Thinking the certification process of embedded ML-based aeronautical components

Filipo Studzinski Perotto

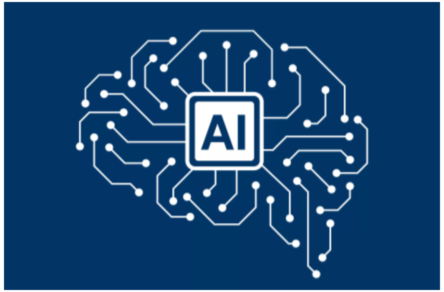
Journée commune du GDR RADIA et du GT IE du GDR GPL
21/11/2024

ML and Aeronautical Certification

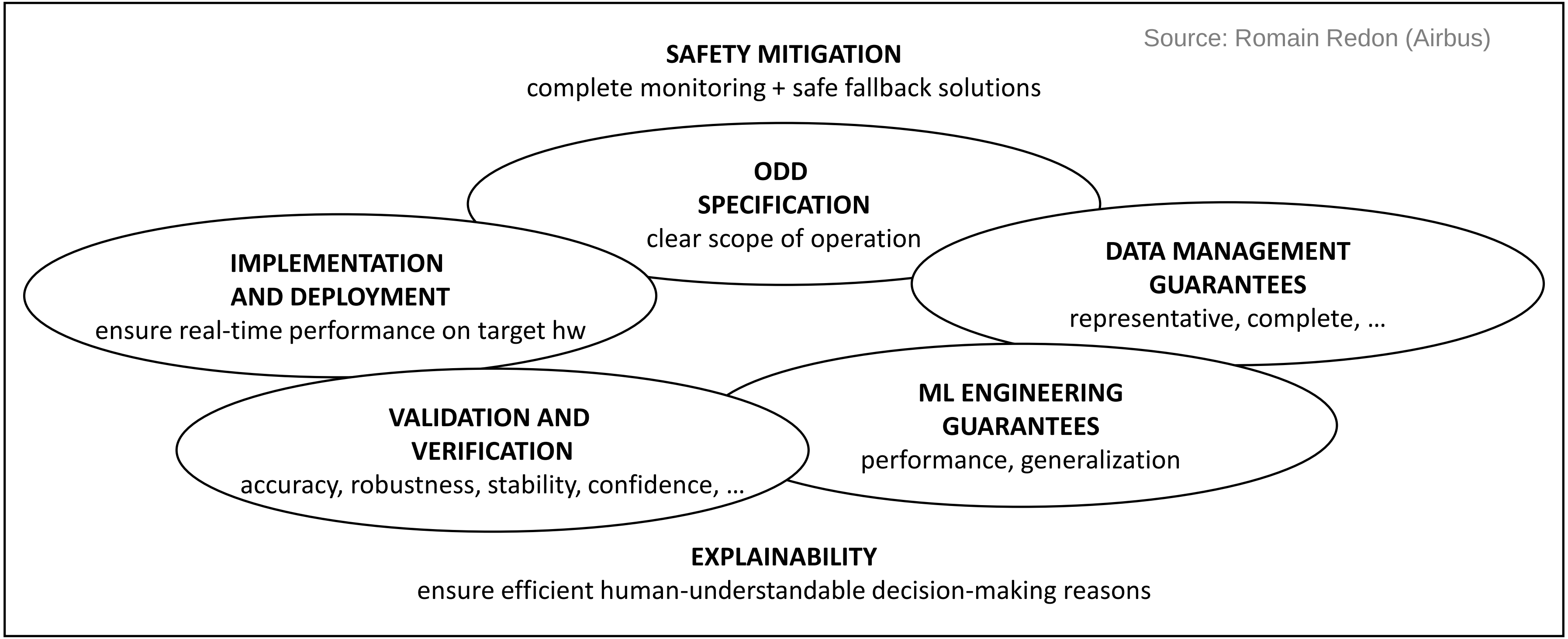
- **Deep Learning / ML (AI) Revolution**
  - >10 years of impressive results on diverse applications
  - Looks like magic !
  - Aeronautical industry is eager for using it

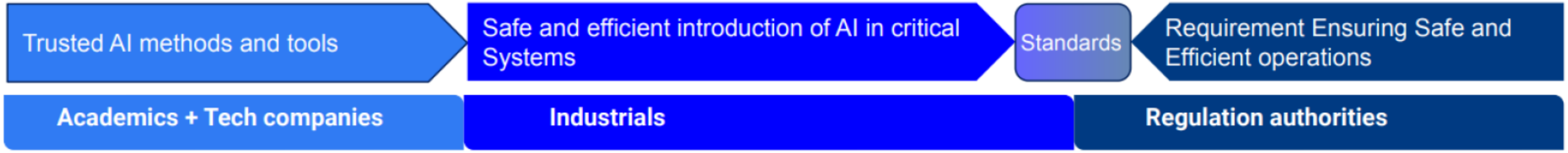
- **Aeronautical (Critical-)Systems**
  - Must be safe, robust, secure, trustworthy
  - Are constrained by strict regulations
  - Software and development process must be certified

- **Are those solutions certifiable ?**
  - Not for now…
  - ML development procedures and tools must be enhanced
  - Certification standards need to be adapted



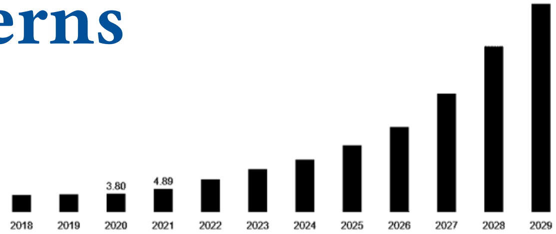
Learning Assurance for Embedded AI





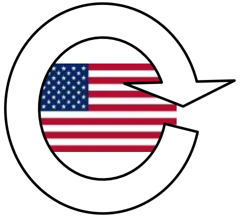
AI: Strategical and Economical Concerns

- **Market Size :**
  - Cloud AI : today ~ US$ 150 billion in 10 years ~ x10
  - Edge AI : today ~ US$ 10 billion in 10 years ~ x5 (x20 ?)



PyTorch TensorFlow Keras

American AI Platforms
(even if open source)



American HW targets

DEEPGREEN

FRANCE 2030

aidge

- **French AI Platform**
  - sovereign, independent, and open
  - competitive performances
  - allowing interoperability
  - adapted backends to French HW components
  - focus on frugality (data, energy, memory)

DeepGreen Project

aidge **Embedded AI Deep Learning Platform**

OPEN & INTEGRATED

Adaptable and reusable tooling to foster innovations. Hosted by Eclipse Foundation.

