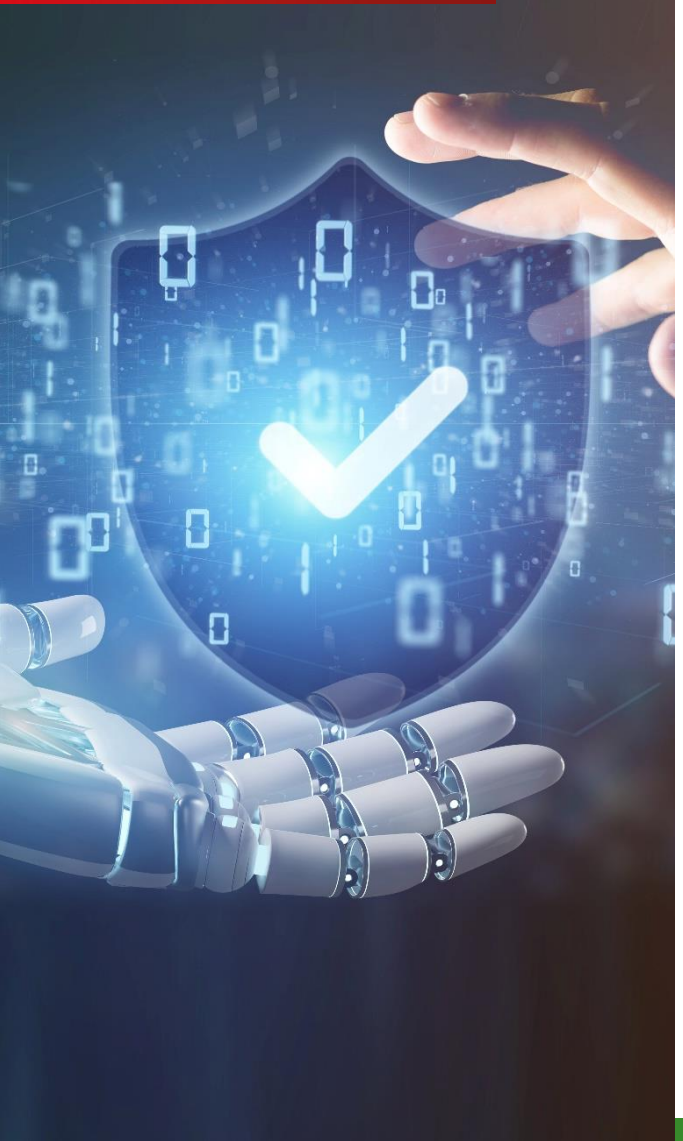
Proved constraint solver



Formal spec. & verification of DNNs



CAMUS: A Framework to Build Formal Specifications for Deep Perception Systems Using Simulators. J. Girard-Satabin, G. Charpiat, Z. Chihani, M. Schoenauer, ECAI 2020



« TRUST »

will make the difference

(Quality, Safety, Reliability, Ethic, Responsibility…)

1.RTCA DO-297 - Integrated Modular Avionics (IMA) Development Guidance and Certification Considerations.
2.ISO 26262-1:2018, Road vehicles — Functional safety
3.EN 50129:2003. Railway applications – Communication, signalling and processing systems – Safety related electronic systems for signalling
4.The ML Test Score: A Rubric for ML Production Readiness and Technical Debt Reduction. Eric Breck, Shanqing Cai, Eric Nielsen, Michael Salib, D. Sculley
5.Assuring the Machine Learning Lifecycle: Desiderata, Methods, and Challenges, Rob Ashmore, Radu Calinescu, Colin Paterson
6.“Driving Policy Transfer via Modularity and Abstraction”., M Muller, A Dosovitsky, et al. 2018.
7.“Concrete Problems for Autonomous Vehicle Safety-Advantages of Bayesian Deep Learning”, McAllister,Ygal.. 2017
8.Making the case for safety of machine learning in highly automated driving. In: International Conference on Computer Safety, Reliability, and Security. Burton, S., Gauerhof, L., Heinzemann, C. pp. 5–16. Springer (2017)
9.“Safe-by-Design Development Method for Artificial Intelligent Based Systems”. G. Pedroza, M. Adedjouma, Int. Conf. on Software Engineering and Knowledge Engineering, SEKE 2019, June Portugal
10.“Self-Driving Cars: A Survey”. C Badue, R Guidolini, et al. Oct. 2019.
11.“End-to-end Driving via Conditional Imitation Learning”. F Codevilla, M Muller, A Lopez, et al. ICRA 2018. Mar. 2018
12.“Uncertainty Quantification with Statistical Guarantees in End-to-End Autonomous Driving”. R. Michelmore, M. Wicker, et. al. Sept 2019.
13.“CAMUS: A Framework to Build Formal Specifications for Deep Perception Systems Using Simulators”. J. Girard-Satabin, G. Charpiat, Z. Chihani, M. Schoenauer, ECAI 2020
14.“Kernel Random Matrices of Large Concentrated Data: The Example of GAN-generated Images”. M. E. A. Seddik, M. Tamaazousti and R. Couillet , ICASSP’19