FRUGAL & EFFICENT

Greener hardware architectures combined to state of the art optimizations

RELIABLE & ROBUST

Compatible with critical functions and regulations

DEEPGREEN

Durée

4 ans

2023-2027

Membres

18

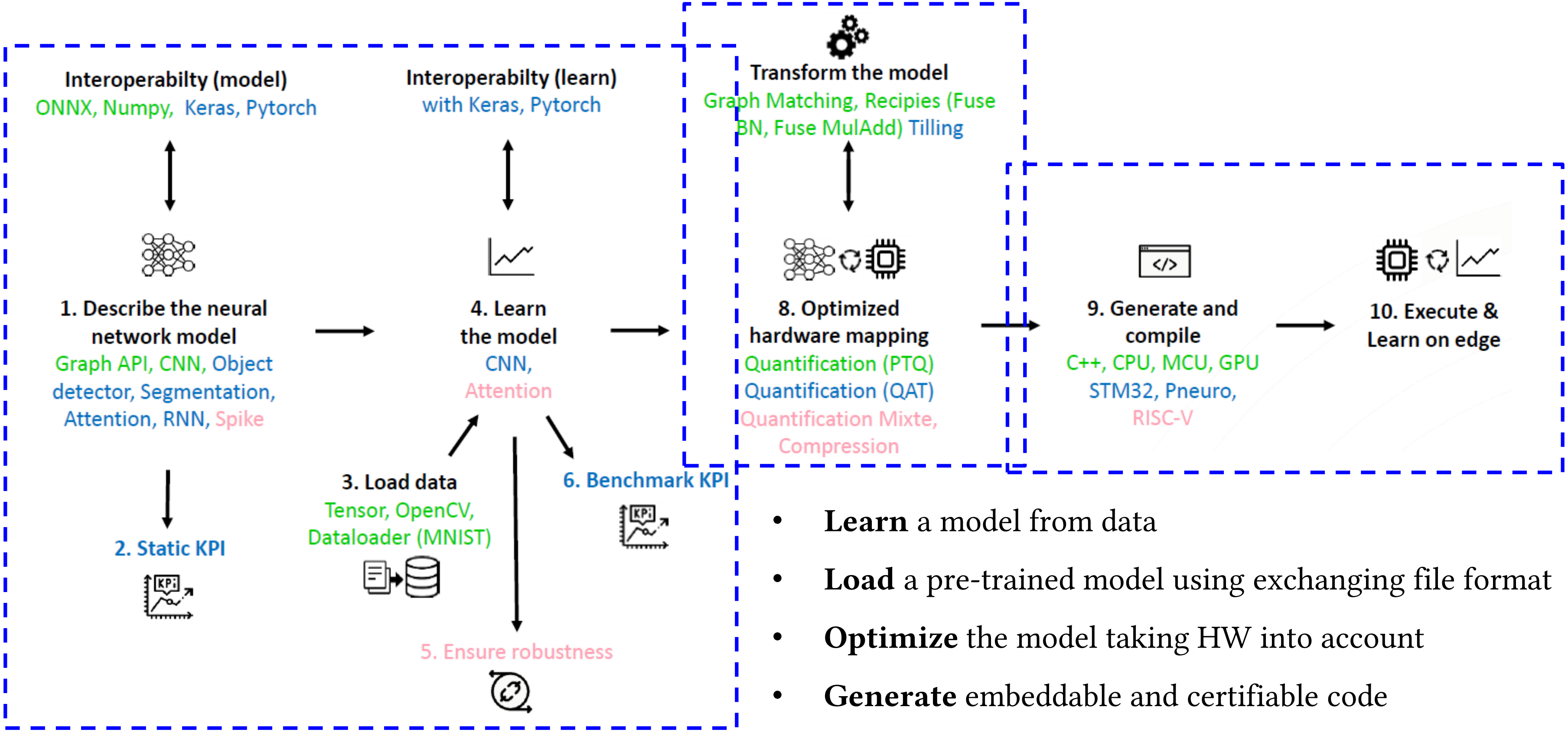
Une plateforme bâtie sur une triple expertise IA, compilation et systèmes embarqués

Des fournisseurs de solutions matérielles français

Des industriels représentatifs des secteurs clefs de l'IA embarquée

Défense | Aérospatial | Usine du futur | Energie | Santé

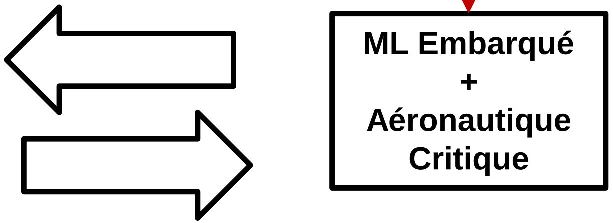
AIDGE Workflow



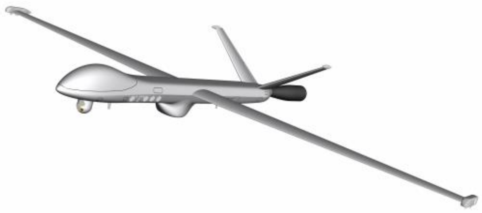
- **Learn** a model from data
- **Load** a pre-trained model using exchanging file format
- **Optimize** the model taking HW into account
- **Generate** embeddable and certifiable code

DeepGreen – Workpackage Aéronautique

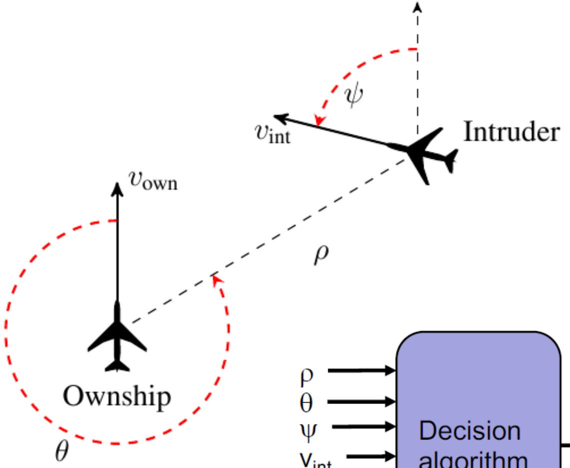
- **Evaluation de la plateforme AIDGE et recommandations concernant :**
  - les contraintes **d'embarquabilité** aéronautiques
  - **la certificabilité DAL C**
- **2 cas d'usage dans une première phase : ACAS-Xu et VBL**



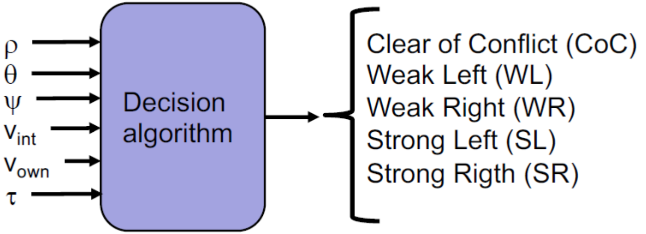
Use Case 1 : ACAS-Xu

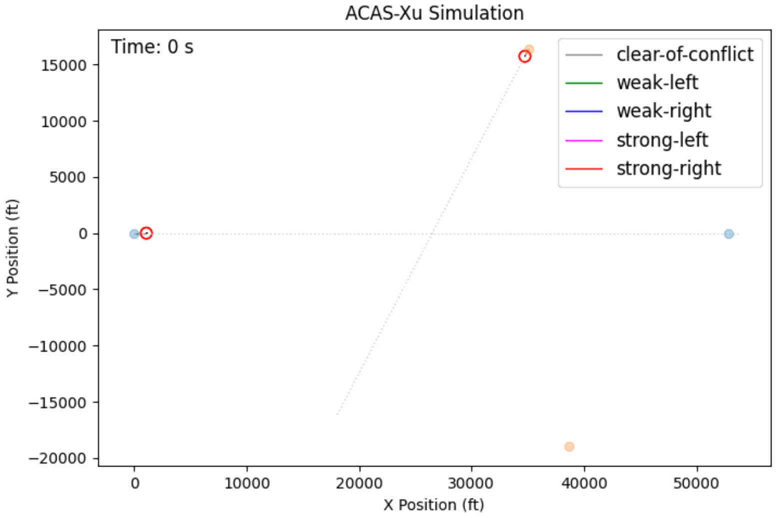


- ACAS-Xu :
  - Next Generation Airborne Collision Avoidance System for UAVs / RPASs
  - 7 discretized dimensions : $\rho, \theta, \psi, v_{own}, v_{int}, a_{prev}, \tau$
  - Q table : (10 x 5 x 12 x 12 x 41 x 41 x 39) x (5) = ~ 2 billion parameters

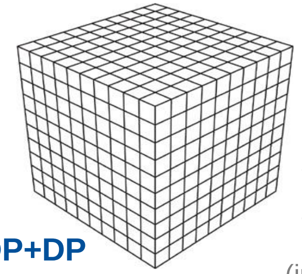


$a \in \{0^o, \pm 1.5^o/s, \pm 3^o/s\}$





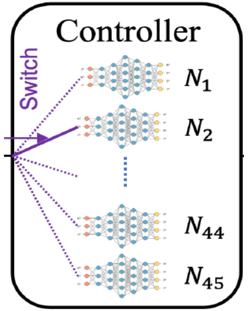
Use Case 1 : ACAS-Xu



MDP+DP
=
LUT

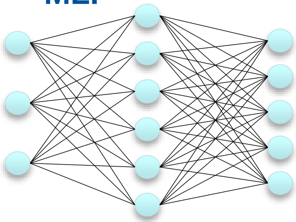
7 dimensions,
2.36 milliards
de q-valeurs
(int16 → 4.4 Go)

découpage en multiples réseaux



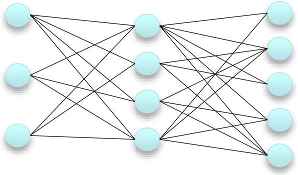
réduction d'empreinte mémoire (1/1000)

MLP

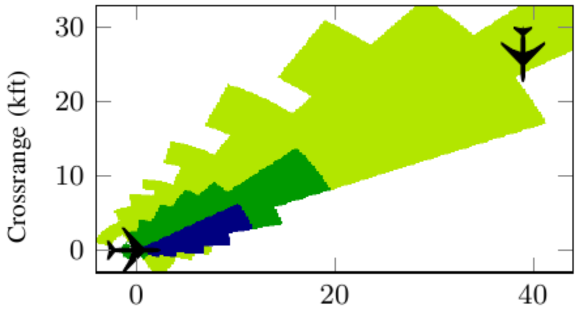


6 layers FC+ReLU
50 neurons per layer

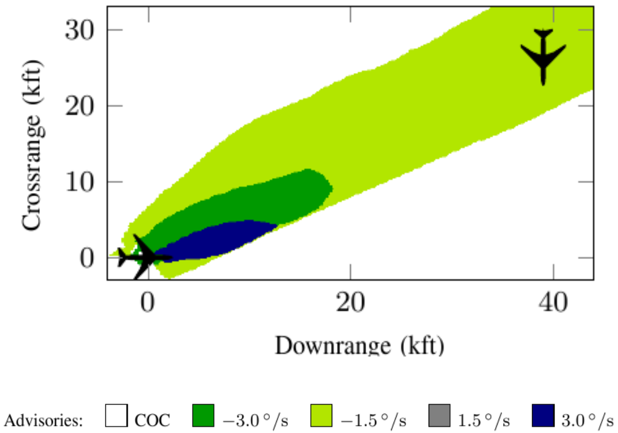
MLP Optimisé



Nearest neighbor interpolation of MDP table

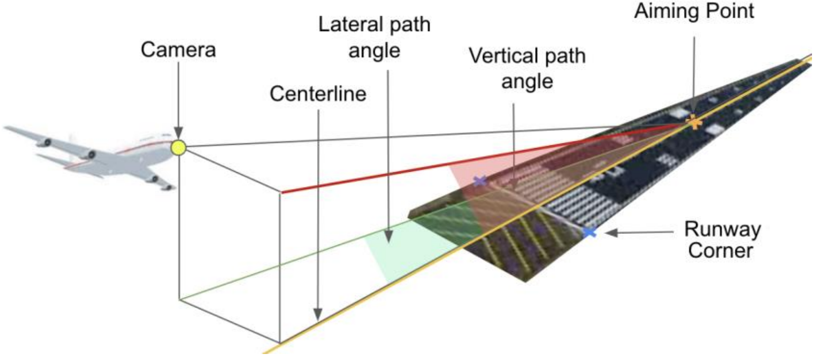


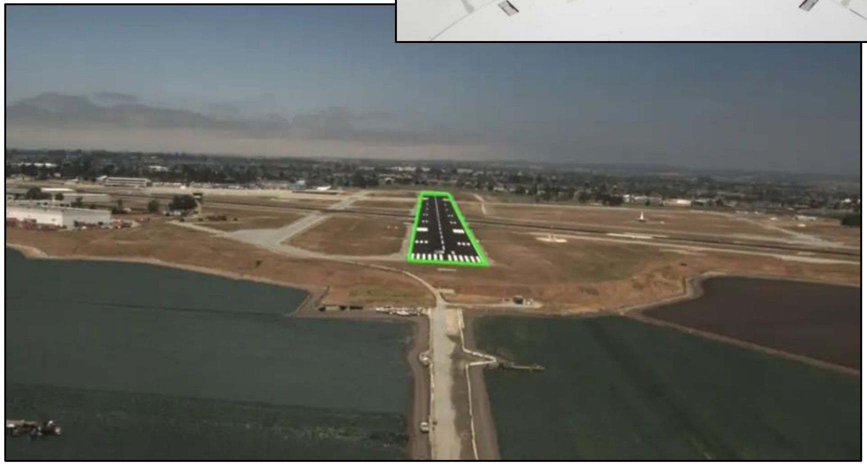
Neural Network



Use Case 2 : Visual Based Runway Detection

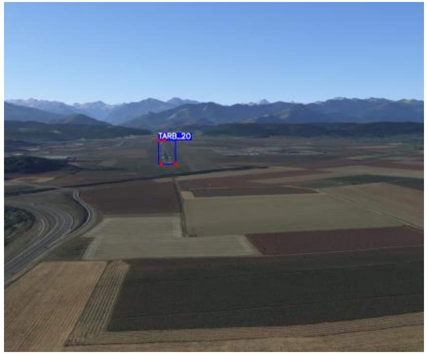
- Détection de la piste d'atterrissage dans des images issues d'une caméra frontal
- Architectures convolutives envisagées: Yolo-v5, Yolo-v8, LeYolo
- Entrainement, optimisation, compression, évaluation
- Vrais enjeux d'embarquabilité et certification

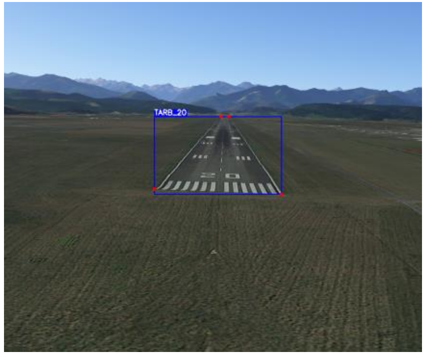


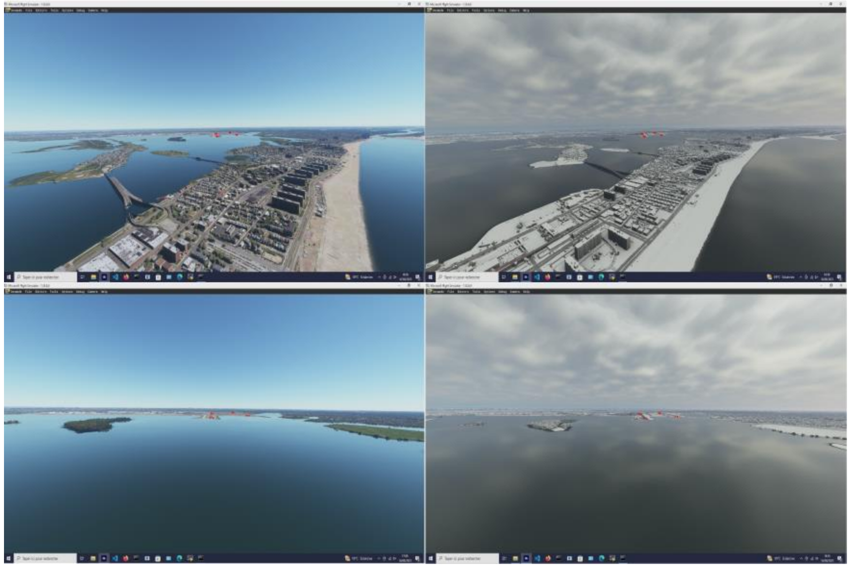


- **Bases d'Images: LARD et BARS**
- **Simulateur: Schemin**

Use Case 2 : VBL - VBRD







Development Assurance Levels

Critical Systems

DAL	Failure Effect	Condition	Risk
A	**Catastrophic**	Continuing the flight, takeoff or landing safely is impossible.	Several fatalities, maybe airplane crash.
B	**Hazardous**	Significant reduction in safety margins or functionality, with considerable increase in crew load.	Serious injuries, maybe fatalities.
C	**Major**	Reduction in safety margins or functionality, and increase in crew load.	Discomfort to the occupants, maybe injuries.
D	**Minor**	Small reduction in safety margins, or light increase in crew load.	Some inconvenience to passengers
E	**Insignificant**	No effect on aircraft operational capability or pilot workload, so no special requirements are imposed.	No particular risks.

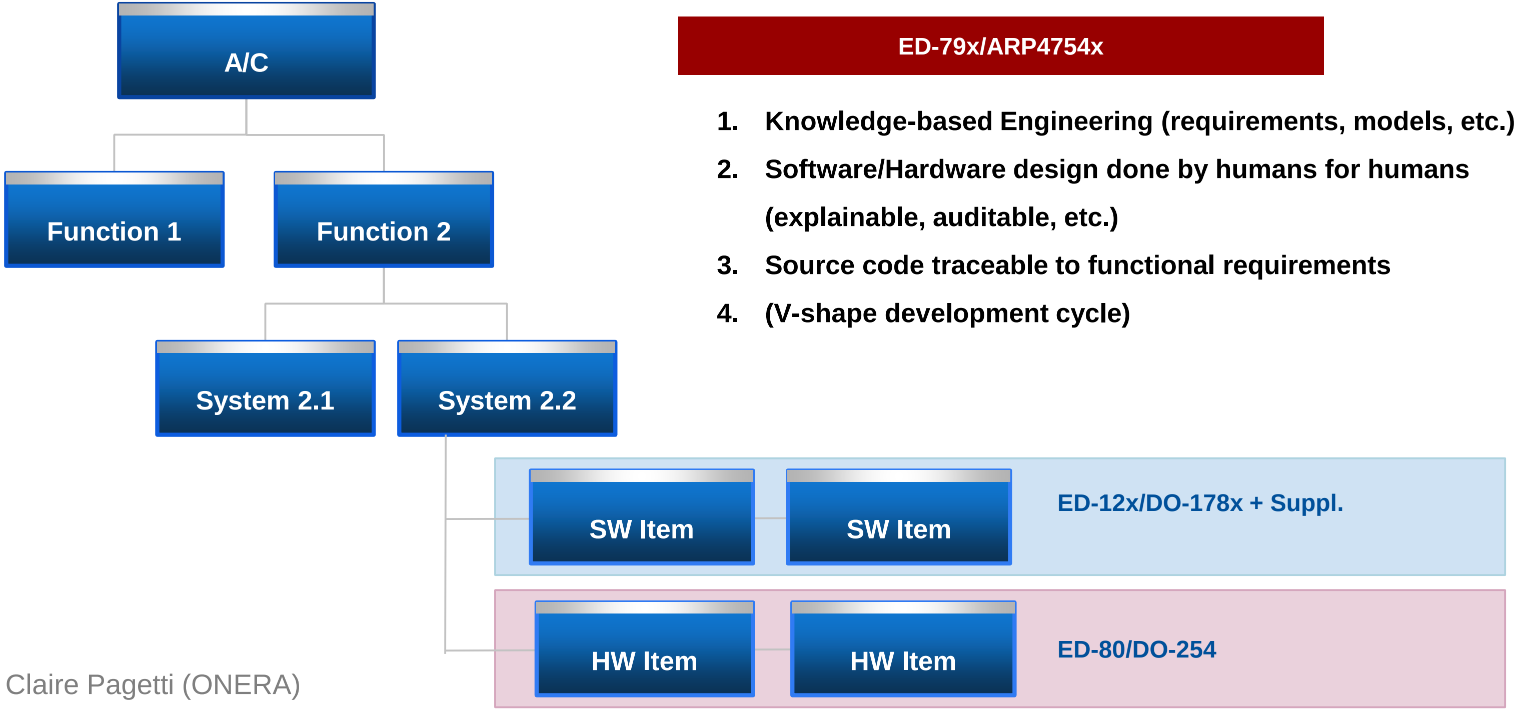
EUROCAE

DO-178C

The more the system is critical…
… Stronger are the certification requirements

Certification standards define specific objectives depending on DAL

Current Certification Approach in Aviation (Airborne)

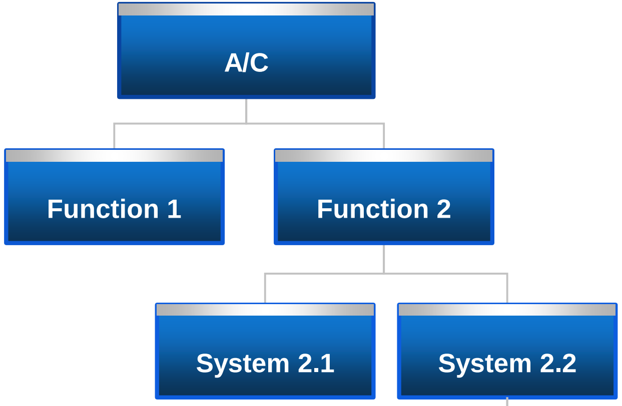


ED-79x/ARP4754x

1. **Knowledge-based Engineering (requirements, models, etc.)**
2. **Software/Hardware design done by humans for humans (explainable, auditable, etc.)**
3. **Source code traceable to functional requirements**
4. **(V-shape development cycle)**

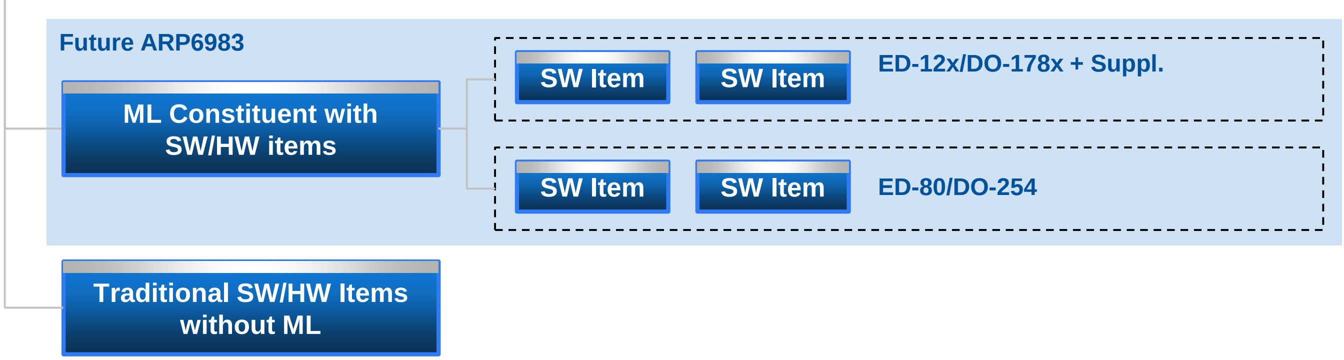
Source: Claire Pagetti (ONERA)

Adapted Certification Approach – in construction



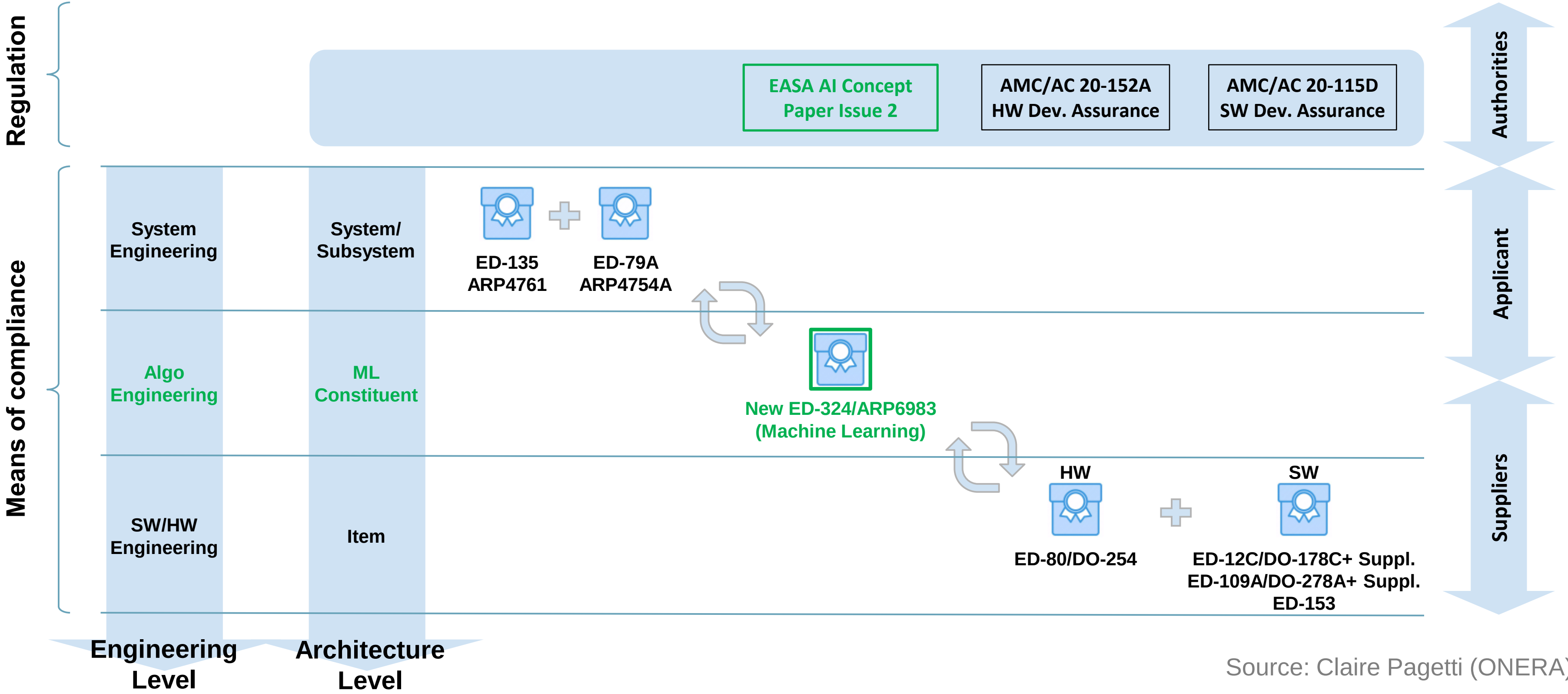
ARP4754x

1. Knowledge-based Engineering (requirements, models, etc.) → Data-Based (driven) engineering
2. Software/Hardware design done by humans for humans (explainable, auditable, etc.) → The complex design of a ML model is not directly understandable by humans
3. Source code traceable to functional requirements → The low level design developped by the machine is not directly traceable to functional needs



Source: Claire Pagetti (ONERA)

Future certification framework (Airborne view)



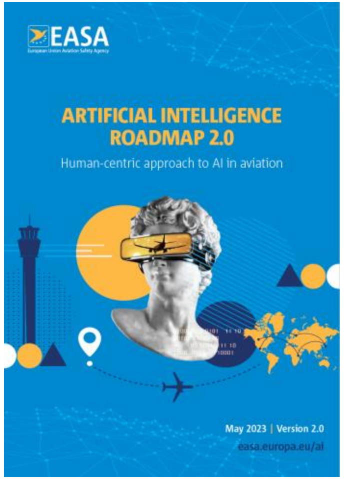
Source: Claire Pagetti (ONERA)

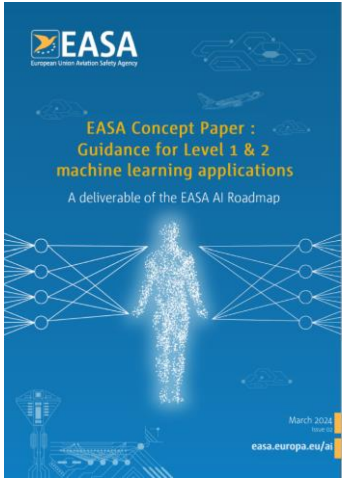
Challenges of a New Paradigm

Data-driven statistical learning:
- Data guarantees
- Learning guarantees
- Performance / Accuracy guarantees
- Robustness guarantees

CPU/GPU embedded implementation:
- Implementation guarantees
- Semantic preservation guarantees
- Runtime performance guarantees

AI Roadmap and Concept Paper





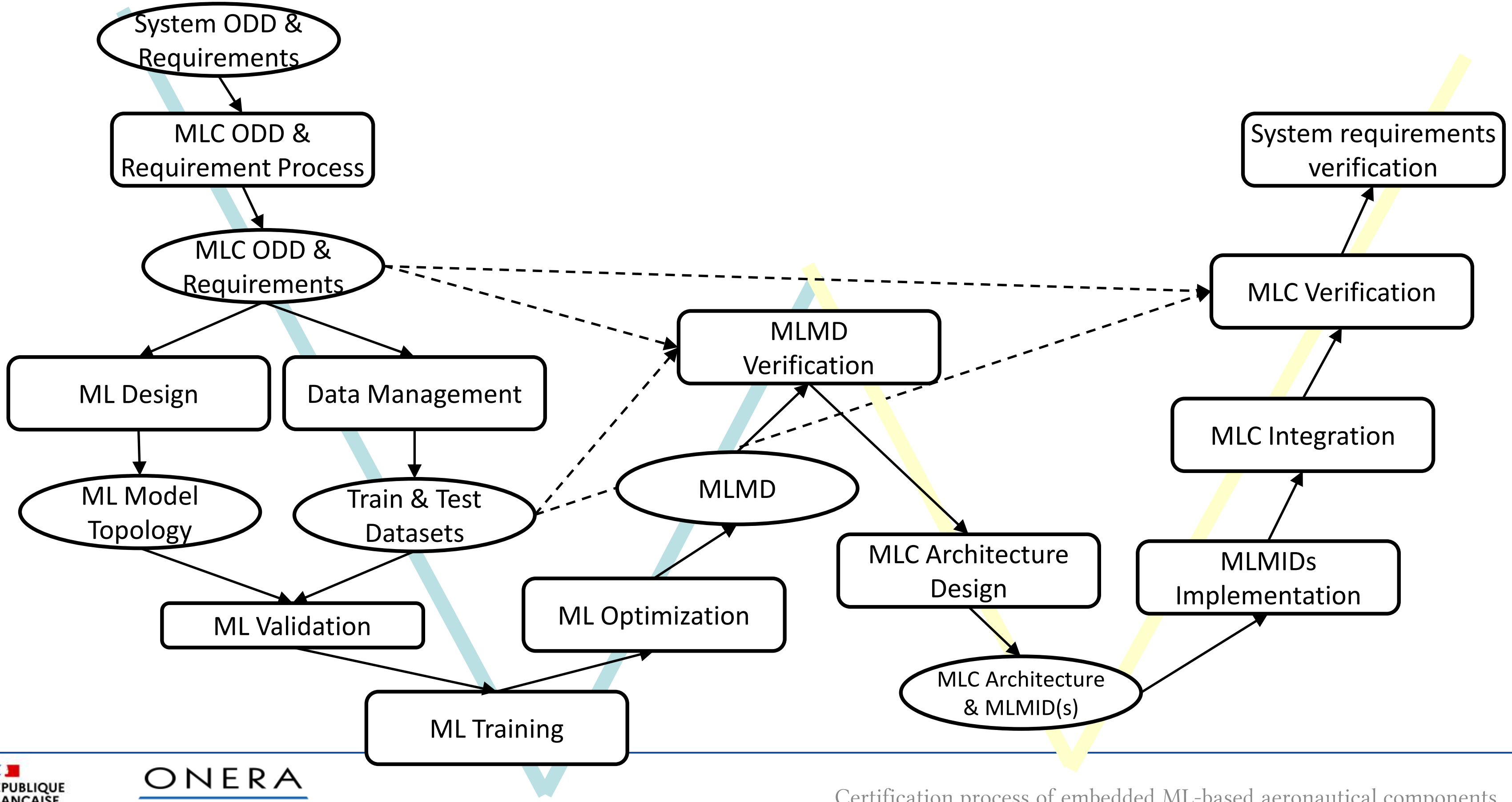
Aeronautic Industry

EUROCAE

SAE INTERNATIONAL

New ED-324/ARP6983 (Machine Learning)

Développement et Certification ML – schéma W



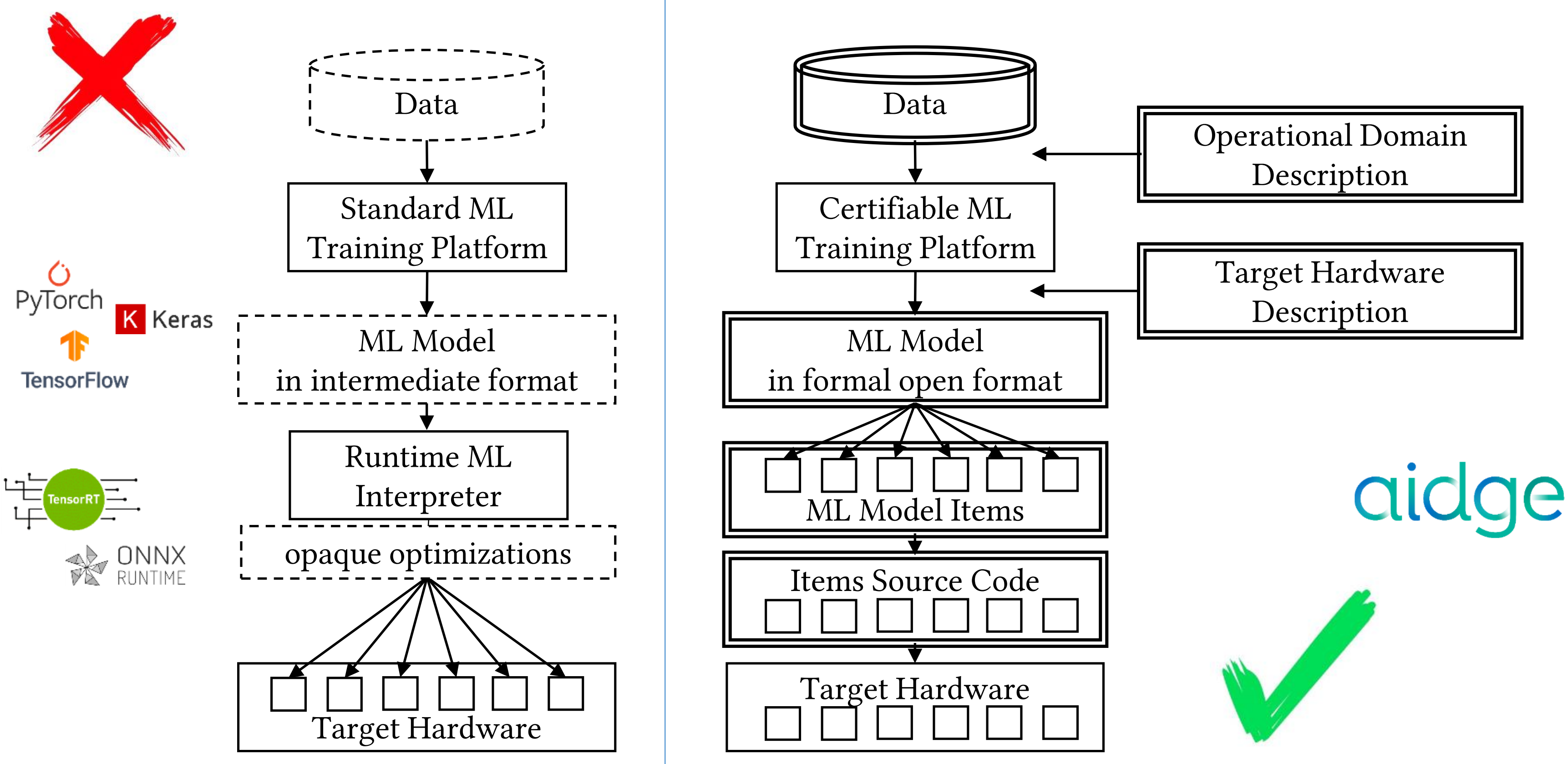
Certification of ML : lightening the black box

ML Tool → **Certifiable ML** ← **Standards and Recommendations**

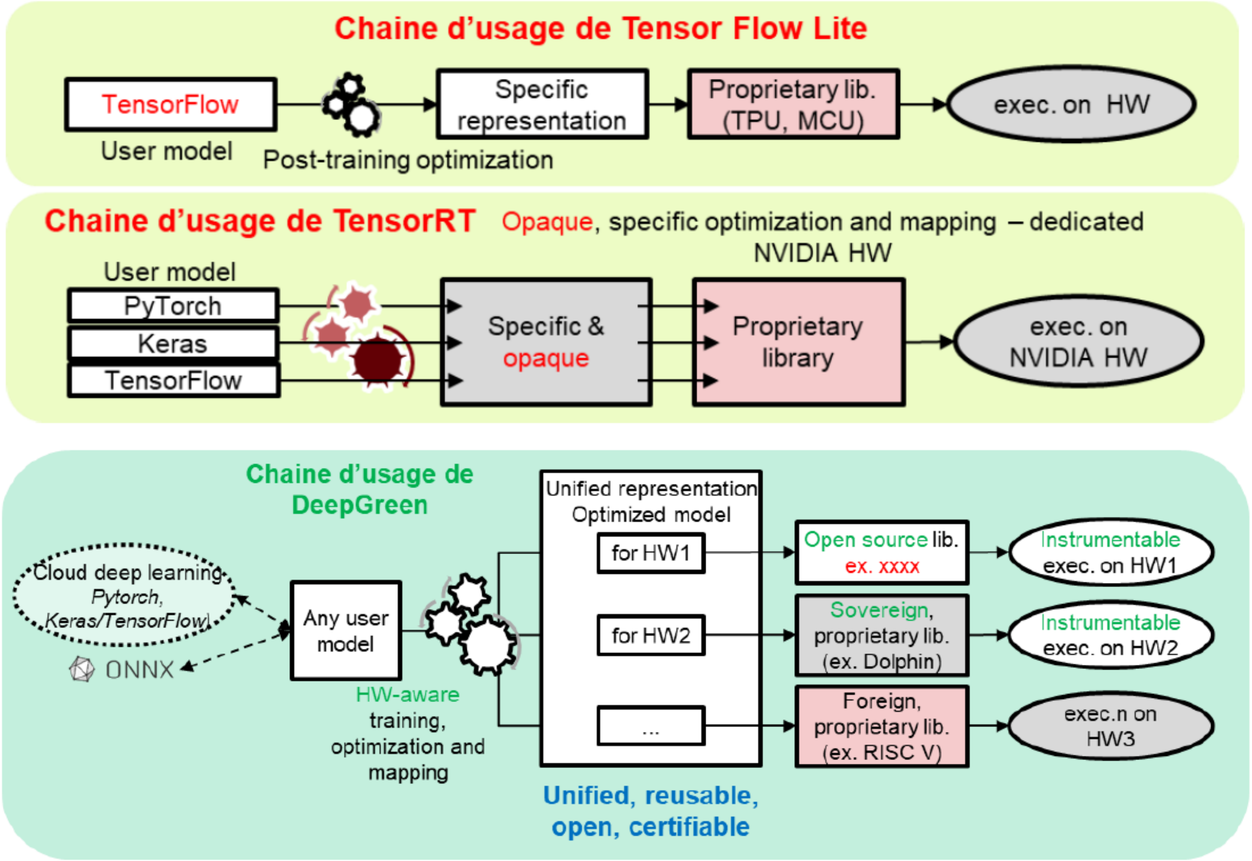
- ✓ can design, train, and optimize neural networks
- ✓ helps to control and follow certification activities
- ✓ can produce evidences for argumentation

- ❑ ARP 6983 – Certification of Aeronautical Safety-Related AI (2025)
- ❑ Usable Guidance for Level 1 (2021) and 2 (2024) ML Applications
- ❑ MLEAP – ML Application Approval (2024)
- ❑ AI Roadmap 1.0 (2020), 2.0 (2023)
- ❑ Towards the engineering of trustworthy AI for critical systems (2022)
- ❑ Concepts for Design Assurance of NN, CoDANN I (2020) et II (2021)

Certifiable ML Platform for Embedded AI



Certifiable ML Platform for Embedded AI



Certification Principles

- **Be compliant**

+

- **Show compliance**

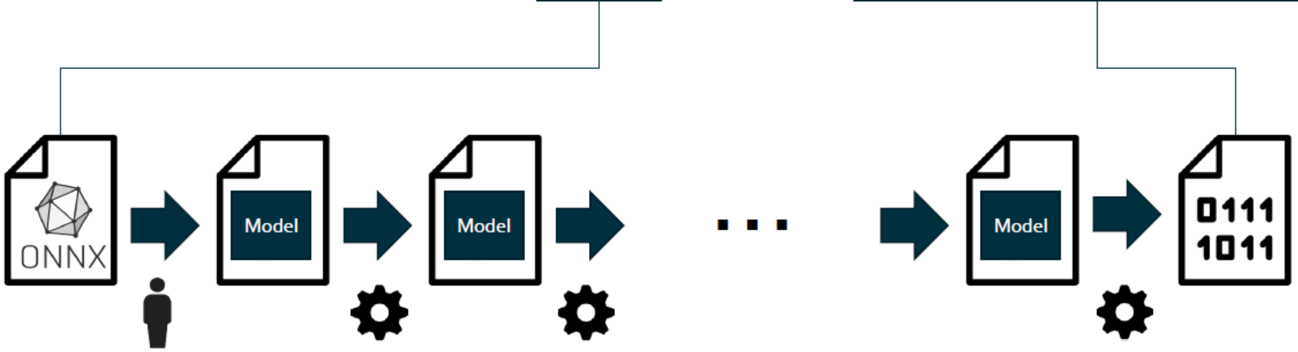
→

- **Traceability**
- **Transparency**
- **Reproducibility**
- **Determinism**
- **Non-Ambiguity**

- **System must be compliant (*be compliant*)**
- **Compliance must be demonstrated (*show compliance*)**
- **Correctness and Robustness must be verified (*tests*)**
- **Assurance of design process (*present development process activities*)**
- **Correlation guarantees: *Model – Code – Binary – Target HW***

Non-Ambiguous Representation : S-ONNX

1. We need to demonstrate that the ML model is correctly implemented on the target



2. To be able to ensure / verify correctness, we need a non-ambiguous description of the model (see later in the presentation)

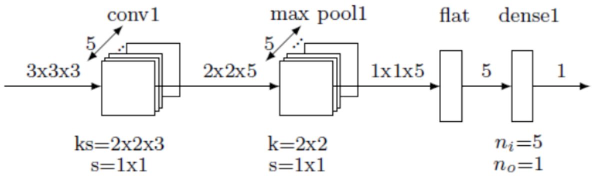
Source: Eric Jenn (IRT-St-Exupéry)

- Each operator associated with a formal definition, e.g. ACSL

Predictable Code Generation

- Traceability between the requirements and the (source) code
- Capacity to estimate or compute tight WCET (worst-case execution time)
- Traceable sequential C code generation from inference model
  - memory layout
  - semantic preservation
- Formal verification
  - ACSL
  - Code C
  - e.g. FramaC

- *Inference function* is **in-lined** and model dependent:



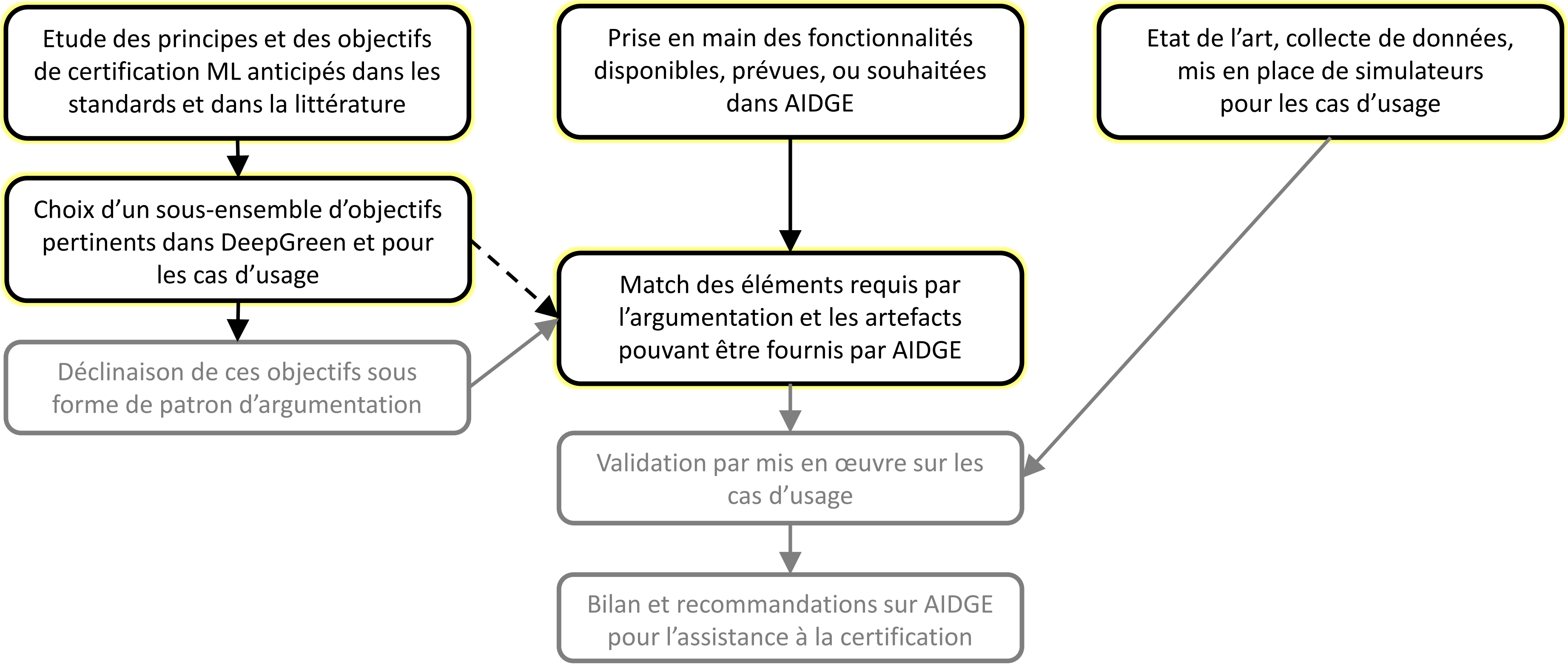

```c
int inference(double prediction[1], double nn_input[27]){
    static double output_pre[27], output_cur[27];
    double dotproduct, sum, max;
    int count;

    // Conv2D_1
    for (int f = 0; f < 5; ++f){
        for (int i = 0; i < 2; ++i){
            for (int j = 0; j < 2; ++j){
            ...

    // MaxPooling2D_2
    for (int c = 0; c < 5; ++c){
        for (int i = 0; i < 1; ++i){
            for (int j = 0; j < 1; ++j){
            ...

    // Dense_3
    for (int i = 0; i < 1; ++i){
        dotproduct = 0;
        for (int j = 0; j < 5; ++j){
        ...
    return 0;}
```

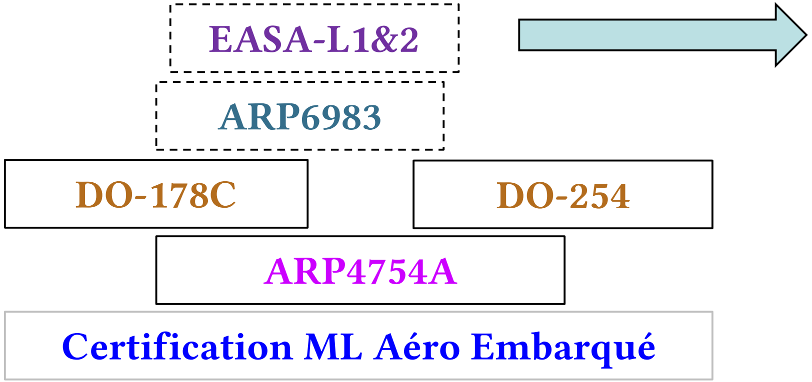
Workpackage Aéronautique – Démarche



Certification

- **Questions fondamentales concernant la certification**
  - **Exposition de la représentation interne du modèle → objet, figure, json, onnx, sonnx ?**
  - **Peut-on certifier une solution en se basant sur des artefacts produits par un logiciel non-qualifié ?**
  - **Certification de code C++ / C, possible intégration d'ACETONE dans AIDGE**

- **Point d'avancement certification**
  - **Analyse de 50 objectifs (sur 130), sélection de 16**
  - **Anticipation d'artefacts productibles par AIDGE**

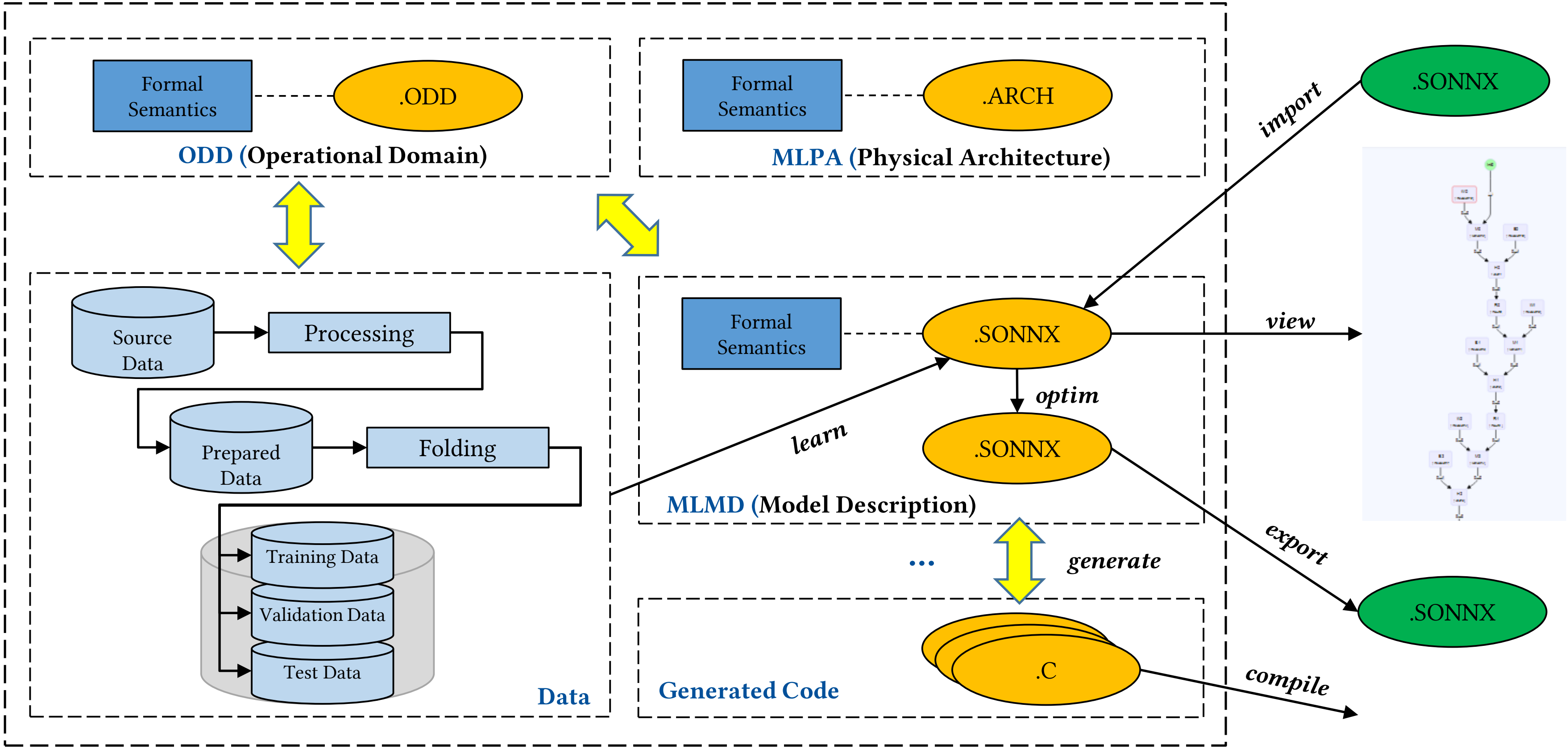


Trusthworthiness	CO - Conception	6	
	CL - AI Level Classification	1	
	SA - Safety Assessment	3	
	ICSA – Instr. for Continuous Safety Assessment	1	
	IS - Information Security Risks	3	
	ET - Ethics	8	
AI Assurance	DA - Development Assurance	10	
	DM - Data Management	**8**	**3**
	LM - Learning Management	**16**	**10**
	IMP - Model Implementation	**12**	**3**
	CM - Configuration Management	1	
	QA - Quality and Process Assurance	1	
	RU - Reusability	3	
	SU - Surrogate Model	2	
	EXP - Explainability	9	
Human-Factors for AI	EXP - Explainability	10	
	HF - Human Factors	34	
Safety Risk Mitigation	SRM - Safety Risk Mitigation	2	
TOTAL		130	**16**

Objectifs de Certification Pré-Sélectionnés

DM-01: Describe all the parameters of the AI/ML constituent **operational design domain (MLC-ODD)** (ranges, nominal data, edge and corner cases, etc.)
DM-02: Describe all the **data quality requirements** (DQR) (data format, accuracy, integrity, completeness, representativeness, independence, …).
DM-03: Describe all **data pre-processing** requirements (pre-processing operations).
LM-01: Describe the AI/ML constituent and the **model architecture** (model structure, layers, operations, …)
LM-02a: Describe **learning management** requirements (selection of family model, learning algorithm, cost/loss function, bias and variance metrics, robustness and stability metrics…)
LM-02b: Describe the **training process parameters** (initialization strategy, hyper-parameters, …)
LM-05a: Describe the **result of the model training** (training curves for the cost/loss functions, error metrics, performance with the validation dataset, …)
LM-05b: Describe the **model** (learned model parameters)
LM-06a: Describe any **post-training model optimization** that affects the model behavior (e.g. pruning, quantization)
LM-06b: Describe the impact of **post-training model optimization** on behavior or performance (comparison with original model on test dataset).
LM-07a: Account for the **bias-variance trade-off** in the model family selection (error comparison on learning and validation datasets, verification with diverse sampling).
LM-07b: Provide evidence of the **reproducibility of the training process** (determinism, complete control over meta-parameters and random generation).
LM-08: Verify that the observed **bias and variance** of the selected model meet the associated **learning** requirements (performance on learning dataset).
LM-09: Verify the **performance** of the trained model based on the **test data set** and document the result of the model verification.
LM-11: Provide an analysis on the **stability of the learning** algorithm (impact of noise, hyper-parameters variation, data shuffle, …).
LM-12: Provide an analysis on the **stability of the trained model**, covering the whole MLC-ODD (against input noise).
LM-15: Describe the **resulting ML model** (resulting MLMD, in a formal structured description).
IMP-04a: Describe any executed **post-training model transformation** (conversion, optimization)
IMP-04b: Describe and validate the impact of **post-training model transformations** on the model behavior and performance
IMP-08: Verify and describe the **performance of the ML component** based on the test data set (deployed inference model verification).
IMP-09: Verify and describe the **stability** of the ML component.

AIDGE project (.aidge) → Dev. Cycle



Thinking the certification process of embedded ML-based aeronautical components

Filipo Studzinski Perotto

Journée commune du GDR RADIA et du GT IE du GDR GPL

21/11/2024